360. revista de alta velocidad

2018 - número 5

INTERNATIONAL CONGRESS ON HIGH-SPEED RAIL (Vol. I)

- Mobile Material
- High Speed Rail Superstructure
- Safety and Signalling Systemas
- High Speed Rail Infrastructure
- Economic Impacts

International Congress on High-speed Rail:
Technologies and Long Term Impacts
Ciudad Real (Spain), October, 2017

Escuela de Ingenieros
de Caminos, Canales y Puertos
de Ciudad Real
UNIVERSIDAD DE CASTILLA-LA MANCHA

FUNDACIÓN DE LOS
FERROCARRILES
ESPAÑOLES

360. revista de alta velocidad

CONSEJO EDITORIAL

EDITOR

Alberto García Álvarez

EDITORES INVITADOS

Santos Sánchez-Cambronero García-Moreno (Profesor E.T.S.I. de Caminos, Canales y Puertos, UCLM)
Ana Rívas Álvarez (Profesora E.T.S.I. de Caminos, Canales y Puertos, UCLM)
José María Coronado Tordesillas (Profesor E.T.S.I. de Caminos, Canales y Puertos, UCLM)

SOBRE LA REVISTA

360. revista de altavelocidad pretende servir de foro de discusión serena y plural, a la vez que profundiza en todos los temas relacionados con la alta velocidad ferroviaria: planificación, efectos económicos y sociales, explotación, tecnología, etc.

El nombre "360.revista de alta velocidad" simboliza la voluntad de "ir más allá" en la aportación social de la alta velocidad (300 km/h es la velocidad máxima actual), y a la vez el deseo de ofrecer una visión panorámica y plural (de 360º de amplitud).

Se articula en tres partes: artículos propios; datos comentados sobre la alta velocidad; y revista de blogs y de prensa, etc., para dar cabida a las opiniones ajenas y ofrecer un termómetro del estado de opinión sobre la alta velocidad.

La revista asume que la velocidad no es un fin, como tampoco lo son las infraestructuras necesarias: el objetivo debe ser el incremento de la sostenibilidad del sistema de transporte y de la eficiencia de la movilidad.

Además se asume que en este campo no hay verdades absolutas ni de validez universal, sino que cada caso debe analizarse individualmente olvidando los apriorismos o ideas preconcebidas.

Los artículos son solicitados a los autores por el Editor (a iniciativa propia o propuesta del Consejo Editorial). Podrán ser publicados en español, inglés o portugués.

Los artículos expresan, exclusivamente, la opinión de sus autores.

Los temas concretos que serán abordados en la publicación; las normas de petición, envío y revisión de los artículos; y las normas de presentación de los mismos podrán ser consultadas en el sitio web de la revista.

Está disponible a través de la página web:

www.360revistadealtavelocidad.es

Edita:
Fundación de los Ferrocarriles Españoles

ISSN: 2174-9655

Depósito legal: M-40482-2011

Coordinación editorial: J. Mariano Rodríguez
Maquetación: Belén Romeo Garre

Foto portada: Patier

FUNDACIÓN DE LOS FERROCARRILES ESPAÑOLES

www.360revistadealtavelocidad.es

Fundación de los Ferrocarriles Españoles

360.revista de alta velocidad

International Congress on High-speed Rail:
Technologies and Long Term Impacts
Ciudad Real (Spain), October, 2017

Escuela de Ingenieros
de Caminos, Canales y Puertos
de Ciudad Real
UNIVERSIDAD DE CASTILLA-LA MANCHA

This Special Issue of the journal joints a set of selected papers presented in the International Congress on High Speed Rail that took place on 4-6 October 2017 at the University of Castilla La Mancha in Ciudad Real to commemorate the 25th anniversary of the Madrid Seville HSR corridor.

The congress plenary and parallel sessions were divided in three days: the first and the third ones in Ciudad Real and the second was held in Sevilla, moving the participants on board of an Special High Speed Train (Miguel de Cervantes) service between Ciudad Real and Sevilla. In particular, the commemorative video that was projected on board of the train can be seen following that link: https://goo.gl/KfH9ZW

Regarding to the objective of the Congress, it was thought to evaluate the usefulness of High-Seed Rail system under a quintuple perspective:

- Long-term implications
- Multidisciplinary scientific approach
- Transport relevance
- Territorial development
- Technology development and diffusion

The social cost and usefulness of the High Speed Rail system has been abundantly studied prior to and shortly after the introduction of this new transport mode. With this fact in mind, this Congress tried to focus on a long-term perspective on its costs and benefits and on a technology development and diffusion perspective. In fact, high speed rail systems have already been in operation for just over half a century in Japan and for 35 years in France. Thus, it was possible to debate its pros and cons considering an important time span.

Finally, this International Congress combined several High Speed Rail system point of view such as superstructure, infrastructure and rolling material optimization, economic, urban and territorial development, sustainability and energy savings, and safety.

INDEX

1. MOBILE MATERIAL

2. HIGH SPEED RAIL SUPERSTRUCTURE

3. SAFETY AND SIGNALLING SYSTEMS

4. HIGH SPEED RAIL INFRASTRUCTURE

5. ECONOMIC IMPACTS

Mobile material

Fundación de los Ferrocarriles Españoles

360.revista de alta velocidad

Mobile material

International Congress on High-speed Rail:

Technologies and Long Term Impacts

Ciudad Real (Spain), October, 2017

Escuela de Caminos
Ciudad Real UCLM

**Escuela de Ingenieros
de Caminos, Canales y Puertos
de Ciudad Real**
UNIVERSIDAD DE CASTILLA-LA MANCHA

Mobile material

360.revista de alta velocidad
número 5 - junio - 2018. Pág 9 - 21

Lightweight primary structures for High-speed railway carbodies

De la Guerra, Eduardo

Talgo[1]

Abstract

The railway industry are looking to increase the capacity of the railway system, bringing flexibility in order to align capacity and demand, to increase availability, reliability and energy efficiency reducing life cycle costs (LCC) and an improvement of the passenger comfort and the attractiveness of the rail transport.

One of the lines of action to solve the above challenges is the introduction of new carbody structures. The weight saving potential of the use of new materials and technologies in the carbody structures would result in reduced power consumption, lower inertia, less track wear and the ability to carry greater payloads.

One example of the employ of light structures is Talgo AVRIL train, with an innovative layout of seats that allows to introduce one extra seat per row (3+2 configuration), increasing dramatically the capacity of the train. It has been possible, respecting the current regulatory framework, with an optimized design of the extruded aluminium structure.

It will be presented different challenges of introducing lighter structures in the carbodies of coaches of high speed train and different studies and projects to describe the future possibilities of new material and methodologies to improve the lightness of the vehicles.

Keywords: lightweight, Rolling stock, carbody.

---

[1] *De la Guerra, Eduardo. Technical Leader of the R&D Projects. Structures Department.Talgo. Email: edelaguerra@talgo.com*

1. Introduction

The railway industry are looking to increase the capacity of the complete railway system, bringing flexibility in order to align capacity and demand, to increase availability, reliability and energy efficiency reducing life cycle costs (LCC) and an improvement of the passenger comfort and the attractiveness of the rail transport.

Different programs inside H2020 framework are working on the development of key technologies to remove already identified blocking points for radical innovation in the field of railway vehicles, as part of a longer term strategy to revolutionise the rolling stock for the future.

These main aims are in line with other transportation sector, like the road transportation (e.g. Superlight Car or ENLIGHT Project) or aerospace (e.g. SESAR, CleanSky…), inside the Roadmap to a single European Transport Area developed by the European Commission to achieve a sustainable transport, also called, White Paper (https://ec.europa.eu/transport/themes/strategies/2011_white_paper_en).

In 2050, it is expected:

- No more conventionally-fuelled cars in cities.
- 40% use of sustainable low carbon fuels in aviation; at least 40% cut in shipping emissions.
- A 50% shift of medium distance intercity passenger and freight journeys from road to rail and waterborne transport.
- All of which will contribute to a 60% cut in transport emissions by the middle of the century.

The third bullet point is a key point for the future of the railway sector, different innovation and new concepts should be developed and implemented to comply with the objective of improved capacity in railway sector of the roadmap.

The amount of passengers that can be transported by the railway system is expected to increase dramatically so all new generation subsystems should be made smaller which will allow a rearrangement of the architecture delivering a better use of space, lighter subsystems (traction, vehicle structure, train control and monitoring system) which will reduce axle load restrictions and will bring more flexibility for the future design of high-capacity and adaptability of vehicle capacity to service conditions.

In addition, new technology development actions targeting energy efficiency and vehicle weight, (Joost, 2012) in a combined manner should result in reductions of energy consumption in operation. The main results expected should be:

- Traction efficiency will increase significantly with new developments in power electronics;
- At the same time, the reduction in vehicle weight derived from lighter structures, lighter traction equipment and elimination of physical equipment will equally support energy consumption reduction;

So it is clear that the new generation of carbodies will be a lightweight carbodies, manufactured in new lighter materials, specially made by hybrid structures (composite and metallic materials) compared with the actual ones made fully metallic (steel or aluminium). This approach allows more passengers per vehicle or more equipment without reaching the maximum allowed axle load reducing power consumption, lower inertia, less track wear and the ability to carry greater payloads.

But as the motto of the White Paper said “The future of our mobility…Today”, all the different bidding process and projects in progress are now claiming for trains with the lowest LCC as possible, maintaining the performance, comfort and bringing high capacity. So now it is designing and manufacturing very optimized carbodies in terms of weight with the current technologies, like the carbody of the AVRIL.

2. State of the art

The function of a carbody is to be the transport passenger container and also the physical link of all the elements of the vehicle.

Historically, passenger coaches were formed by a frame normally made of wood or steel, which received the loads coming from the track and the other coaches, and a cover which had incorporated the doors, windows and gangways (if any).

Progressive improvements were made and self-supporting steel and aluminium carbodies were created in order to reduce mass and improve crashworthiness. In addition, standardised solutions and subassemblies, see Figure 1, have been introduced to become more cost effective.

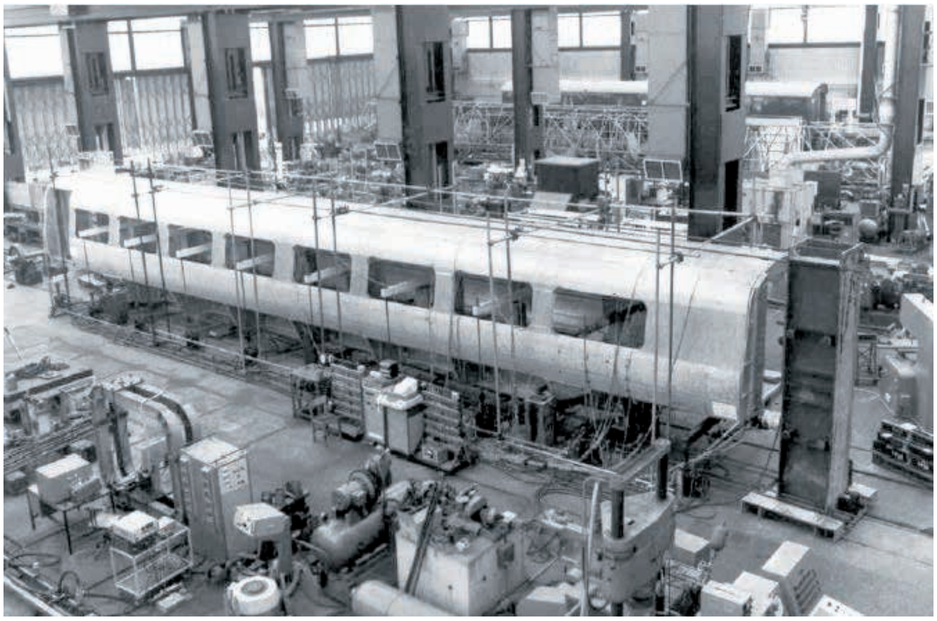


Figure 1. Assembling of the main frame, the side walls and the roof takes place in a positioning and welding station.

Nowadays, the railway carbodies are self-supporting structures made of steel and/or aluminium alloys and usually consist of a set of sheets and profiles joined together by different classical bonding technologies such as welding, riveting or bolting forming a set of great stiffness and currently highly optimized regarding weight per axle, Molinari (2016).

Focusing on aluminium carbodies, which were later developed, but due to the manufacturing process firstly open and then closed extrusions profiles, can improve stiffness and also avoid extra reinforcements (see Figure 2) reducing weight compared to steel carbodies. Over time, aluminium has reached a balanced use compared with steel and is in use in metros, regional and high speed trains in more or less the same ratio than steel.

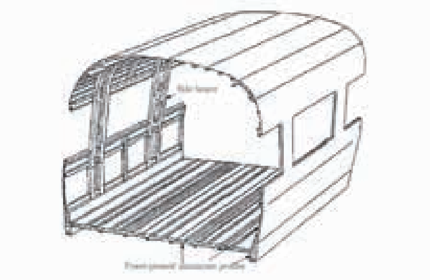


Figure 2. Typical aluminium carbody.

Progress has also been made in steel carbodies, with continuous reductions in steel sheet thicknesses and smarter designs.

2.1 Characteristic of the AVRIL train

Talgo AVRIL (Figure 3) is the most advanced train and the best solution for railway operators with high capacity demands in the Talgo's portfolio. AVRIL (Light Independent Wheel High Speed or "Alta Velocidad Rueda Independiente Ligero" in Spanish) combines maximum speed, low energy consumption and a high capacity with more than 600 passengers in an extremely lightweight single-deck train, trying to answer the needs for the railway industry explained above, Rodríguez (2016).

Figure 3. Talgo AVRIL.

Talgo's design for its trains, which are based on short articulated vehicles together with the Talgo's running gear (rodals or Talgo trucks), allows for the use of lighter carbodies. Talgo trucks use a system with an independent axle per wheel separating the rotation on each wheel.

In addition, in the case of AVRIL offers a greater width for the operator but without exceeding the loading gauge as sketched in the Figure 4. This optimises the train's features since it can seat more people without having to resort to a double-deck layout, which makes boarding and disembarking the train more complicated and brings about an unwelcome feeling for passengers of being boxed-in, mainly in the upper level. Talgo AVRIL guarantees passenger comfort by allowing enough space for each individual passenger without changing the width of the seat.

The ability to carry more passengers improves rolling stock utilisation, allowing operators to offer more competitive prices while decreasing overcrowding on the most popular routes. As a result, operators stand to increase profits and provide a greater return to investors.

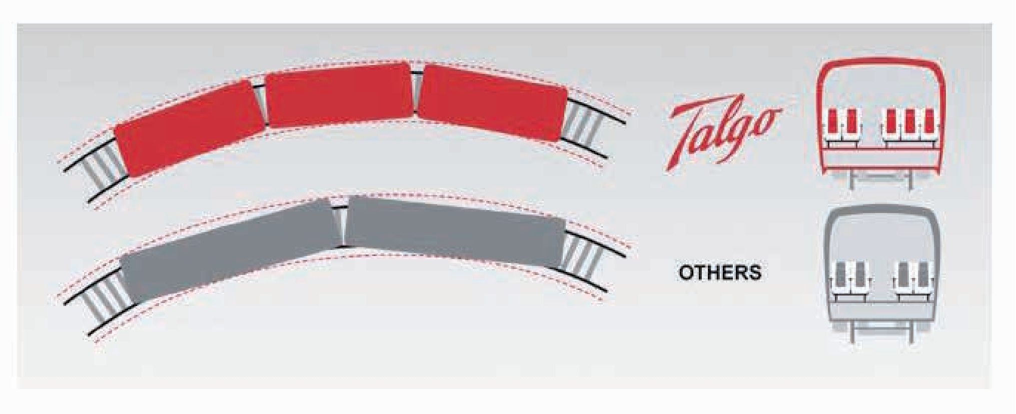


Figure 4. Widebody Talgo AVRIL.

Due to the more space available, it can offer various seating arrangements:

- Customised layout offering several possibilities: 3+2 (see Figure 5), 2+2 or 2+1.
- Seats adapted to all uses: a different configuration is available for all types of passengers: individuals, couples, or groups of three to six people.

Figure 5 Seat layout with 3+2 configuration.

Thanks to the single floor and the low height floor, at platform level as can be seen in Figure 6, AVRIL is the most accessible transport solution. Passengers carrying luggage can get on and off the train without having to use the stairs. Inside they can move along the train without restriction and without steps due to continuous low floor. But the benefits aren't just limited to increased passenger satisfaction, the design of the train means that the time spent at stops is reduced, which helps to optimise the use of rolling stock and infrastructure, as said previously.

Figure 6. Improved accessibility without stairs.

Talgo AVRIL train is designed to be adaptable to the individual needs of every operator. Withmaximum reliability, high interior comfort and low operating costs as baselines, the client is able to decide which solution suits them.

- AVRIL can operate on non-electrified lines by using diesel engines, allowing operators to reach more destinations by simplifying last mile operations.
- The train can be equipped with Talgo's automatic variable gauge system to allow operation on tracks with different gauges.
- AVRIL is compatible with the most common electric voltages used on railway tracks anywhere in the world, including direct and alternating current.
- It can be equipped with any train control or monitoring system, including, ERTMS, TVM, LZB, Indusi, ALSN, ATC, BACC, CONVEL.
- By utilising Talgo's natural tilting system, the train can offer increased speeds on conventional lines which can help to avoid costly infrastructure upgrades.

It is the lightest high-speed train, which means it can offer drastic reductions in operating costs:

- Thanks to its high capacity design and its running gear, it is possible to optimize the AVRIL's energy consumption compared to other high-speed train and previous series.
- On top of that, the design of its components and the ability to use generic systems reduces maintenance costs and the amount of time the train has to spend in the workshop.

In Table 1 is collected the main features of AVRIL (https://www.talgo.com/en/rolling-stock/very-high-speed/avril/).

Table 1. Main technical features	
Maximum speed	365 km/h
Operating speed	330 km/h
Passenger capacity	ca. 500 - 600
Trainset configuration	Typically two powerhead and 12 coaches but adaptable to client specifications
Track gauge (mm)	1435-1668
Floor height (mm)	760 above TOR
Length (m)	201.9
Width (mm)	3200
Tare weight (tonnes)	317
Number of axles	21
Passengers/WC	40-55
Power supply	25kV AC (tri-voltage optional)
Power	2 x 4400 kW (tri-voltage optional)
Brakes	Powerhead: regenerative, rheostatic and air brakes Coaches: air brakes
Standard	UIC/ETI

3. Challenges of new lighter structures

All the above requirements imposed to the new trains usually imply the reduction of the weight of the primary structure in order to accommodate new payloads, be more efficient, etc.

The design of the lighter structures that can withstand more than the previous one, usually supposes challenges when the structure is analysed in term of stiffness (natural frequencies, damping ratios and mode shapes) and noise isolation. Lighter primary structures under more restrictive loads (without changing concept design) are prone to have low natural frequencies that imply poorer ride comfort values for passengers, (Dumitriu, 2017; Chatti, 2006).

In general, the running gear interacts with the track producing vibrations with different amplitudes and frequencies depending on the line, tracks, defects or maintenance program. The more speed, the more amplitudes and higher frequencies levels (see Figure 7). To avoid passenger discomfort, the carbody should be designed with sufficient stiffness in the vertical, lateral and torsional directions to avoid coupling of frequencies between the running gear and carbody, both should not display similar Eigen frequencies. In addition the interfaces between bogie and carbody should be carefully analysed in order to avoid undesired bypasses of forces through them.

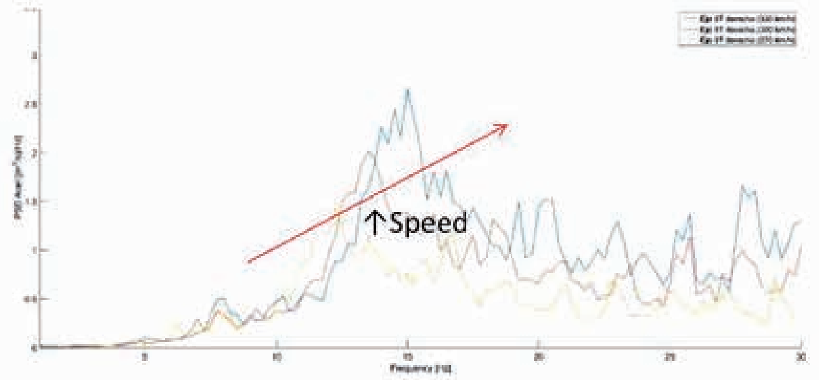


Figure 7. PSD of acceleration at unsprung mass level depending on speed for HS track Madrid-Barcelona.

Anderson et. al (2005, 2007) suggested values with some typical carbody Eigen frequencies fora high speed train, Table 2, and as can be seen are very closed to the maximum excitation produced by the track.

Table 2. Typical Eigen frequencies

Mode direction	Frequency (first mode)
Vertical	9-10 Hz
Lateral	10-11 Hz
Torsional	11-12 Hz

In term of strength, rail vehicle design criteria in Europe are based mainly in EN 12663-1 if no other specific reference is given, usually applied category P-II for High-Speed passenger vehicles, Wennberg (2010).

The stresses to which carbodies are subjected are of various types and defined in EN 12663-1:

- The stresses due to longitudinal, vertical and extraordinary forces caused by the normal movement of trains.
- Also, there are the stresses of vibration of the carbody itself, which occur by the effect of its mass and dynamic loads from the track and its effect on fatigue.
- Finally, stresses due to accidental collisions related with passive safety.

Furthermore, for high speed trains, it is also necessary to consider overpressure/underpressure load cases due to the crossings with other trains particularly in tunnels, and stresses caused by lateral winds.

Regarding carbody vibration behaviour, in the standard EN 12663-1 is only stated that the natural modes of vibration of the carbody should be separated sufficiently, or otherwise decoupled, from the bogie suspension frequencies, so as to avoid the occurrence of undesirable responses and to achieve an acceptable ride quality according EN 12299. The fundamental Eigen frequencies of the carbody in lateral and vertical directions should be above NHz. Commonly the value of N is given by the fundamental resonance frequencies of the bogie frame and other influencing factors like the track or the passenger load.

As a general rule, the natural frequencies of the vertical mode for a typical bogie with the carbody installed are in the range f0,bf=6-8Hz. Therefore, the frequency separation requires that the fundamental carbody Eigen frequency complies f0,cb>√2·f0,bfHz. With this values, it is important to note that in some cases the ride quality according to EN 12299 could not be achieved, due to possible peaks of frequencies that could appear on the track in the range of 10-15Hz (depending on the track default, as can be seen in Figure 7). Then, the stiffness of the carbody must be optimised to increase as possible the corresponding carbody Eigen frequencies o achieve the comfort requirements, i.e. as higher as possible, de la Guerra (2016).

Regarding the interface between bogie and carbody, as rule of thumb the input mobility of mounting points for bogie elements (dampers, rods) shall be below -80dB re 1m/Ns. Alternatively, the input impedance of mounting points for bogie elements (dampers, rods) shall be above: 80dB re 1Ns/m, Shabalin (2013).

4. Numerical and Experimental Analysis for design

The use of new methodologies for design (FEM+Multibody Simulation) together with of operational/experimental modal analysis allows the understanding of the behaviour of the new structures regarding vibration and comfort. In addition, it is also checked the effectiveness of the general rules explained above applied in a very specific product like Talgo's trains (non-standard running gear, short, wide and light vehicle).

The structures of the AVRIL's coaches are made by an optimum combination of open and closed profiles with variable inertia to achieve the strength of the carbody. In order to improve the stiffness of the carbody stiffeners and reinforcements inside and outside are welded to the main structure to achieve an assembly very stiff and light. With the new profile architecture is possible to have a bigger structure (mainly wider) saving around 10% in weight compared to the structure of the previous high speed Talgo's coaches (S102/112).

First step is performed a complete modal analysis with a finite element program to estimate the natural frequencies and mode shapes. Mainly we are focused on the first mode shape, frequency and if it is a local mode or a mode for the complete structure (modal mass). The analysis is done in different mass condition (full and empty).

For example in the Figure 8 is shown the 2nd flexural mode of the main frame of the prototypcarbody, in this case at 20Hz.

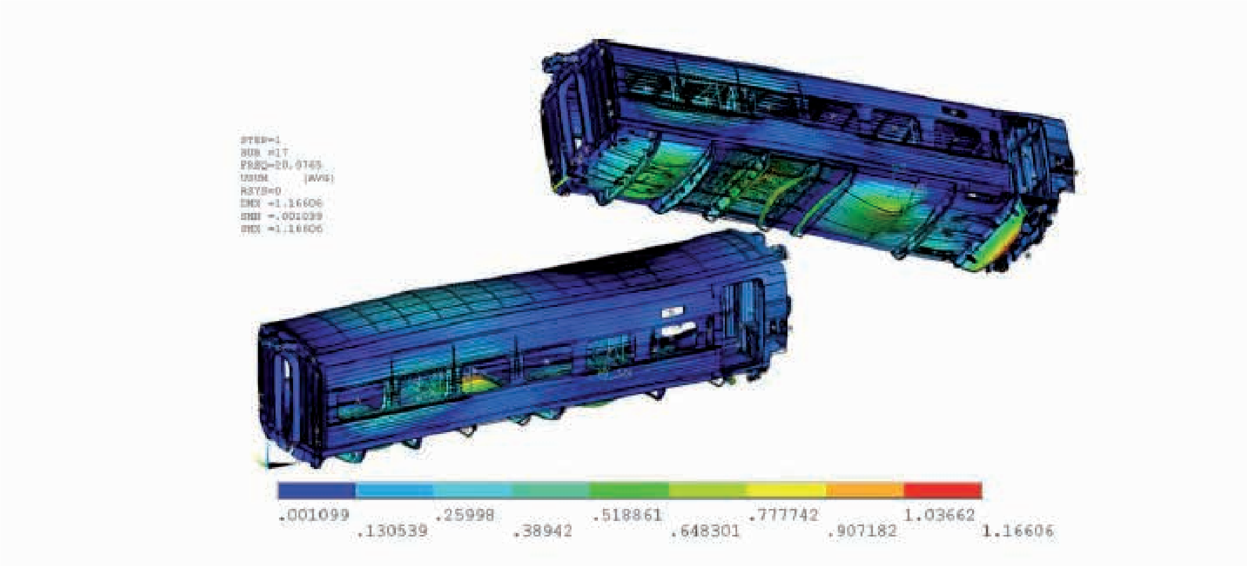


Figure 8. Modal FEA. 2nd flexural mode at 20Hz of the main frame of the prototype carbody

This approach is very accurate when the Eigen frequencies are far from the running gear excitation. For this reason, together with modal FEA, a combined analysis with multibody and finite element is done to assess the separation between the Eigen frequencies of the running gear at train level and the flexible carbody. It is very important when you have a lot of interfaces with the carbody like in the Talgo case: guidance system, secondary suspension, dampers.

In addition to computer design different analysis was made in the prototype carbody. Operation modal analysis (OMA) and Experimental modal analysis (EMA) would be a useful tool to obtain the real mode shape and frequencies. OMA is a modal analysis using operating conditions as excitation in contrast to EMA that uses defined and controlled excitation.

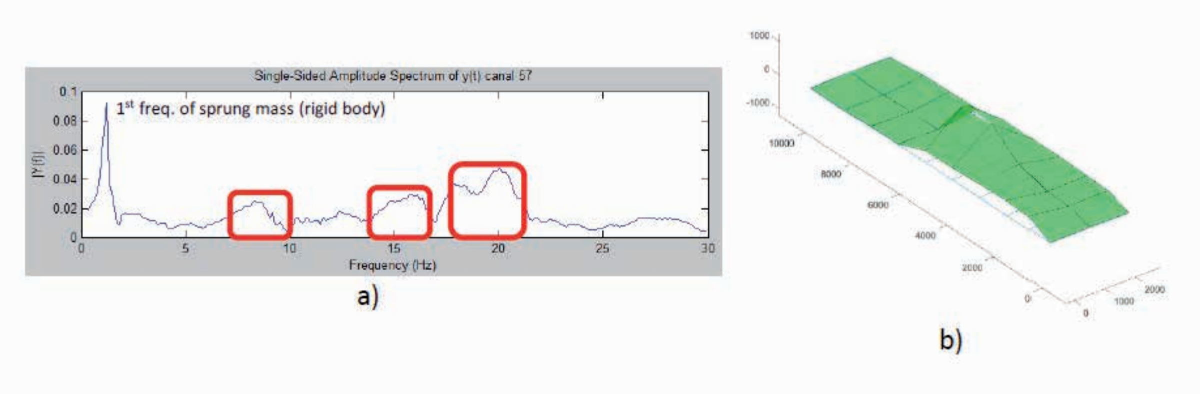


Figure 9. Eigen frequencies measured over the main frame of the carbody

In the Figure 9 a different accelerometer signal of the carbody can be seen, three Eigen frequencies groups can be distinguished, 8-10Hz around 15Hz and a third group in the 20Hz environment. With a more detail analysis is possible to extract shape modes (Figure 9b) and detect each frequency in the groups.

The occurrence of frequencies in the vicinity of 8-10Hz is not obtained during the FE calculations but appear due to the guidance system excitation as checked with the testing and Multibody simulation.

5. R&D Projects

In addition to the customer project, different R&D projects are in progress regarding lightening of the structures, Figure 10.

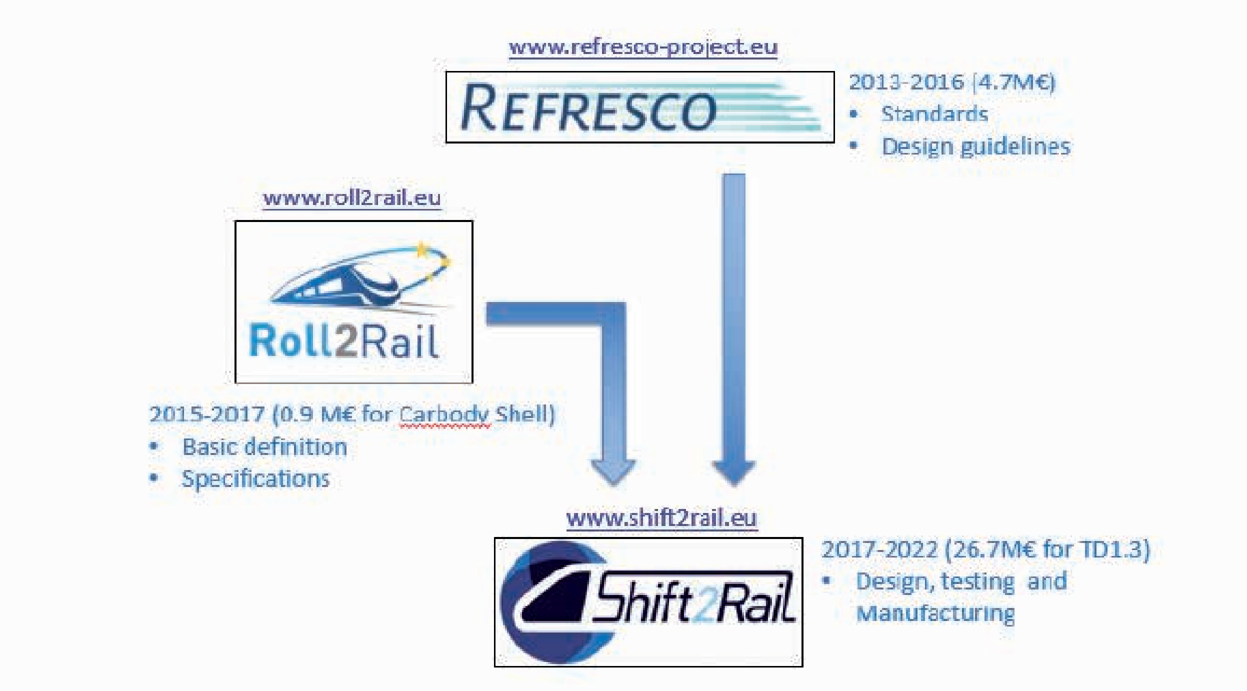


Figure 10. R&D Project map

In these case an approach completely different introducing a complete new concept and new materials for the primary structures in railway sector, Peris (2017).

The goal of the work carried out as part of Shift2Rail is to develop lighter carbody shells which make full use of the possibilities of composite materials including integration of functions. This is linked to identifying the specific design principles, materials and manufacturing processes that fulfil the requirements set in previous projects such as REFRESCO and Roll2Rail, in terms of material properties, manufacturing cost and certification. To this end this plan envisages to develop a carbody made of hybrid materials (mainly composites) in the primary structure that will achieve:

- Between 15 and 30% weight reduction.
- Energy savings in operation, resulting from the weight reduction.
- Improvements in maintainability, coming from new concepts.
- The introduction of a specific health monitoring process.

With today's materials and manufacturing technologies, manufacturers of rolling stock are very close to the border in terms of weight optimization (Molinari, 2016), so a change of concept in the structure of the car is imposed in order to reduce the weight of the structure primary. This objective is aligned with the objectives set in the H2020 to reduce the life cycle cost (LCC or Life Cycle Cost) globally of rail transport and to increase the capacity of the rail sector.

The use of material different from metal has been proved in the aeronautic industry, where composites are increasingly being used in structural parts after having passed all tests regarding safety.

In conclusion lighter carbodies could be made with industrial processes, provided that adequate joint methods are used and there is compliance with rail safety standards.

Different conceptual studies have been developed based on topology optimisation and structural/ acoustical calculation of materials and joints in order to lay the foundation for the designing phase of Shift2Rail.

An example of the topology optimisation is shown in figure 2. The influence of the geometry and location of the cut-outs in the weight of the structure was studied. For an urban model, different configurations were analysed achieving a weight improvement of up to 20% in the case of decreasing the width of the door 300mm (-15% in width). For a High Speed model, up to a 14% weight reduction was achieved when the service door is placed at the centre of the coach.

At the end of Shif2Rail project a full carbody demonstrator with a hybrid carbody will be presented fulfilling the requirement of weight reduction mentioned before maintaining or improving the current performance of the metallic primary structures in terms of safety, durability and maintainability.

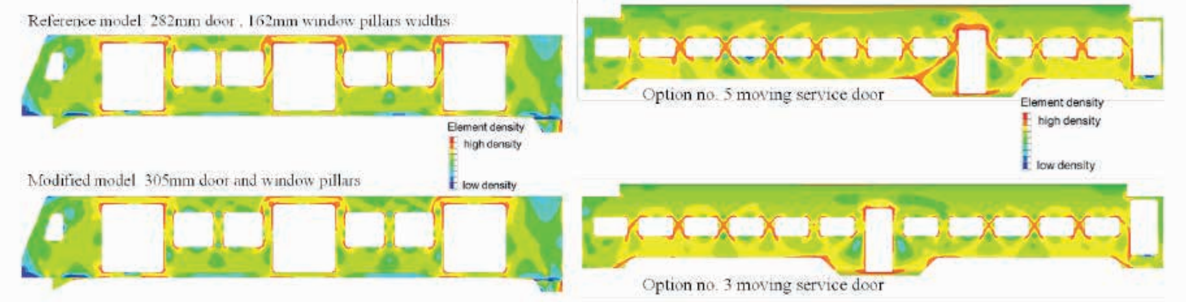


Figure 11. Topology optimisation for urban trains (left) and high speed trains (right)

6. Conclusions

The weight reduction and optimization of railway vehicles have several positive effects including energy savings, increase payload capacity and reduction of rail damage.

However in order to maintain the performance of the current carbodies regarding comfort some extra cautions should be taken.

A study in detail like done in AVRIL train is necessary to avoid discomfort for the passenger at high speed. FEM calculation together with Multibody simulation are powerful tool for design phases. Prototypes to assess the hypothesis employed are very useful to measure and obtain real mode shapes and Eigen frequencies to compare with the track excitation.

7. References

- Andersson E., Berg M., Stichel S. (2005) Rail Vehicle Dynamics. Railway Group KTH, Div. of Railway Technology, Dep. of Aeronautical & Vehicle Eng, the Royal Institute of Technology.
- Andersson E., Berg M.(2007) Spårtrafiksystem och spårfordon, Del 2: Spårfordon. Railway Group KTH, Div. of Railway Technology, Dep. of Aeronautical & Vehicle Eng, the Royal Institute of Technology.
- Chatti S. (2006). Production of profiles for lightweight structures. BoD–Books on Demand.
- De la Guerra, E. et al. (2016) Car bodyshell specification, Deliverable 3.1 Roll2Rail Project.
- Dumitriu M., Cruceanu C. (2017) Approaches for reducing structural vibration of the carbody railway vehicles. MATEC Web Conf. 112 07006 DOI: 10.1051/matecconf/201711207006
- Joost WJ. (2012) Reducing Vehicle Weight and Improving U.S. Energy Efficiency Using Integrated Computational Materials Engineering. JOM, Vol. 64, No. 9, DOI: 10.1007/s11837-012-0424-z.
- Molinari. (2016). Optimization - Light weight design for rolling stock vehicles. http://www.molinari-rail.com/wp-content/uploads/2016/06/optimization-en.pdf
- Peris E., de la Guerra E., Ramirez JM. (2017) Lightweight structural materials for railway carbodies. Compositi Magizine no. 43 (XII)
- Rodríguez A. (2016) AVRIL de Talgo, nuevo tren de alta velocidad de Renfe. Vía Libre Separata del no. 613.
- Shabalin D. (2013). Theoretical and experimental comparison of point mobility measurement methods. Master thesis Aalborg Universitet, Brüel&Kjaer. Mobility Measurements. Application notes
- Wennberg, D. (2010). A light weight car body for high speed trains: Literature study. KTH Royal Institute of Technology.
- EN 12299 (2010). Railway applications - Ride comfort for passengers - Measurement and evaluation.
- EN 12663-1 (2011). Railway applications - Structural requirements of railway vehicle bodies - Part 1: Locomotives and passenger rolling stock (and alternative method for freight wagons).
- https://ec.europa.eu/transport/themes/strategies/2011_white_paper_en the Roadmap to a single European Transport Area, White Paper EC.

- http://www.roll2rail.eu/ Official Webpage of Roll2Rail Project. H2020 Program Grant nº: 636032.
- http://www.shift2rail.org/ Official Webpage of Shift2Rail Project.
- http://www.refresco-project.eu/ Official Webpage of REFRESCO Project. FP7 Program Grant nº: 605632.
- https://www.talgo.com/en/rolling-stock/very-high-speed/avril/Official Webpage of Talgo AVRIL

Mobile material

360.revista de alta velocidad
número 5 - junio - 2018. Pág 23 - 31

New technologies applied to railway maintenance

Coca, Marcos

Alstom[1]

Abstract

Rarely in the history of rail maintenance has there been such a clear opportunity to introduce new technology and to leverage information and communication systems aiming to search for improvements in equipment maintenance, infrastructure or in maintenance operations.

It is currently possible to find-out in the market a wide range of monitoring devices, interfaces or communication systems configurable at a moderate cost, as well as computing technology with sufficient processing capacity to diagnose in real time and even anticipate the future behavior of the systems.

This initially favorable situation also poses two major challenges: the new *on-hand* digital solutions will not bring the desired added-value without suitable integration and probable adaption to the existing business processes and tools, which is not usually easy according to the new dimension of the project and the greater uncertainty in the outcome. On the other hand the wider diagnostic and interconnection possibilities of digital systems, usually able to provide huge volumes of data, make the identification of key parameters difficult enough to draw effective conclusions in terms of maintenance optimization, thus data structuring, data filtering and prediction algorithms play a fundamental role.

Alstom, with its global experience in railway maintenance and more than 25 years in maintaining Very High Speed Trains, will show few examples of new technologies brought to equipment maintenance and operational management, as well as those difficulties and relevant findings in this process.

Keywords: optimal railway maintenance, augmented reality, predictive plataforms, drones, big data, machine learning, real time equipement diagnostics.

---

[1] *Coca, Marcos. Alstom. Email: marcos.coca-sola@alstomgroup.com*

Since its very beginning, railway maintenance has evolved at a pace marked by technology innovations integrated in rail and infrastructure equipment as well as the set of methods and tools acquired by the industry. However just in the third quarter of the twentieth century electronic and communication systems burst into the railway industry to accompany and mark the future of this activity.

An important milestone which is part of this Third Industrial Revolution or era of automation is the incorporation of digital systems. The leap from analog to digital control systems, apart from improving accuracy and response times in the regulation of systems, significantly increased the diagnostic capacity by progressively developing controllers and interfaces capable of managing larger volumes of information per unit-time and to extend the scope of connected devices at even increasingly competitive costs of acquisition and operation. Even though it was a real source of improvement for railway maintenance execution, the processing and communication speeds already achieved could not lead to real-time data sharing and remote systems diagnosis.

At this time we navigate within the Fourth Industrial Revolution [Industry 4.0 or Connected Industry], reaching outstanding levels of computing power, intelligence and development of communication systems, including wireless, which are allowing to an unprecedented qualitative leap in the execution of maintenance operations and in the organization of the railway maintenance activity.

Opportunities are extraordinary: from planning the execution of the maintenance activity in advance after predicting the future behavior of the systems, to considering the results of previous decisions and feed the process through Machine Learning and Artificial Intelligence, and to detect new business opportunities or optimization possibilities benefitting our clients or the company itself from massive data analysis [Big Data].

But the advantages also bring important challenges to be mastered so that the introduction of the new technology adds real added-value to the maintenance activity and does not become a mere digital showcase. The vertiginous technological development puts on the table an enormous volume of data provided by the equipment that needs to be interpreted, filtered andtransformed into useful information for decision making, and which in turn encourages the continuous appearance of data analysis tools which must be observed. On the other hand, this scenario provides complementary information to the maintainer (service demand, infrastructure status, environmental parameters, etc.) that may imply new opportunities or changes in maintenance management models. This requires adequate skills and competencies: from the global vision necessary to direct investments in digital transformation, to the ability to interpret the technical and operational data available to transform them into useful information for business decisions.

Currently more than 50% of the companies adopting Internet of Things are not sure of the return of their investment. New technologies are routinely evaluated and deployed in such a fragmented way that it is not possible to assess in advance the joint effect of their integration. 27% of companies are not sure of the questions they should ask around data, and 31% do not store the information that is generated. This indicates that the level of maturity is not yet high and we are only at the beginning of the learning curve.

How the organization adapts itself to this new scenario, how information is shared, how resources are organized, and how processes are designed to make operational decisions reliable and efficient enough will be key to the success of the system, and in any case the possible solution to the challenges posed.

1. Technological offer

The market is currently able to provide technological gadgets to particular users or companies combining remarkable processing capacity and communication interfaces compatible with the main standards for domestic and industrial communication.

Many of these gadgets will soon incorporate communication interfaces allowing to operate in bands compatible with 5G technology, therefore they will multiply the communication speed by 50 to 100 times, and will be considered fully connected devices. The future of railway maintenance will go through the proper integration and use of these devices.

Among the multiple angles in which one can approach to railway digitization, one could be the expected usage in maintenance operations. In this way we can have:

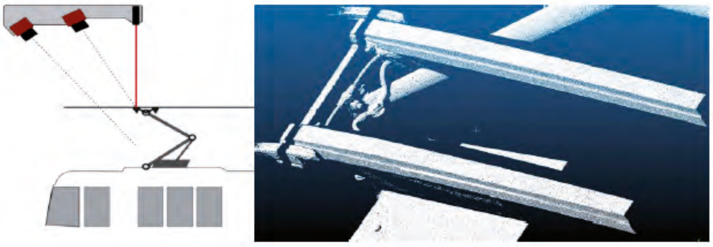
- Asset monitoring and diagnosticssystems
- Maintenance supportdevices
- Integration Platforms, data processing and interface with business managementsystems

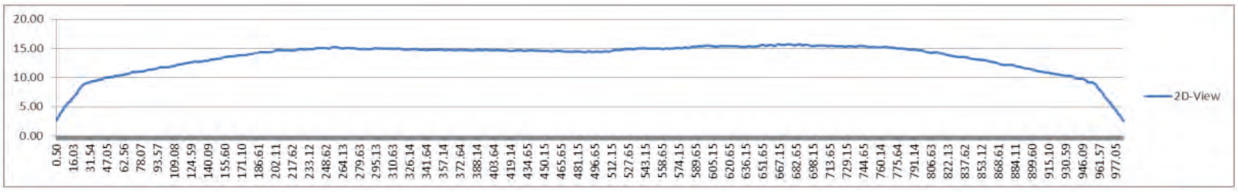
Asset monitoring and diagnostics systemsprovide to the maintainer built-in technology or later integration possibilities, as well as monitoring solutions that minimize physical interventions in the asset. All of them have the objective of continuously and reliably retrieve key parameters to determine or to infer the equipment performance.

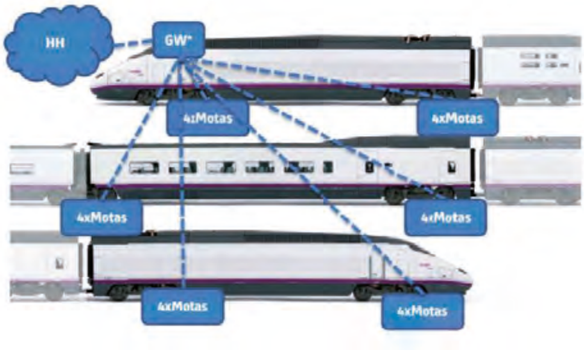
The expected performance of these systems is primarily measured in terms of reliability and quality to monitor key parameters, which can be physical (distance, temperature, pressure, vibration, voltage, current) or presence and usage (number of actions) and in terms of the required time for data transmission. Therefore, reliability, precision and communication performances in relation to investment and operating costs will be relevant factors to consider in your selection.

Some examples of monitoring technology are optical and laser systems to perform dimensional checks, thermographic technology, inductive, electrical or optical systems to confirm the presence or position of an element, number of maneuvers or displacement, instrumentation for measurement pressure, etc.

These systems currently allow, for example, temperature and vibration monitoring at grease-box bearings even in conventional and high-speed vehicles. They also provide means to measure and to identify degradation patterns in elements subjected to wearing such as pantograph contact strips or the brake disks, or to monitor interlocking control and driving systems (point machines) through electrical parameters.







Technology evolution is in turn providing monitoring devices with the necessary intelligence to produce and transmit in real time only the relevant diagnostic information, thus reducing the frequency and volume of raw data transmission to optimize the battery autonomy when they lack external power. In turn they allow to download, if required, the full-recorded data postponed in time.

Devices or systems supporting maintenance execution can be directly used as a complement or replacement of maintenance activities, either to improve the safety, reliability or ergonomics of operations, or to enhance the efficiency and optimize the costs of the intervention by automating the process.

They are in fact the most popular gadgets in the Connected Industry or IIoT since many of them find eventual applications in maintenance or repair workshops and could be part of what is usually designated as Workshop of the Future.

These devices include Robots and Cobots -Collaborative robots- Additive Manufacturing and 3D Scan, Augmented and Virtual Reality (AR / VR), Connected Tools, Wearables or smart technology integrated in clothing, Drones, Smart Glasses, etc.

The main value of these tools is the possibility of improving the process efficiency and safety by automating labor-allocated tasks, as well as the possibility to reinforce both process control and process quality.


Robots


Pit Inspection Robots


3 D printing


Big data


Automated storage


Augmented reality


Connect lights


Surveillance Robot


Motion capture


Drones


Connected tools


Wearables

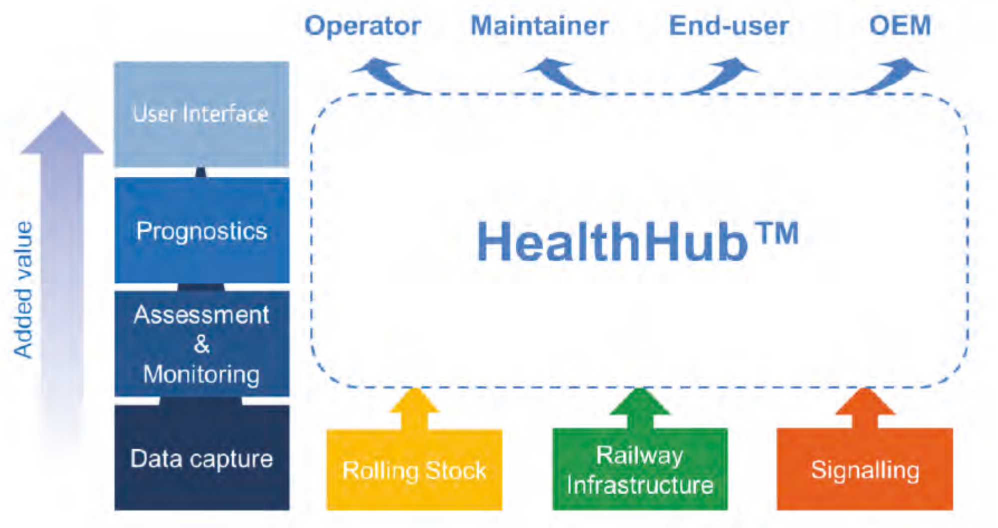
Data processing, integration and interface platformsare the nexus in the Connected Industry leveraging integrated systems and corporate applications to provide real added value to maintenance activities.

Large technology providers, communications and industrial services are developing data integration platforms and associated Cloud Computing services, Reporting, Security Management, intending to concentrate activity and to become an industrial benchmark. In parallel Open Source platforms are also emerging intending to provide roughly equivalentservices.

On the other hand, data transmission protocols are also being adapted to the Connected Industry, for example Sigfox, LoRa standards improving the efficiency and optimizing the energy consumption of the interface.

Integration Platforms are the core of Big Data, setting-up the basic structure to provide integrity and analysis potential to exchanged data, from which it is possible to launch algorithms and processing routines to extract the relevant information.

In the field of railway maintenance, this basic data structuring is even more important because of the great heterogeneity in the sources (customer, infrastructure, vehicles, maintenance) and in the information provided, sometimes lacking the minimum required reliability to subject to analysis.



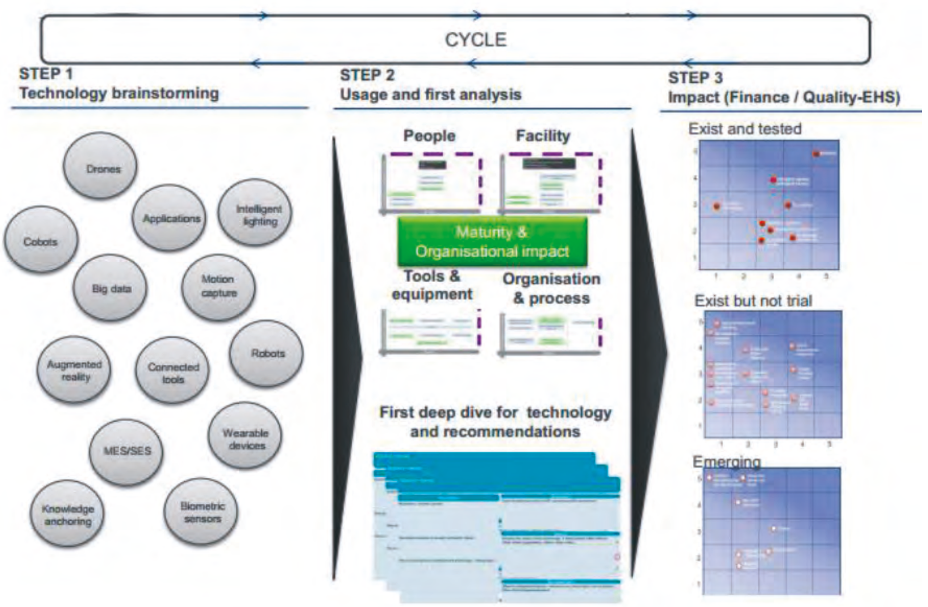
Decisions surrounding the selection and management of these platforms are complex and due to the current dynamism that they must be supported by work groups with adequate vision and technical skills.

2. Adoption strategy

Defining a clear strategy for introducing new technology, providing it with a plan and specific resources for its achievement and involving all levels of the organization are key factors to secure the integration project.

To this end, it may be appropriate to carry out an initial evaluation of the possibilities of the new technology to solve those relevant problems and inefficiencies noticed by the organization together with the required changes or adaptations. On the other hand, segmenting new tools according to their state of maturity and risk for evolution, together with the ratio [outcome quality -cost of the investment], will provide better insight into the possibilities of technological devices and into the associated risks.

Raising the introduction of Connected Tools in a rail maintenance workshop would require assessing how the new technology will bring value and will help to optimize the process together with an estimation of the expected return, likewise identifying the adjustments or redesign requirements in the maintenance process, staff training needs and logistics support to get the tool fully operational whenever it is required by the maintainer.



The action plan derived from the integration strategy must be ambitious and shall consider all the optimization possibilities offered by the new technology, but at the same time it should be developed in continuous and small steps, without stopping.

It is common practice to develop Proofs of Concept in the early stages of that process to confirm that the performance of the new devices is in line with expectations and the investment is justified. They are usually small-scale tests intending to get closer to the current operating conditions. But the outcome may not be conclusive or could be distorted if testing conditions are not representative, or if the device is forced to integration into the existing process rather than into a redesigned one.

3. The management of the new process will determine the outcome

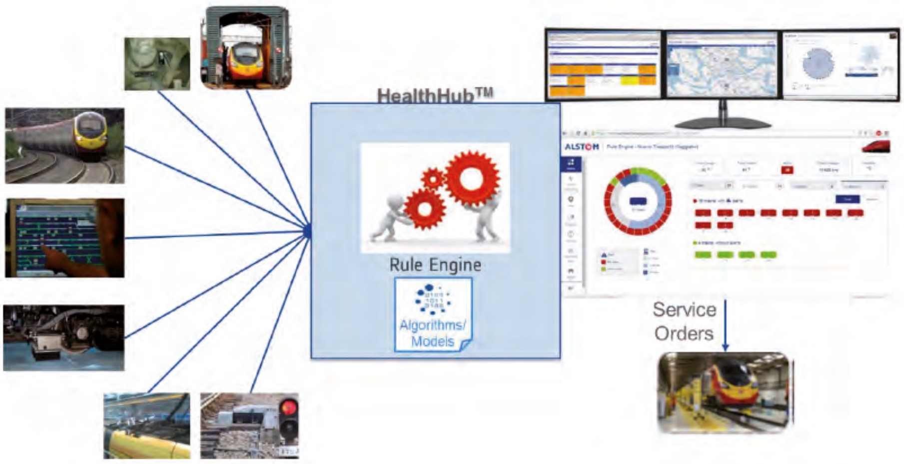
Although the integration plans are bult in small steps, the strategy should be conceived over the final management vision of processes and tools required in the maintenance activity, including in-between synergies and the need for integrating platforms and data processing tools.

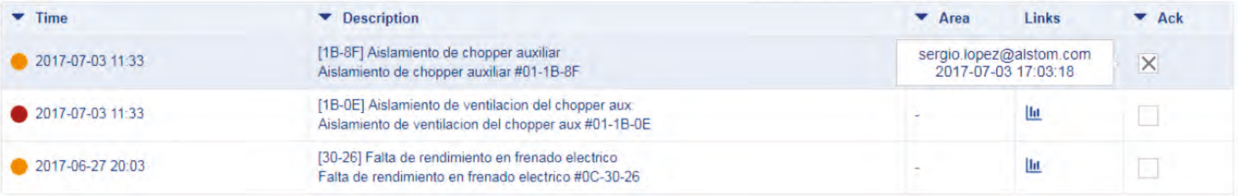
An example of this vision is found over the use of asset monitoring systems and the development of Predictive Maintenance algorithms.

Scheduled preventive maintenance task are being progressively relieved by Condition Based Maintenance (CBM), insofar as it has been possible to access to new reliable diagnostic information for railway equipment. Currently technology allows going a step further, applying Predictive Maintenance techniques in order to model and to predict the behavior of the system till failure, and even applying Proactive Maintenance to detect and act in advance in the causes that generate degradation performance of the component or equipment.

In this way, it would be possible to dynamically plan maintenance tasks, to avoid unexpected Service-Affecting Failures and to optimize maintenance intervals, maximizing the reliability and availability of the fleet.

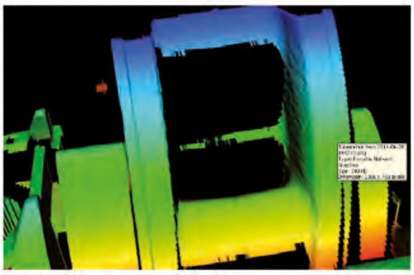
As a result, it may be eventually required to modify the maintenance process to reconsider the execution of corrective tasks in order to keep the vehicle in operation as per the joint evaluation of the predictive model and the real-time reported alerts and events.





Time	Description	Area	Links	Ack
● 2017-07-03 11:33	[1B-8F] Aislamiento de chopper auxiliar Aislamiento de chopper auxiliar #01-1B-8F	sergio.lopez@alstom.com 2017-07-03 17:03:18		☒
● 2017-07-03 11:33	[1B-0E] Aislamiento de ventilacion del chopper aux Aislamiento de ventilacion del chopper aux #01-1B-0E	-		☐
● 2017-06-27 20:03	[30-26] Falta de rendimiento en frenado electrico Falta de rendimiento en frenado electrico #0C-30-26	-		☐

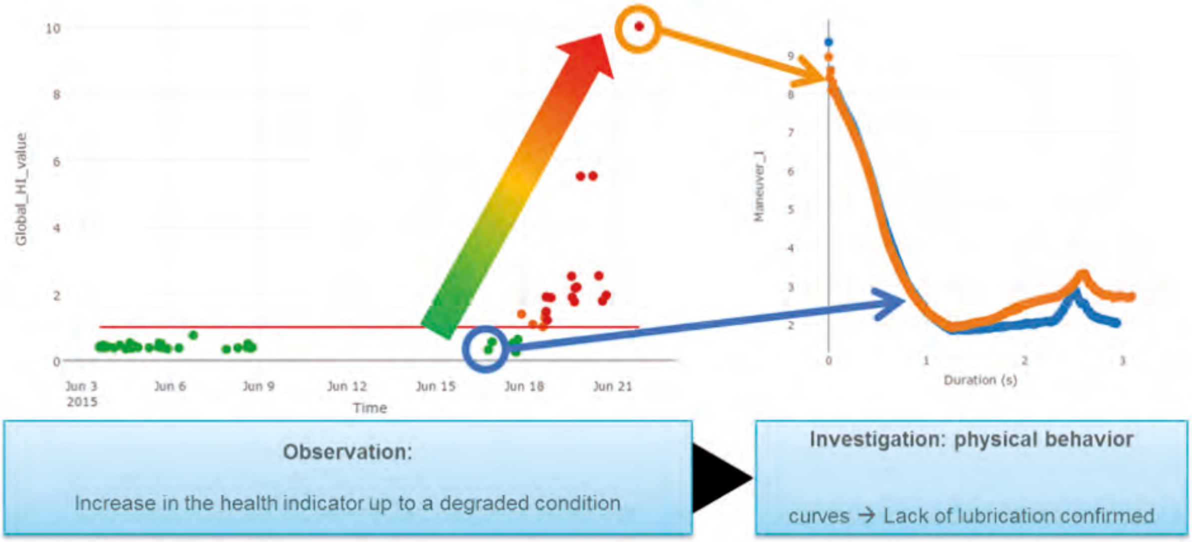
At present it is possible to diagnose the condition of a good number of rolling stock and infrastructure equipment, for which predictive algorithms are being developed to adapt the maintenance strategy and ease operational decisions, including:



- Wear of brake pads and brakediscs
- Wear and detection of defects in pantograph contacttrips
- Wheel profile scan to identify defects, and determine the limits ofuse
- Missing or incorrect positioning of exterior elements: fairings, doors,etc.
- Axle-box and motor bearings condition.
- Components diagnosis in main equipment: Traction, Brake, Doors, Train ControlSystem
- Operating condition of PointMachines

The removal of a significant number of maintenance interventions at first-level visits comes true, especially those associated with integrity and wear checking, thus improving the fleet availability by reducing the immobilization periods.

At infrastructure level, one another case-study comes from predictive algorithms already developed for RFI Point Machines after analyzing more than 100,000 maneuvers in 20 devices, allowing the identification and validation of suitable Health Indicators. Detection patterns of lack of lubricant, presence of obstacles and non-compliant deformations were identified, making possible to update the maintenance strategy.



Observation:

Increase in the health indicator up to a degraded condition

Investigation: physical behavior

curves → Lack of lubrication confirmed

There could come also changes in the maintenance processes when a new device is integrated to complement and sometimes provide more accurate information to the maintenance activity.

Specific monitoring devices allow remote data collection from equipment maintenance ports, often providing wider and more precise information than the one retrieved from the train control system.

Maintenance data currently collected from AVE S100 Motor Blocks are being managed to develop and validate models for traction degraded operation, so that automatic alerts could be configured and the root cause for possible failures could be more rapidly investigated, improving troubleshooting and providing means to enhance the fleet availability.

The new decision-making process consistently manages the former available information and established criteria, jointly with the information provided by the new monitoring system, thus providing means for questioning the maintenance interventions and decisions made so far.

4. Conclusions

The progress in processing technologies and the emergence of the Connected Industry are incentives to reconsider the way in which the railway maintenance activity is organized and executed.

Pioneering actions such as the use of augmented reality in the execution of maintenance tasks, the development of predictive platforms and technologies or the use of drones to perform infrastructure maintenance and inspections at electric substation are some examples of initiatives that are starting to shape this change.

But the real leap will come from reliable data interpretation by Big Data techniques and Machine Learning to effectively cross-check commercial service information, maintenance interventions, real-time equipment diagnostics, infrastructure condition and external conditions. At the horizon one can expect a notable increase in the performance of railway vehicles in terms of service availability and reliability as well as opportunities to optimize maintenance costs.

An adequate integration of technology devices and the revision of maintenance processes, together with the development of competencies and skills for the effective data processing will be key-drivers to succeed.

5. References

- The 2016 MHI Annual Industry Report Accelerating change: How innovation is driving digital, always-on supply chain, Deloitte 2016
- Why Industry 4.0 is not only about IoT devices - Ben Merton, April 2017 Errores al implementar un proyecto de IoT, mcpro - Elisabeth Rojas 2016 Industry of things world. Survey report 2017
- Alstom Predictive and IoT researchs and outcomes.

Mobile material

360.revista de alta velocidad
número 5 - junio - 2018. Pág 33 - 43

Technological, economic and sociological factors on the maximum design speed of high speed trains

Schumann, Tilo
Meyer zu Hörste, Michael
Heckmann, Andreas
Lemmer, Karsten

German Aerospace Centre[1]

Abstract

The maximum design speed of high speed trains is the maximal speed the train is operating in regular traffic. Due to certification and testing purposes the real maximum speed is higher. Historically the maximum design speed was understood as a constant value, which is depending on factors like distance to be travelled between two stops, traffic volume, energy cost etc. The study to be presented here aims to show, that the maximum design speed is a variable value depending as well on factors as willingness to pay, which is again depending on variables like airplane ticket prices or even petrol prices.

Keywords: Very High Speed, Aerodynamics, Signalling

---

[1] *Tilo Schumann. German Aerospace Centre (Deutsches Zentrum für- Luft- und Raumfahrt - DLR), Institute of Transportation Systems. Email: Tilo.Schumann@dlr.de*
Michael Meyer zu Hörste. German Aerospace Centre (Deutsches Zentrum für- Luft- und Raumfahrt-DLR). Institute of Transportation Systems. Email: Michael.MeyerzuHoerste@dlr.de. (corresponding author)
Andreas Heckmann. German Aerospace Centre (Deutsches Zentrum für- Luft- und Raumfahrt - DLR), Institute of System Dynamics and Control. Email: Andreas.Heckmann@dlr.de
Karsten Lemmer. German Aerospace Centre (Deutsches Zentrum für- Luft- und Raumfahrt - DLR). Email: Karsten.Lemmer@dlr.de

1. Introduction and current discussion of high-speed mass transport systems

When Elon Musk published his Hyperloop-Alpha paper in 2013 [1], he explicitly referred to the California High-Speed Rail project, but ironically enclosed the term “high-speed” by quotation marks. This way he wanted to express his disappointment on the intended speed level, which is 350 km/h at most or 264 km/h on average between San Francisco and Los Angeles.

He even finally concluded: “How could it be that the home of Silicon Valley and JPL, doing incredible things like indexing all the world’s knowledge and putting rovers on Mars, would build a bullet train that is both one of the most expensive per mile and one of the slowest in the world?” With this background, the spectacular solution Mr. Musk is envisioning, is targeted to run at a maximum of 1220 km/h and is supposed to operate in sealed partial-vacuum tubes in order to substantially reduce the aerodynamic drag.

Surprisingly the discussion of the Hyperloop concept does not comment at all on the actual technology leader in terms of speed which is the MAGLEV system that initially was targeted on 500 km/h operational speed not being the end point of its technical potential [2], [3].

A lesson to be learned from the MAGLEV experiences is about the application of a customized track system elevated and supported by pylons that the Hyperloop-Alpha paper assumes to be a major item to save money compared to conventional rail track systems.

However, the incompatibility to existing rail infrastructure either requires to purchase premises for stations where they are in particular expensive if available at all, i.e. in downtown areas, or to accept access times similar to planes which in turn compromises optional travel time gains by higher running velocities.

In order to point out the significance of this drawback, opponents here may refer to the fact that several prominent plans to install long-distance MAGLEV lines have been abandoned in favor of wheel-rail technology in the past [4], although the MAGLEV technology has proven its technical maturity since the 1980’s. Examples are the connections from Beijing to Shanghai or from Hamburg to Berlin.

The potential counterexample is Chuo Shinkansen from Tokyo to Nagoya that, by the current state of knowledge, will be the first long-distance MAGLEV line and open in 2027 [3], [5]. However even there, the approval of the Japanese government to construct this new line was given under the condition, “it could be rebuilt to a conventional high-speed line later, if necessary” [6].

There is no doubt, the existing rail infrastructure, its pure construction value on a global economic scale, its availability in urban centers, defines the competitive edge of the traditional wheel-rail technology.

However in view of the challenges posed by the mobility megatrend very high speed is nevertheless an issue for the steel-on-steel technology.

In fact, the pure technical feasibility of classical trains is not limited to today’s maximums speeds of to say 350 km/h. The TGV world record of 2007, when 574.8 km/h maximum speed were reached, is surely the outstanding example to substantiate this statement. But actually it is only the leading one in a series of records of experimental or commercial train layouts since 1980, in which competing suppliers and operators showcase their capabilities, see Figure 1.

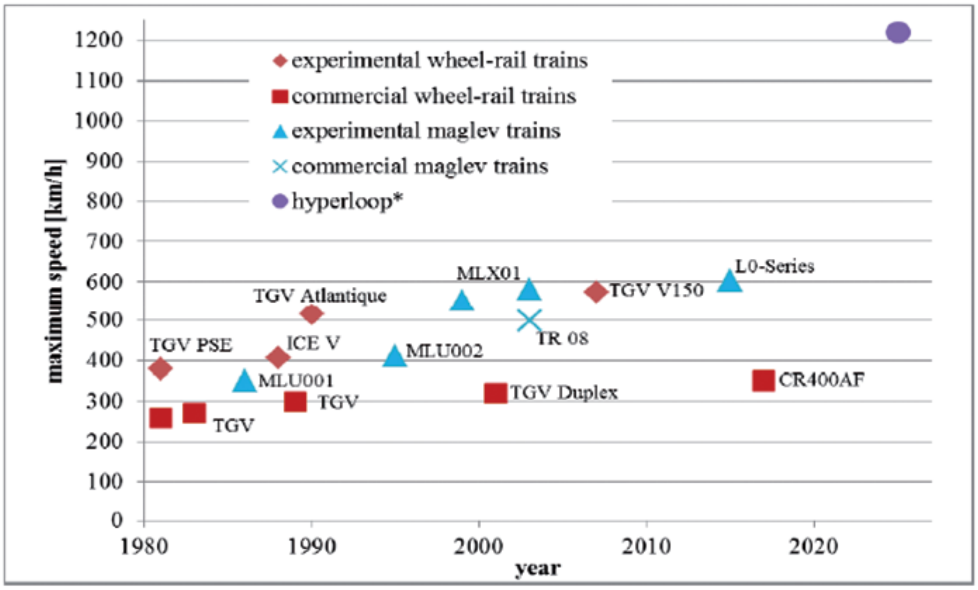


Figure 1. Maximum speed of passenger trains in the last 40 years,
**concept study experimental = technical modified or prototype vehicles commercial = vehicles are used in daily operation.*

Even though the record runs of experimental vehicles in Figure 1 each document a certain status of feasibility, which is still to come for the Hyperloop idea, their relevance for every-daypurposes is actually limited. The TGV world record for instance was conducted on a brand new high speed line exploiting downhill segments with increased catenary tension and voltage, the test vehicle was assembled with additional powered axles, larger wheels and deployed aerodynamic improvements. Since these circumstances cannot be transferred to regular operation, an interesting question in consideration of the summary in Figure 1 and in competition to the Hyperloop and the MAGLEV concept still remains unanswered, namely:

What is a reasonable upper speed limit for high speed wheel-rail systems in daily operation?

This question was adopted to be elaborated within the project to be introduced in the following section.

2. DLR's Next Generation Train Project

In 2007, DLR initiated a long term research project on a future railway vehicle called Next Generation Train (NGT). Eight high level objectives were specified:

- Increase in the permitted speed in daily operation to 400 km/h and additionally explore the velocity range up to 600 km/h
- Halving the specific energy demand compared to the ICE 3 at 300 km/h
- Noise reduction
- Increase in comfort
- Improvement in vehicle safety
- Improved wear behaviour and life cycle costs
- Cost-efficient construction using modulisation and system integration
- Improvement in efficiency of development and approval processes

With this background, a concept for a train set of 202 m length consisting of eight coaches and two train heads was developed and named NGT HGV , see Figure 2.

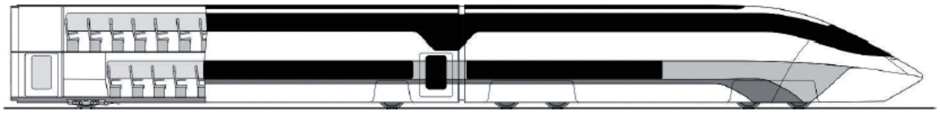


Figure 2: Side view of the NGT HGV with leading train head and first intermediate coach

Important features considered by this train concept may be summarized as follows:

- Double deck configuration in order to increase the number of seats per train length and in turn reduce energy consumption per seat
- Ambitious light weight design and streamlined construction of carbodies and running gears accordingly
- Use of running gears with independently rotating and driven wheels throughout the complete train set in order to distribute traction effort, actively control running stability and reduce wear and noise

Due to the organization of DLR, the NGT project could be organized as a joint effort of 11 institutes involving disciplines such as structural mechanics, vehicle dynamics, aerodynamics, electrical engineering as well as systems and operational engineering.

The same wide-range expertise could also be exploited to explore the speed range up to 600 km/h and elaborate on the “reasonable upper speed limit” question posed in the section above.

3. Technical Aspects related to higher train speed

3.1 Vehicle Dynamics

Modern high speed trains are complex systems and quantities such as forces, accelerations or wear result from the interaction of many components and environmental as well as operational influences. In order to master a fundamental question as given, it is helpful to subdivide the transportation task from the vehicle-dynamical point of view, which leads to three sub-tasks each focused on one specific direction of motion, cf. [7]:

a. load bearing, that is mainly related to the vertical dynamics of the vehicle,

b. guidance, which is associated to the lateral dynamics along curved and tangential track and

c. traction in order to transmit propulsion and braking forces in longitudinal direction.

As regards a., i.e. vertical dynamics, the comfort of the passengers is dominated by forced vibrations of the bounce and pitch motion of the vehicle, cf. [7, p. 14], which in turn are ruled by the quality of the track and the train speed. Due to human perception, comfort is in addition a function of the vibration frequency with the main emphasis on frequencies between 4 and 8 Hz [8].

On the one hand, measurements of non-ideal , real tracks expose a rise of irregularity amplitudes for increasing wavelengths. On the other hand the excitation frequency, the vehicle is exposed to, is a linear function of the train speed divided by the wavelength. These two relationships together constitute the following effect: as regards a specific frequency under consideration, the wheels are excited by larger amplitudes, if the train runs faster on the same, non-ideal track. In turn, the vibration comfort will be compromised and higher dynamic contact wheel -rail forces will stress wheels and rails for higher train speeds.

However, passenger railway vehicles use two levels of suspensions and it is a matter of concept and design tuning to organize a comfortable train ride and keep the dynamic forces at the wheel-rail interface within acceptable limits. There is a large, so far unexploited potential to deploy active or semi-active components which allow for online adaption of suspension characteristics to the train speed and the track quality is a tuning parameter as well. In summary, train speeds up to the range of the TGV record seem to be feasible even in daily operation concerning vertical dynamics.

As regards b., the lateral dynamics of railway vehicles is strongly related to the so-called hunting motion[1] : lateral track irregularities initiate lateral oscillations of the wheel-sets that are intended to be damped out in order to fulfill the guidance task. However, the stability of this hunting motion depends on the vehicle speed, which defines a requirement for vehicle design. Measures such as low equivalent conicities of the wheel-rail contact geometry [16], [9], adjusted stiffnesses of the primary suspensions [10], additional yaw dampers [11], [12], long wheel-bases [10] among others increase the stability region or the maximum speed a vehicle is capable of running safely, respectively. It can be concluded that a proper mechanical design provides a stable running of a railway running gear even at very high speeds. As Delfosse mentioned in [13], the simulation of the modified TGV unit, which set the former speed record of 515.3 km/h, showed that the critical speed of the train was above 700 km/h.

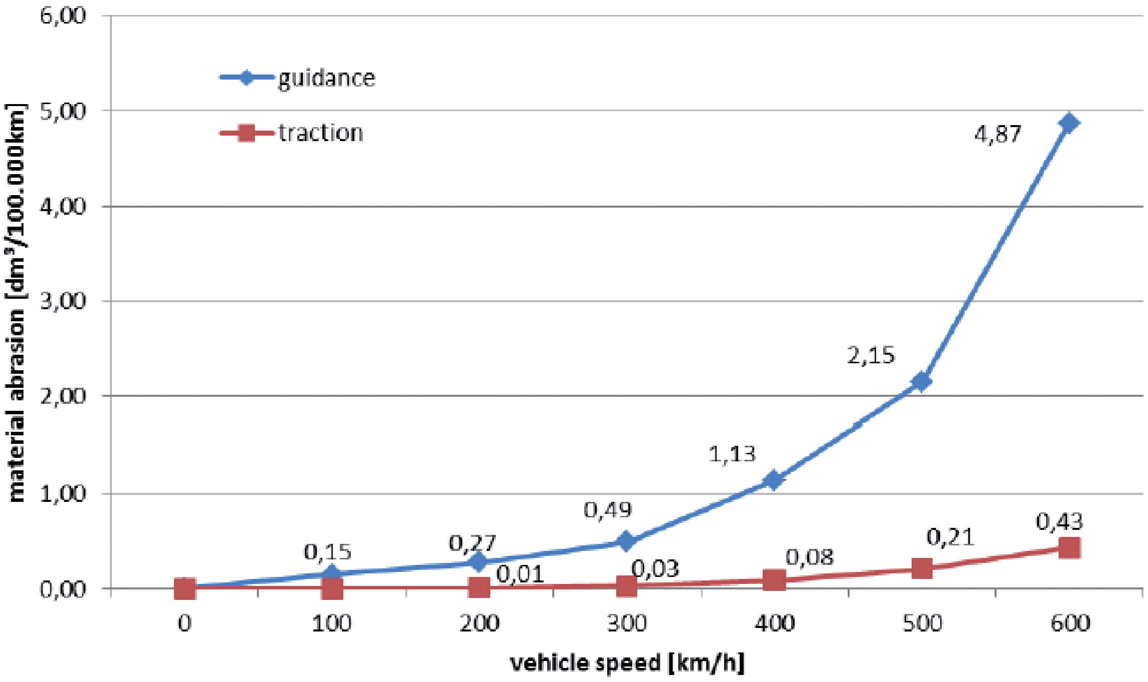


Figure 3: Exemplary estimation of wheel wear as a function of vehicle speed

---

[1] *Streaktly speaking, this applies to the vast majority of railway vehicles that use wheel-sets, only.*

Wear is another aspect to be taken into account. In order to transmit tangential forces it is required to violate the ideal rolling condition at the wheel-rail interface, i.e. there is a small relative motion between the contact partners, which leads to sliding friction [14]. Wear or more precisely the volume of abrasive removed material is related to the associated frictional work [15]. However, wear and rail corrugation are very complex processes, so that a general statement how they depend on the train speed is difficult and out of reach in the given context. Just in order to get an idea, a multibody simulation of a today's articulated highspeed train with 6 cars, two bogies each car, 16*t* load per wheel-set, running on a straight track was performed considering track irregularities. The material abrasion at the wheels per traveling distance as a function of vehicle speed was evaluated on a trial basis. The exemplary results in Figure 3 expose a rough trend: wheel wear for guidance grows progressively with the train speed.

The wear partition associated to traction, i.e. longitudinal dynamics, grows less intense but still with the fourth power of the vehicle speed as shown in Figure 3. This characteristic is based on the fact that the longitudinal forces at the wheel-rail interface depend on the resistance forces which in turn are dominated by the aerodynamic drag at very high velocities.

The traction potential itself is as well a function of the running speed. According to the prominent historical survey by Curtius and Kniffler [16], the friction coefficient converges asymptotically against $\mu = 0.16$ on average for very high speeds, while the lower bound of the measurements indicates $\mu = 0.1$ to be a very reliable figure. The aerodynamic drag may approach values of approximately 200 kN at 600 km/h, which requires 13 wheel axles each loaded with 16 t be counterbalanced by traction with $\mu = 0.1$. This appears to be feasible but indicates the necessity to power as much wheels as possible in order to fully exploit the available traction potential. That's why one intermediate car of TGV 150 that set the world speed record of 574.8 km/h in 2007 was equipped with additional powered running gears [17].

3.2 Aerodynamics

Usually the aerodynamic forces like the drag scale with the stagnation pressure $\rho V^2/2$, where ρ is the density of the air and V is the incident flow velocity in the reference frame of the vehicle. In still air V corresponds directly to the driving speed U. It follows that the power which is required to equalize the aerodynamic drag is proportional to U^3, and at higher speeds the aerodynamic drag will exceed the effect of mechanical friction [18].

However, today's driving speeds up to 600 km/h correspond to a Mach number of $M \approx 0.5$, so that new aerodynamic effects associated to the compressibility of the fluid enter the picture. The critical Mach number specifies the lowest Mach number at which the airflow over some point of the train reaches the speed of sound. Above this critical Mach number the aerodynamic quality of the vehicle will degenerate rapidly. To push the critical Mach number above $M = 0.5$ a train requires a relatively long pointed nose similar to the Japanese Maglev train [19].

A highly safety relevant aerodynamic aspect concerns the crosswind stability in particular if lightweight design is under consideration. Although newer train head designs show elements to reduce the cross-wind forces [20], [21], the optimization potential is limited in general. Investigations with the NGT train concept support the assumption that operational train speeds beyond 400 km/h require a specific device to prevent the lift-off and overturning of the vehicle as it is proposed in [22] or as it is conceptually given by the MAGLEV guidance system. An alternative way to deal with the cross-wind issue at higher speeds is to protect the train from strong gusts using wind fences. Such fences could act as sound barrier at the same time and help to reduce noise emissions of high speed trains.

Another safety relevant aspect concerns the aerodynamic loads which the flow around the train induces on its surrounding. These loads typically as well scale with the square of the driving speed. Either the track-side objects like noise barriers or signal installation are placed at a larger distance from the track or the objects are designed more sturdily to withstand the higher loads.

In the range of 200 to 300 km/h the aeroacoustic emissions of a typical high-speed train are in the same order as the wheel-rail sound. At higher speeds above 300 km/h the aeroacoustic effects dominate [23]. The flow around structural elements like the parts of a pantograph cause a dipole type sound emission whose intensity scales with U^6 [24] [25]. This means a pantograph at 600 km/h radiates about 64 times more acoustic power to its surrounding than at 300 km/h. The measurements presented by Kurita [26] showed that a reasonable noise reduction can be achieved by using aerodynamically optimized pantograph geometries, acoustically absorbing surfaces, and so called noise insulation plates on the roof of the train (see also Yamada et al. [27] and Ikeda et al. [28]). The insulation plates shield the acoustic emissions from the pantograph in lateral direction. The experiments of Baldauf et al. [29] showed that by using an actively controlled single-arm pantograph it is possible to reduce the pantograph noise about 10dB compared to the standard pantograph installed at the German ICE trains. Combining new pantograph designs and noise insulation plates, it appears not unrealistic that the radiation intensity of the pantograph noise at 600 km/h can be reduced to today’s standard level at 330 km/h. Thus the aeroacoustic emissions of pantographs do not constitute an insuperable obstacle for train speeds above 400 km/h.

3.3 Signaling and Train Control

In an approximation the braking distance grows quadratic with the speed, which means that the minimal braking distance increases from 2800m to estimated 11.2 km at 300km/h. Hence it is state of the art, that drivers cannot control the train by trackside signaling at speeds above 160 km/h. Nevertheless all the elements required for a suitable train control and signaling system are available today:

- Cab Signaling
- Safe and reliable radio connections
- Safe on-board Localization
- Train Integrity Supervision in multiple units
- Continuous control and supervision of speed
- Train separation by moving block
- Automatic Train Operation (ATO)

Most of the Elements are part of the European Rail Traffic Management System ERTMS and proven in use. Only two of the required technologies are currently objective of ongoing research activities: moving block and ATO. Therefore it can be stated that the signaling needs to be adopted for very high speeds but it is not limiting the development.

4. Operational Aspects related to higher train speed

Travel time savings are only possible if the train uses high-speed lines (HSL) for a big part of the journey. Therefore new tracks are necessary to increase the speed to over 320 km/h, which is the current speed maximum in Europe.

Furthermore an efficient operation is only possible with lines exclusively used by high-speed trains during the operation time of day. The higher the speed difference on mixed-traffic lines is (with high-speed, freight and regional trains) the more line capacity gets lost. For a speed over 200 km/h it is difficult to operate mixed traffic lines [30]. Many of the HSLs in the world are used exclusively by high-speed trains: Ligne a grande vitesse in France, Shinkansen in Japan, Passenger Dedicated Lines (PDL) in China and Lineas de Alta Velocidad (LAV) in Spain. The different gauge between HSL and the old network in Spain and Japan prevents these lines of being used by conventional trains. In Germany most of the HSLs are built to allow mixed traffic all the day. The advantage is a better line utilization and a more efficient freight train operation due to shorter route length, low gradients and possibly longer trains. Due to safety reasons in tunnels and capacity restrictions the freight traffic is limited to the night time when there is no passenger traffic. The disadvantage is this concept is the lack of time for maintenance works which has to be done with track or total line closures [31]. In France or Japan HSLs are completely closed for intensive maintenance all the night.

To increase the capacity of an HSL the trains should be operated at the same speed. For this also trains with more stops (comparable to the Kodama trains on the Tokaido Shinkansen in Japan, which stop every 15-50 km) have to use strong motorized vehicles [32].

A general problem of increasing the speed of passenger trains is a growing disparity between the operational effort (energy, wear, etc.) and the travel time savings. The travel time reduces in a hyperbolic way, the additional benefit diminishes with higher speed, but the effort grows exponentially.

4.1 Optimized Traction force with very high speed

For the operational analysis special train models for 300, 500 and 600 km/h are derived from the specified NGT 400. For the 300 km/h level also a special version of the NGT is used and not existing HSTs to preserve comparability.

The dimension of the engines increases drastically with the speed. The 400 km/h version has to handle 18 MW driving power whereas the 600 km/h version has to be designed with at least 40 MW. The latter value doesn't include efficiency losses and power demand of auxiliary and comfort systems so the electrical systems have to be designed with significantly more power. Usually trains are designed with additional traction force for instance to handle gradients. In view of the enormous propulsion power to install the idea is to dispense with a reserve. A simulation showed that the effect of this saving is not significant. For a line like the new built one from Stuttgart to Ulm there are some sections with 25‰ gradient. Considering a quite low slack time percentage of 3% the journey between the two cities would be only 20 seconds longer. Thus other use cases look similar and it can be stated that a design without a traction force reserve is acceptable.

4.2 Demand analysis for the reference line Paris-Vienna

To gather information about the effect on passenger demand of increased travel speed, the NGT reference line from Paris to Vienna is chosen for analysis. An operational concept for this line for 400 km/h was developed at an earlier stage of the NGT project [33]. The model is reused and modified for the following speed levels: 300, 400, 500 and 600 km/h. The passenger demand model is based on the European rail network. It includes almost all cities with 50 000 inhabitants and more. These are 1900 cities with 237 Mio inhabitants in countries totaling 525 Mio inhabitants. 120.000 kilometers of rail lines (50% of the real network) are used by 2000 routes with an accurate modelling of travel times and stops. Statistical data for the calibration originates from Eurostat [34]. It has accuracy on the NUTS-2 level. These are smaller countries in Germany or regions in France. Additional data from the UIC statistics was used to complement and verify the Eurostat values [35]. The model is used with four operational scenarios corresponding to the speed levels.

Table 1 shows the results of the analysis. With 600 km/h a train journey between the stations in Paris and Vienna would take approx. three hours. 83 Mio passengers would use the 1145 km line with intermediate stops in Strasbourg, Stuttgart, Munich and Salzburg. The 400 km/h train covers the distance within around 4 hours and attracts 69 Mio passengers, while the 300 km/h version takes around 5 hours and attracts 55 Mio passengers.

The transport performance is raised by two thirds between 300 and 600 km/h from 19 to 31 bn. pkm/year. The acceleration from 300 to 400 km/h generates almost half of this benefit (5 bn. pkm/year). The increase from 500 to 600 leads to an increase of 2.6 bn. pkm/year. So the benefit gets smaller with rising speed, which is directly related to the travel time savings.

Table 1: Growing Demand on the reference line Paris - Vienna depending on the speed

	300 km/h	400 km/h	500 km/h	600 km/h
Travel time Paris-Vienna [h:min]	4:42	3:51	3:23	3:03
Passengers [Mio/year] on reference line	55.3	69.1	77.2	83.0
Passenger km [Bn.Pkm/year] on reference line	19.00	24.84	28.38	30.99
Mechanical energy consumption (at wheel level) [MWh] for one run Paris Vienna	19.2	31.6	44.8	60.8

Hence the doubling of the travel speed from 300 to 600 km/h results in a reduction of travel time by 35%, an increase of travel demand by 63% and an increase of energy consumption by 216%.

The heaviest usage of the line can be found between Paris and Strasbourg, between Stuttgart and Munich and on the Austrian part of the line. Despite the strong national traffic volumes, the international traffic profits most of the new travel speed. The traffic volume between Austria and Germany as well as between Austria/Germany and France is growing much stronger than the national ones.

4.3 Impact of very high speed on the operation

The increasing traffic volume effects a more intensive train operation as shown in Figure 4.

One train per hour and direction with a capacity of 800 passengers is necessary when the line is designed for 300 km/h. With 600 km/h two and a half trains per hour and direction have to circulate. Additional trains run between Paris and Strasbourg and between Stuttgart and Vienna. More additional trains are necessary between Stuttgart/Munich and Salzburg/Vienna.

A train every 15 minutes will run between Stuttgart and Munich with 300 km/h. This increases to a frequency of every 10 minutes between the two cities with 600 km/h. All these values are valid in the morning peak hour.

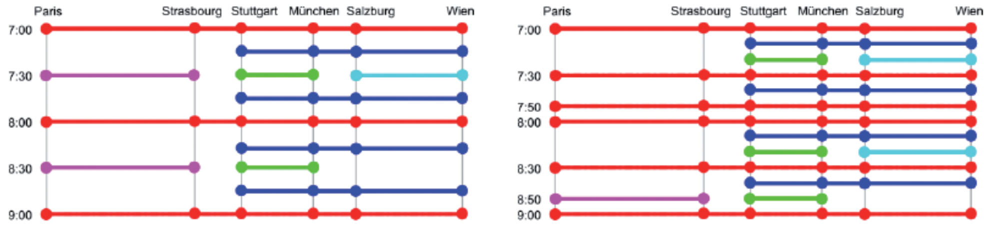


Figure 4: Timetable scheme for 300 km/h (left) and 600 km/h between Paris and Vienna. the most densely traffic occurs between Stuttgart and Munich

Table 2 shows the results of the operational analysis of the different speed levels. The average speed between Paris and Vienna is increasing by 53%, whereas the average speed of all passengers using a part of line is increasing by 23%. The big percentage of using classical railway and the access and egress times reduce this kind of speed. But this value tells us to require big efforts for a network-wide increase of speed and the combination of fast access transport modes like urban public transport or private car or car-sharing in less-dense populated areas.

The number of necessary trainsets is similar over the speed levels, because the higher speed allows a more intensive use of trains which compensates the higher demand. The traffic (operational) performance increases by 57%, though the speed is doubled and the energy consumption triples. This figure shows the faster growing effort compared to a regressive benefit. The operational performance of the trainsets is impressively high; at 600 km/h the project-defined limit of one million kilometres per year is exceeded. A very intensive maintenance is necessary. Probably this will be a big part of the operational cost. The life expectancy probably won’t be 30 years as those by other rail vehicles, especially taking the lightweight construction into account. Thanks to a compulsory reservation and the relative small vehicle size, it’s possible to reach high seat utilization of above 80%.

Table 2: Compilation of operational aspects for the speed levels

	300 km/h	400 km/h	500 km/h	600 km/h
Average speed between Paris and Vienna [km/h]	244	297	338	375
Average speed of all passengers using the line at least for a part of the journey including dwell/access/egress times [km/h]	119	130	138	146
Specific energy consumption at wheel level [Wh/(km * seat)]	21.0	34.5	48.9	66.4
Number of NGT trainsets (incl. 10% operational buffer)	37	40	42	44
Operational performance on the line Paris-Vienna [Mio trainset-km/year]	29.4	36.3	40.7	46.3
Average operational performance of one NGT trainset [km/year]	786 000	916 000	973 000	1 052 000 *(above limit)*
Seat utilization (reservation compulsory)	81%	86%	87%	84%

5. Conclusions

However even if the exact quantitative values are treated with reserve, Figure 5 nevertheless presents a common trend: whatever effort is considered, there is a progressive rise with respect to the maximum operational speed. The higher the considered level the more costly each additional speed step-up turns out to be.

This is contrary to the expected benefit in terms of travel time reduction, which performs on a diminishing scale. Traveling at 400 km/h maximum speed results in a travel time reduction of 18% compared to travelling at 300 km/h, while 600 km/h maximum speed reduces the travel time just by 10% compared to 500 km/h. Note, already the time span $t(v)$ required to travel a fixed given distance s is a hyperbolic, i.e. a declining, function of the velocity v, reviewed under steady-state conditions temporarily disregarding operational aspects.

As a final result, the authors expect the maximum velocity in operation to tend against a saturation point, but which is depending on external factors, too. The balancing of the benefits, efforts and issues such as crosswind stability and energy supply substantiates the assumption that a further increase of the maximum speed of wheel-rail systems beyond 400 km/h will depend on external influences and looks apparently not reasonable today, which could change in the future under changed conditions. The technology will be there to support to run with even higher speeds.

6. References

- [1] E. Musk, SpaceX, “2013. [Online]. Available:

  http://www.spacex.com/sites/spacex/files/hyperloop_alpha-20130612.pdf. [Zugriff am August 2017].
- [2] - H. Hochbruck, High-speed trains, “ETR - International Edition, pp. 6-16, 1 / 2015.
- [3] - L. Sharpe, Japan prepares to build 500 km/h MAGLEV Line, “Engineering & Technology, 2013.
- [4] - K. Ebeling, High-speed railways in Germany, “Japan Railway & Transport Review, Bd. 40, pp. 36-45, 2005.
- [5] - K. Sawada, Outlook of the superconducting Maglev, “Proceedings of the IEEE, Bd. 97, Nr. 11, pp. 1881-1885, 2009.
- [6] - H. Geerlings, The rise and fall of new technologies: Maglev as technological substitution?, “Transportation Planning and Technology, Bd. 21, Nr. 4, pp. 263-286, 1998.
- [7] - K. Knothe und S. Stichel, Schienenfahrzeugdynamik, Berlin: Springer-Verlag, 2003.
- [8] - D. Thompson und C. Jones, Handbook of Railway Vehicle Dynamics, “S. Iwnicki, Hrsg., Taylor \& Francis, 2006.
- [9] - B. Bergander, G. Dendl, A. Nefzger und D. Nicklisch, Die Entwicklung von Rad- und Schienenprofilen, “ZEVrail Glasers Annalen, Bd. 127, pp. 482-493, 2003.
- [10] - G. Sauvage, Running Quality at High Speeds, “Vehicle System Dynamics, Bd. 15, Nr. sup.1, 1986.
- [11] - A. Wickens, Fundamentals of Rail Vehicle Dynamics - Guidance and Stability, Lisse: Swets 6 Zeitlinger Publishers, 2003.
- [12] - R. Kratochwille, Zum Nutzen schaltbarer Schlingerdämpfer in Trassierungselementen mit veränderlicher Gleiskrümmung, Hannover: Dissertation, Universität Hannover, 2004.

- [13] - P. Delfosse, Very-High-Speed Tests on the French Railway Track, “Vehicle System Dynamics, Bd. 20, Nr. 3-4, pp. 185-206, January 1991.
- [14] - J. Kalker, Three-dimensional elastic bodies in rolling contact, Bd. 2, Springer, 1990.
- [15] - F. Braghin, R. Lewis, R. Dwyer-Joyce und S. Bruni, A mathematical model to predict railway wheel profile evolution due to wear, “Wear, Bd. 261, Nr. 11, pp. 1253-1264, 2006.
- [16] - E. Curtius und A. Kniffler, Neue Erkenntnisse über die Haftung zwischen Treibrad und Schiene, “Elektrische Bahnen, Bd. 21, Nr. 9, pp. 201-210, 1950.
- [17] - K. Barrow, French Smash World Speed Record: Test Train Reaches 574.8 km/h on TGV Est, “International Railway Journal, Bd. 12, Nr. 5, 2007.
- [18] - J. A. Schetz, Aerodynamics of high speed trains,“Annu. Rev. Fluid Mech”, Bd. 33, pp. 371414, 2001.
- [19] - N. Shirakuni, Y. Endo, K. Takahashi und K. Yamamoto, Overview of new vehicles for the Yamanashi Maglev Test Line, “in Proceedings of the 17th International Conference on Magnetically Levitated Systems (Maglev 2002), Paper 5104, Lausanne (Switzerland), 2002.
- [20] - J.-L. Peters, How to Reduce the Cross Wind Sensitivity of Trains, “in The Aerodynamics of Heavy Vehicles: Trucks, Buses, and Trains, Lecture Notes in Applied and Computational Mechanics Vol. 19, Springer, 2004, pp. 453-467.
- [21] - M. Schober, A. Orellano, A. Tietze, M. Weise und S. Steilen, Vehicle front end with reduced sensitivity to side winds,“. Patent EP 2383161A1, 2011.
- [22] - A. Heckmann, B. Kurzeck, T. Bünte und S. Loose, Considerations on active control of crosswind stability of railway vehicles, “Vehicle System Dynamics, Bd. 52, Nr. 6, pp. 759775, 2014.
- [23] - C. Talotte, “Aerodynamic noise: a critical survey,” Journal of Sound and Vibration, vol. 231, pp. 549-562, 2000.
- [24] - E. Latorre Iglesias, D. J. Thompson und M. Smith, Component-based models to predict aerodynamic noise from high-speed trains, “in Noise and Vibration for Rail Transportation Systems, Volume 126 of the series Notes on Numerical Fluid Mechanics and Multidisciplinary Design, Springer, 2015, pp. 481-488.
- [25] - E. Iglesias Latorre, D. J. Thompson and M. Smith, “Experimental study of the aerodynamic noise radiated by cylinders with different cross-sections and yaw angles,” Journal of Sound and Vibration, vol. 361, pp. 108-129, 2016.
- [26] - T. Kurita, “Development of external-noise reduction”, Journal of Environment and Engineering, vol. 6, no. 4, pp. 805-819, 2011.
- [27] - H. Yamada, Y. Wakabayashi and T. Kurita, “Noise evaluation of Shinkansen high-speed test train (FASTECH360S,Z),” in Proceedings of the 8th World Congress on Railway Research, Seoul, Korea, 2008.
- [28] - M. Ikeda, T. Mitsumoji, T. Sueki and T. Takaishi, “Aerodynamic noise reduction in panthographs by shape-smopothing of the panhead and its support and by use of porous material in surface coverings,” Quarterly Report of the RTRI, vol. 51, no. 4, pp. 220-226, 2010.
- [29] - W. Baldauf, R. Blaschko, W. Behr, C. Heine and M. Kolbe, “Development of an actively controlled, acoustically optimised single arm pantograph,” in Proceeding of the World Congress of Railway Research WCRR, Colonge, Germany, 2001.
- [30] - S. Andersen, Zur Problematik des Mischbetriebes von Reise- und Güterzügen bei

hohen Geschwindigkeitsunterschieden, “Eisenbahntechnische Rundschau, pp. 351-360, 05 1995.

- [31] - O. Grein, Success Factors of the German High Speed Rail System, “in Conference High Speed Rail in Europe and the Czech Republic, Prague, 2014.
- [32] - S. Andersen, Hat der Hochgeschwindigkeitsverkehr in Deutschland noch eine Zukunft?, “ZEV rail Glasers Annalen, Bd. 136, Nr. 5, pp. 186-193, 2012.
- [33] - T. Schumann, Operational concept Paris to Vienna, “RTR Special - Next Generation Train, 12 2011.
- [34] - Eurostat, Eisenbahnverkehr - Innerstaatlicher und Grenzüberschreitender jährlicher Eisenbahnpersonenverkehr nach Einsteige- und Aussteigeregion, “2005. [Online]:

http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=tran_r_rapa&lang=de. [access: 15.09. 2011].

- [35] - UIC, International Railway Statistics, Synopsis, 2009.

High speed rail superstructure

FUNDACIÓN DE LOS
FERROCARRILES
ESPAÑOLES

360.revista de alta velocidad

High speed rail superstructure

International Congress on High-speed Rail:
Technologies and Long Term Impacts
Ciudad Real (Spain), October, 2017

Escuela de Caminos
Ciudad Real UCLM

Escuela de Ingenieros
de Caminos, Canales y Puertos
de Ciudad Real
UNIVERSIDAD DE CASTILLA-LA MANCHA

High speed rail superstructure

360.revista de alta velocidad
número 5 - junio - 2018. Pág 49 - 69

Current situation and prospects of electric traction systems used in High-Speed railways

Martínez Acevedo, José Conrado
Berrios Villalba, Antonio
Peregrín García, Eugenio

Adif[1]

Abstract

For their operation High-Speed railways use electric power systems with very specific characteristics and different from other electric consumers. The electric traction currently used in High-Speed networks is characterized by being very energy efficient. This is mainly due to the use of AC systems instead of DC systems. On the other hand, the electrical systems used in High-Speed networks are a source of disturbances, both for the railway system itself and for external power supply networks (public grid). The traditional technologies developed in the past continue to be used in the new railway lines, especially in the case of having a sufficiently robust electrical network. On the other hand the expansion of the High-Speed in zones in which this network is not sufficiently powerful, has begun to introduce new solutions based on power electronics equipment. This technology, very implanted in the railway vehicles, has hardly had implantation in the railway infrastructure. The operational advantages of these equipments are several. In addition to the use of lower quality power grids (lower short-circuit power), they also introduce new improvements in traction energy collection. For example, the catenary neutral zones can be removed, the electrical parameters can be regulated and the system performance can be improved in case of operating in a degraded scenario (a problem electrical substation). Although current catenary power systems are quite efficient and operational, integrated operation with these power electronic equipment means that the final conditions improve.

Keywords: Traction Power System, Disturbances, Grid, Unbalanced, Power Electronics, AC feeding systems.

---

[1] *Martínez Acevedo, José Conrado. Adif. Email: jcmartinez@adif.es*
Berrios Villalba, Antonio. Adif. Email: aberrios@adif.es
Peregrín García, Eugenio. AdifF. Email: eugenioperegrin@adif.es

1. Introduction

The electric traction offers, against diesel traction, advantages such as the possibility of building vehicles of great power and speed, better efficiency from the point of view of energy consumption, and less environmental impact. Undoubtedly this traction is the traction of the present and the future in the railways, occupying the first place in the railway systems of the developed countries with the only important exception of the United States of America. In developing countries this type of traction is also the one that tends to be installed in all of its major railroads.

On the other hand electric traction requires large economic investments in its own facilities (electric power lines, substations and electric power transmission lines for train power), so it requires important economic studies. In any case in railway lines with high traffic speed and high traffic density, the use of electric traction is always necessary.

Traction Power System (TPS) can be distinguished between Direct Current (DC) and Alternating Current (AC) systems with different nominal voltages and power frequencies. Currently the most commonly used TPS is based on 25 kV nominal voltage and 50 Hz power frequency (industrial frequency). Such systems have inherent advantages like simple substation design and low transmission losses compared to DC systems.

1.1 The use of the AC system in High-Speed railways

It can be said that the speed of circulation (v) at which a railway line is designed conditions the electrification system to be used. It is evident that as this speed increases the power demanded by the train is also greater. This is justified considering that in the general formula of drag resistance the term representing the aerodynamic drag is proportional to the square of the velocity:

$$R_a = A + Bv + Cv^2$$
(1)

The term $A + Bv$ represents the rolling resistance while Cv^2 is the term corresponding to the aerodynamic drag. R_a is expressed in [kN] and v in [kph].

On the other hand, in High-Speed railway traction, the fundamental equation of the dynamics applied to a train (with mass M) and characterized by an acceleration γ can be written as:

$$F_j - R - M\,0{,}00981\,i = kM\gamma \qquad (2)$$

In this equation Fj [kN] represents the total effort on the wheels of the locomotive with all its motors; $Mg \sin a$ represents the gravity component; i is the slope expressed in [‰]; k represents the coefficient of inertia of the rotating masses. It is dimensionless, slightly higher than 1. The term kM, therefore, represents a fictitious mass referenced to the wheels of the locomotive.

In summary a train of mass M and drag resistance R_a has an acceleration γ on a line of profile *í*. The motor vehicle must develop in its wheels a total effort Fj which is calculated by equation (2) for each speed and slope.

If it is considered that the power P_r in the wheel of the locomotive is expressed by the following equation:

$$P_r = \frac{1}{3,6} F_j v \qquad (3)$$

This power can be related to the electric power P_{TR} absorbed by the pantograph. To do this, one must know the overall performance of the locomotive which is usually supplied by the manufacturers. In equation (3) the power P_r is expressed in [kW], F_j in [kN] and v in [kph].

As an example if it is considered the case of a Renfe S/100 train (8 cars and a total mass of 393 t), equation (1) is expressed as follows:

$$R_a = 2{,}54 + 0{,}0336v + 0{,}000504v^2 \qquad (4)$$

Considering a speed of 300 kph, a drag resistance of 58 kN is obtained. Taking into account that for this train k has a value on the order of 1,04 and considering a non-slope zone and an acceleration of 0,25 ms^{-2}, it obtains a F_j of 158 kN. Using equation (3) P_r acquires a value of 13 MW.

From this example it can be deduced that the power which is necessary to provide from an electric line for feeding a High-Speed train is very high. If it consider a performance near to 1 (losses in the traction link of the locomotive are negligible), it has that the electrical power demanded by the pantograph of the train are also 13 MW. Considering the following equation:

$$S_{TR} = V_{TR} I^*_{TR} = P_{TR} + jQ_{TR} = |V_{TR}||I_{TR}| \cos\theta_{TR} + j|V_{TR}||I_{TR}| \sin\theta_{TR} \qquad (5)$$

Taking into account a pure resistive load, the electrical power demanded by the train in the pantograph is:

$$S_{TR} = V_{TR} I^*_{TR} = P_{TR} + jQ_{TR} = |V_{TR}||I_{TR}| \cos 0 + j|V_{TR}||I_{TR}| \sin 0 = |V_{TR}||I_{TR}| \qquad (6)$$

It follows that to transport a given power to the train could be acted on the voltage or current so that the required value is obtained. Because it is not in the interest of the intensity value to be high, because the losses that would arise in the transport are proportional to the square of its value, always tends to increase the value of the voltage and to decrease, as much as possible, the of the current. Thus, when working with high values of voltage, it is necessary to always use alternating current.

Considering that, as explained above, the current value is the one that is decreased, a catenary will be required mechanically lighter because the required conductors will be less heavy because they do not need a significant geometric section (a characteristic of this type of catenary is in which only one contact wire is required). The lightness characteristic is essential at High Speed because the contact between the pantograph and the catenary improves.

For all of the above, High-Speed railways should always use AC systems. Unlike the generation of electric power in the power plants, where alternating current is always generated in three phases, the use of alternating current in railways uses only one phase (single-phase alternating current).

1.2 Electric power regenerated

Due to the use of AC systems it can be intuited that the energy efficiency of a high speed line is much greater than in a DC line. This is mainly due to the existence of lower losses in the system.

On the other hand, the management of the electric power regenerated in the braking of the trains has great incidence in the energetic improvement with respect to a line with DC system DC. The braking process a train has to carry out, either to make a stop or to lower the existing slopes along the path, or even to succeed in reaching the speed limits imposed, may lead to important consequences in the final calculation of the energy required by an electric railroad line. Indeed, energy is dissipated in the braking process, some of it is lost in friction brakes (pneumatic brakes), that has no useful use, and some of it is dissipated in the dynamic brakes.

For the dynamic brakes in particular, in the case of having an electric traction train (or a dieselelectric traction one), the braking process involves the generation of electricity. At present, the electricity generated in this type of brake can have multiple destinations:

- Provided the train incorporates regenerative braking, energy is dissipated as heat into electrical resistors provided on board (rheostat brake).
- Provided the train does not incorporate regenerative braking, energy is returned to the catenary. In this case, if there is another train being fed from the same power sector requiring energy, the train may consume the returned energy. This particular example constitutes an optimal process from an energy standpoint. In the case there are no trains demanding energy, two additional cases arise:
    - In case of having a single phase AC power (for high-speed lines), the energy generated is returned to the national grid and can be used by other consumers connected to it.
    - In the case of having DC power, thus existing a rectifier group in the substation, energy is dissipated in the resistors of the rheostat brake provided in the train.

That is, in DC electrifications, energy cannot be returned to the national grid taking into account the current situation, since the substations are equipped with rectifier groups that do not allow current flow into the grid. It also should be noted that the energy is regenerated to the catenary when the train that is stopping has previously fed its auxiliary services (heating, air conditioning).

2. Traction Power System (TPS)

A Traction Power System (TPS) is a system in which it is possible to absorb or generate energy and distribute it to trains in an efficient and safe way. This system represents in itself a power electrical system with its own characteristics. In most cases the TPS is interconnected to the country’s general electrical system (Figure 1). It can also constitute its own electrical system. In the first case the AC railway system will operate at industrial frequency while in the second case they will operate at a special frequency (case of some countries in Central Europe).

A common feature is that the electric energy, from its generation to its delivery to the trains, goes through different stages of adaptation and transformation.

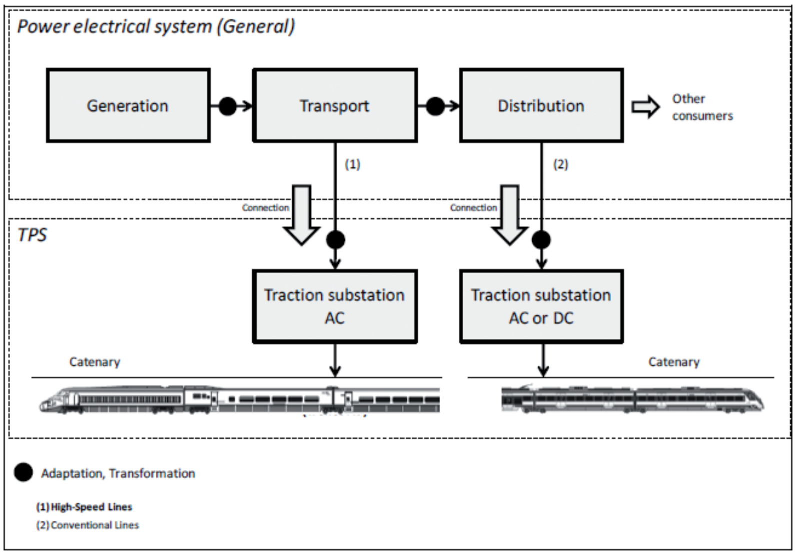


Figure 1. Basic diagram of a power electrical system (general and railway system). Diagram is represented by all the elements necessary for a train with electric traction to operate. In the most common case: 1) Generation sub-system (central of generation); 2) Sub-system Transport (transmission line); 3) Sub-system Distribution (distribution line); 4) High-Speed TPS (traction substation, single-phase electric transmission line to the train (catenary) and train); 5) DC Conventional TPS (traction substation, DC catenary and train). (Source: Author).

The function of the transmission line is to transport large powers from the generation plants to the centers of the load and to the large industrial consumers that exceed the normal limits of the distribution lines. This would be the case of the High-Speed rail, which, due to its power demand, requires a direct connection to the transmission line. This type of line has sufficient short-circuit power to ensure the correct operation of the rail. In this case, therefore, the distribution line is dispensed with and the transmission line is used directly as the distribution line.

3. Source of disturbances

High-Speed trains that operate with industrial frequency (network frequency) are a source of disturbances in the power lines and the own railway environment. It is a load powered by single-phase alternating current, variable in space and time, and power electronics of locomotives. This electronics produces harmonic components of the traction current that flows through the catenary and then returns to the nearby terrain. This fact complicates the operational scenario, taking into account that the rest of the railway systems require electrical cables for their operation.

Although the single-phase alternating current offers an important advantage over the direct current as is its ease of transformation, as a disadvantage is its property of inducing voltages in parallel conductors. Note that in the normal use of alternating current in three-phase systems, the inductions of each phase are compensated by the inductions of the other phases. This fact does not occur in single-phase electrification as there is an electromagnetic disturbance that may be important for other railway installations.

For all of the above it can be said that electrification causes disturbances in the electrical environment of the High-Speed line. These disturbances occur both on the transmission line (as a consequence of being connected to it) and in all the electrical and electronic installations of the railway line.

3.1 Disturbances in the transmission line

These disturbances are mainly the following (Figure 2):

1. Imbalances of currents and voltages at different points in the network, due to the character of the single-phase load.
2. Harmonic currents and voltages that produce the phenomenon of harmonic distortion. They are usually originated by the electronic equipment of the locomotive and even by the electric arcs produced in the interaction between the pantograph and the catenary. Harmonic distortion affects the characteristics of the mains voltage, deteriorating the quality of the service.
3. Voltage fluctuations in the network. It has its origin in the variations of the regime of the load, that is to say, in the randomness of the circulations of the trains and in the variability of the power demanded by each one of them. Voltage fluctuations cause variations in the supply voltage to users near the common connection point.
4. Voltage sag. They are caused by defects in the traction circuit causing harmful effects that may become unacceptable by users connected in the environment of the common connection point.

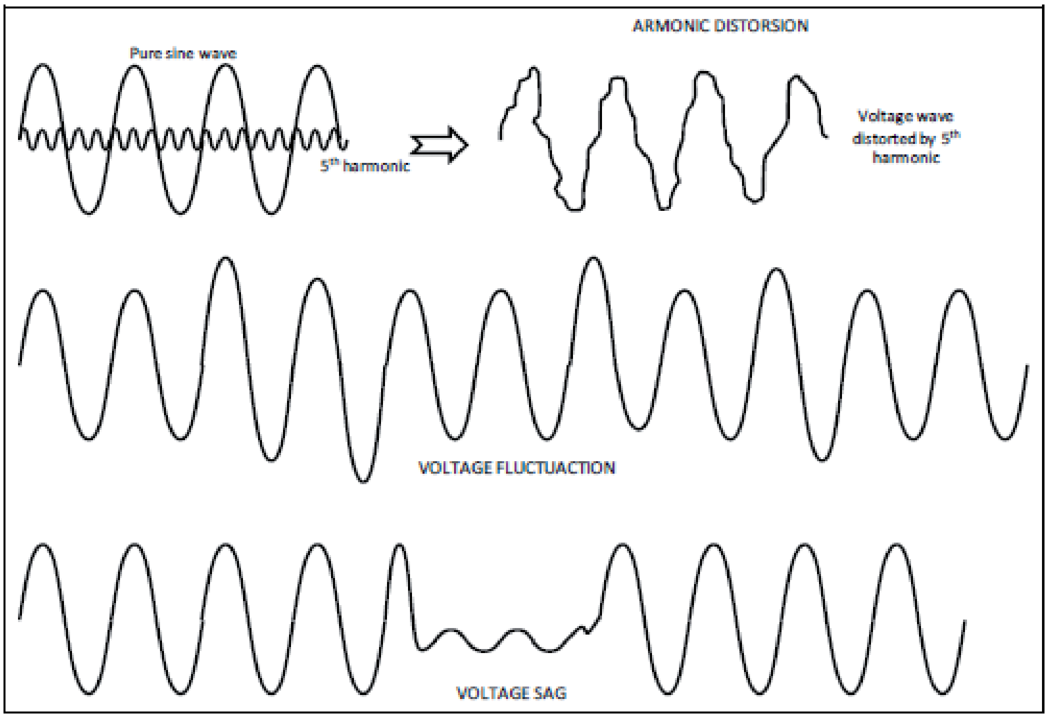


Figure 2. Disturbances. (Source: Author).

3.2 Imbalances of currents and voltages

Of all the disturbances produced on the transmission line, the imbalances of line currents and voltages are the most important (Figure 3). If the same electrical phase of the network were always used, the three-phase system would be decompensated. Therefore, in the electrical supply of High-Speed lines, the phases are always alternated or rotated.

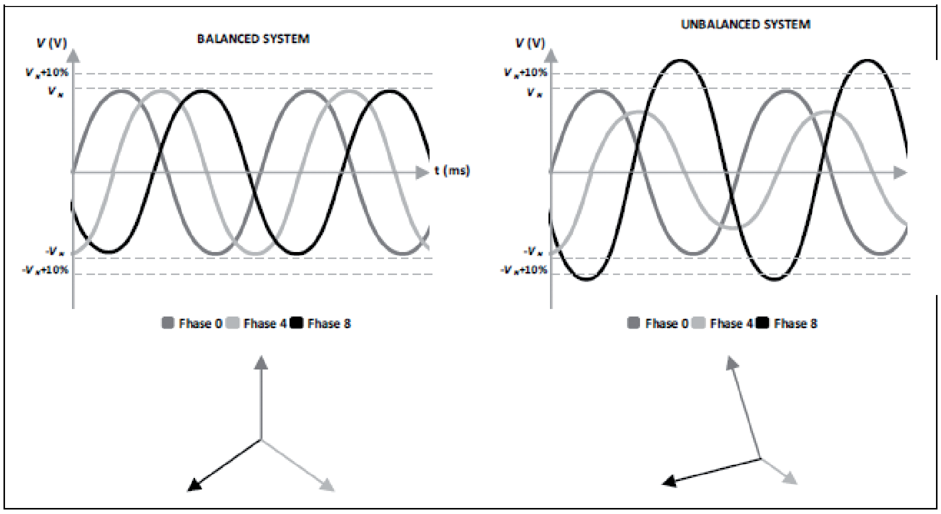


Figure 3. Left) Balanced system; right) Unbalanced system. (Source: Author).

Figure 4 shows a three-phase network to which a single-phase load (represented by a HighSpeed train) is connected. If it considers the unbalanced systems of currents and voltages of a point A in any of the network, both can be represented by two symmetric systems balanced, one direct and one inverse. The ratio of the respective symmetrical components expressing the degree of unbalance. Applying the Fortescue transformation to the calculation of the symmetrical components, it obtains the corresponding values, direct and inverse, of the phaseto-phase voltage at point A. The following equation gives the approximate practical approximate value of the voltage unbalance factor:

$$\tau = \frac{V_{0i}}{V_{0d}} \cong \frac{S_{TR}}{S_{CC}}$$

In this equation τ is the voltage imbalance factor; V_{0i} is the inverse voltage; V_{0d} is the direct voltage; S_{TR} is the single-phase power demanded by the train; and S_{CC} is the short-circuit power of the transmission line.

This relationship between the apparent power of the single-phase load and the apparent short-circuit power at the different points in the network, greatly determines the type of electrical connection to be used between the railway traction substation and the transmission line.

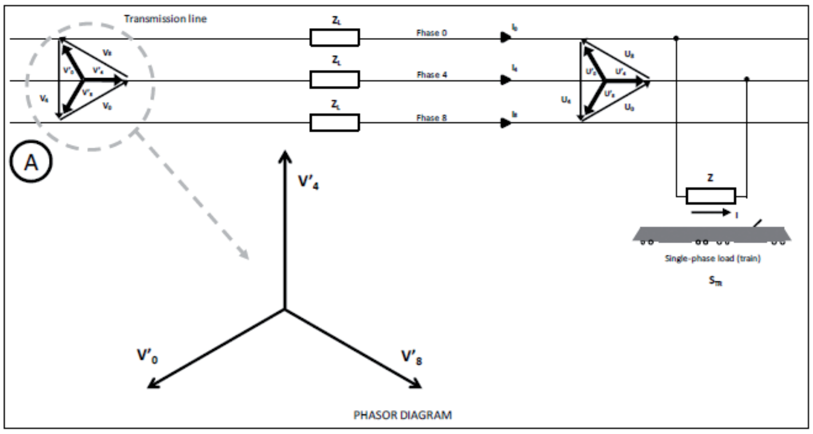


Figure 4. Schematic representation of a single-phase load (train) connected to the three-phase transmission line. Schematic of a balanced three-phase system. (Source: Author).

3.3 Passive techniques to reduce imbalance

Different passive solutions have been introduced for minimizing the effect of voltage unbalance. Although the ideal solution for the railway operation would be to connect the traction substations according to Figure 5A, in practice the substations are connected alternately to the phases of the network as shown in Figures 5B and 5C. Such connections are often referred to as pure single-phase connections. Consequently, the voltages between different feeding sections are 60° or 120° out of phase, so that the network must be sectioned for isolating the paths fed from different substations or different transformers. The phase separation of different feeding sections is done by so called neutral sections.

In the case of Figure 5B it is said that the phase rotation is by electrical substation (the two transformers are connected to the same phase). In the case of Figure 5C the rotation is by transformer.

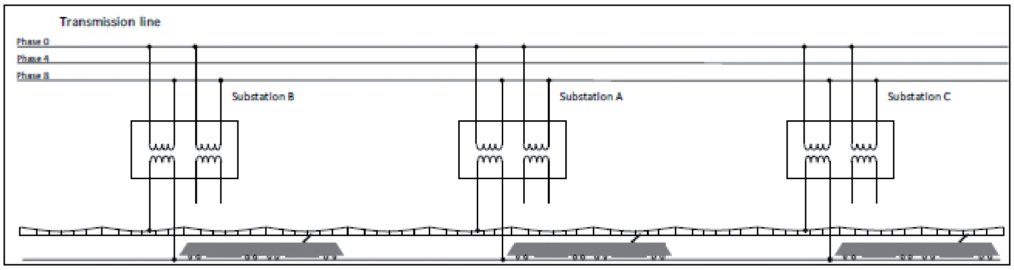


Figure 5A. Pure single-phase connection (ideal but not possible). (Source: Author).

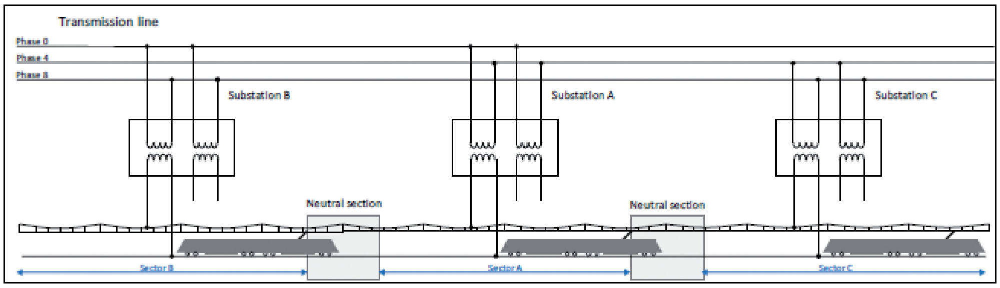


Figure 5B. Pure single-phase connection (phase rotation by electrical substation). (Source: Author).

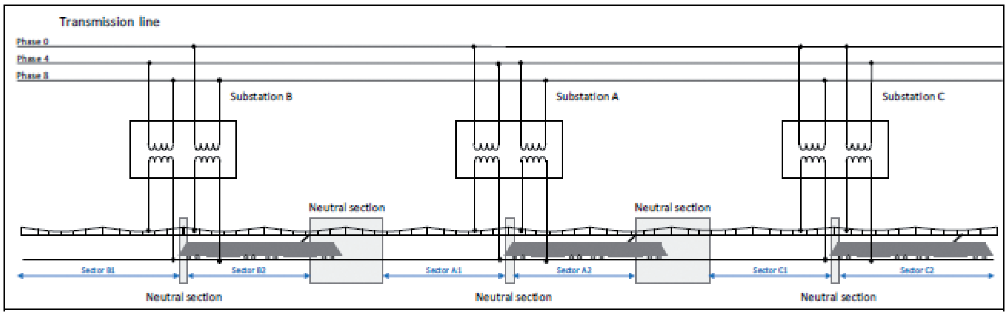


Figure 5C. Pure single-phase connection (phase rotation by transformer). (Source: Author).

3.3.1 Special connections

It is also possible to reduce the voltage unbalance by use of different special types of transformers. The remaining voltage unbalance depends on the type of special transformer used as well as on the distribution of power on the transformer's secondary side terminals.

The connections seen in Figures 5B and 5C are usually those used when there is sufficient short-circuit power in the three-phase substation power supply network. It is observed that the power transformer of the railway substation does not connect to the three phases.

However, if the three-phase supply network does not have sufficient short-circuit power, a single-phase pure connection can not be used. There are other special connections (Scott connection, V-V connection, Le Blanc connection, etc.) that partially solve the presented problem. This would be the case of the Japanese High-Speed network where such connections are often implemented because sometimes the power to High-Speed trains (Shinkansen) is supplied from low-power networks.

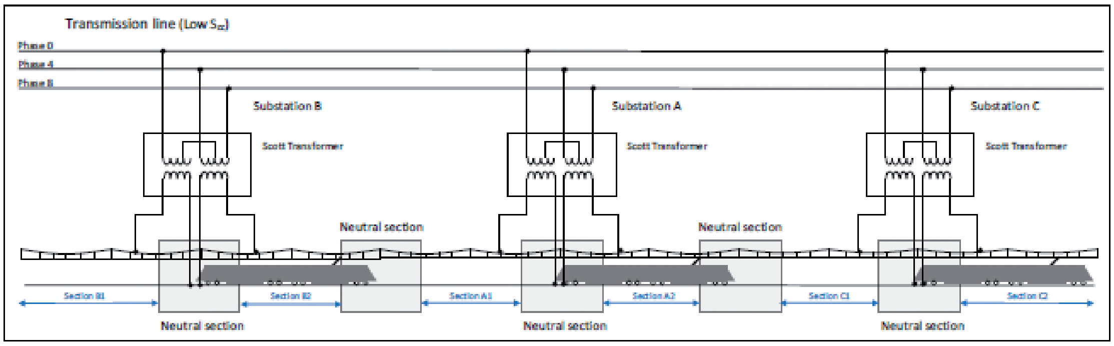


Figure 6. Scott connection (special). (Source: Author).

3.3.2 Neutral section

As seen in Figure 5 neutral sections are installations that are necessary in High-Speed lines as they prevent electrical phase unbalance in the three phase grids that power them. The existence of neutral sections may entail some occasional operating problems and therefore an attempt is made to minimize how many of them there are and improve operation.

Considering the case of the Spanish High-speed network, a neutral section is built with two insulated overlaps between which the no-voltage catenary is installed (Figure 7). According to the indicated diagram, catenary 1 is powered from the electrical phase 2 (voltage V1) while catenary 2 is powered from phase 2 (voltage 2). The de-energised catenary is installed between the two (catenary 3).

The train that is coming along catenary 1 enters into contact with catenary 3 via the first sectioning. Therefore, when the pantograph has lost contact with catenary 1, it only rubs catenary 3 until it reaches the second sectioning, when it starts to come into contact with catenary 3. It should be pointed out that the train enters the neutral section moving by mechanical inertia as its traction switch has been disconnected. It will reconnect when it leaves to be able to be powered from the new stretch again.

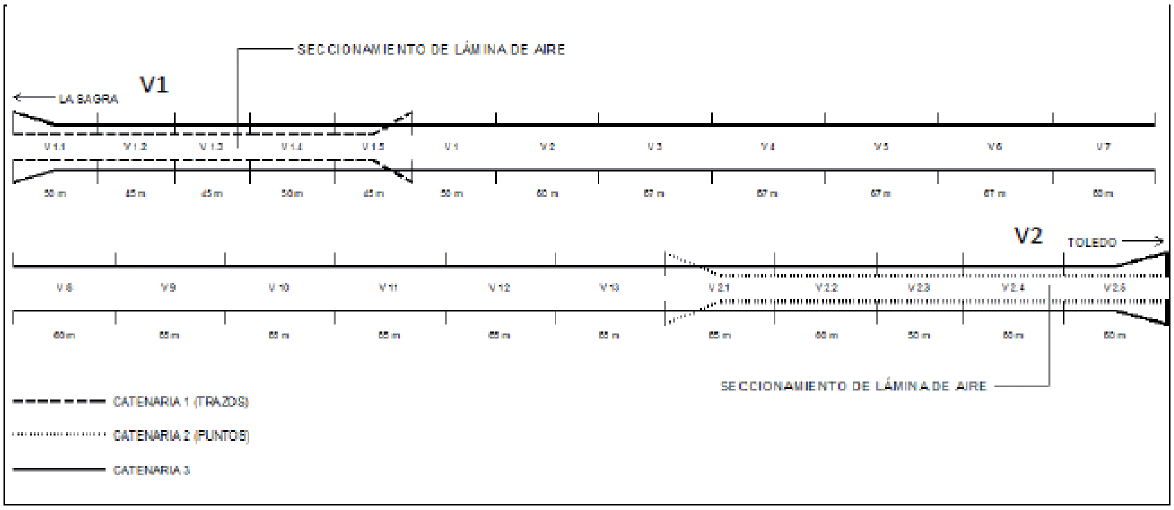


Figure 7. Mechanical layout of a neutral section of catenary. (Source: Adif).

Logically, on passing from catenary 1 to catenary 3, the pantograph will connect to voltage V1. The same will happen with the V2 voltage on passing on the sectioning of catenary 2. The rest of the time, the voltage of catenary 3 will vary depending on the difference of the electrical phases of the voltages of the ends and on the impedance of the de-energised section.

It can be seen that if the time rates of the voltages of the collateral voltage transformers are consecutive (60º phase difference), the electric voltage between the ends of the neutral section is roughly the nominal voltage of the secondary of the transformers. If the rates are not consecutive (phase difference of 120º), the voltage between ends will be roughly √3 times the nominal voltage of the secondary.

A neutral section may be announced with enough time for the train to be able to carry out the opening operation of the traction switch[1]. As seen above, the opening requirement is due to the pantograph energizing the de-energised catenary for a few seconds using the end sectionings, an action that may involve the electrical bypassing of both routes if the train has two pantographs that connect catenaries 1 and 2 via catenary 3 when passing over the sectionings. Furthermore, in short neutral sections, it has been seen that a very high speeds the arc extinction might not have occurred when the pantograph is already in the second sectioning. In any event, the greater the length of the de-energised section, the fewer problems will occur although the loss of speed the train suffers will generally be higher.

The need for installing neutral sections in high-speed lines may entail the following problems:

- Breakdowns on the catenary if a train enters with the traction switch closed. In the event of circulating under ATP this type of incidents will not occur.
- Trains that stop in the section de-energised (for different reasons) and which cannot start up again on their own. Normally a train stopping affects the regularity of the rest of the trains circulating behind it, as some minutes are required for energising the de-energised section and being able to start to run again.
- Generation of fatigue in the train's traction equipment due to continuous openings and closings of the circuit.

It should be pointed out here that the train's lost speed when passing through the neutral

---

[1] *Automatic process if circulating under ATP.*

section[2] is approximately 9 kph with a total time passed since traction of roughly 22 s and having run roughly 1600 m. Even so, it has been seen that the existence of neutral sections on the line does not affect the total route time and the total amount of electric energy consumed.

Even so, and as to be expected, the ideal situation for operation would be to have as few neutral sections as possible.

3.4 Active techniques to reduce imbalance. Introduction of Power Electronics

Furthermore power electronics applications can be used to actively reduce the voltage unbalance in the public grid. For this purpose different technologies and applications are introduced. Furthermore other advantages can be achieved like reduced voltage drop at the TPS. For example static VAR compensators (SVCs) connected to the three-phase grid in parallel to the TPS reduces the voltage unbalance imbalance but require large harmonic filters due to the switching of the thyristors. Then synchronous static converters (STATCOMs) connected to the three-phase or single-phase traction network were used where they also allow to filter the harmonics produced by the traction loads.

Static frequency converters have also been used to provide the total power required by the substation, although in a smaller number due to the lower cost of the SVC or STATCOM since these are dimensioned for a fraction of the total power of the substation. These frequency converters have evolved from back-to-back converters to the current modular multi-level converters (MMC). There are two configurations for MMC converters: AC/AC direct converters that allow to convert power from a three-phase network from 50 Hz to a single-phase network at 16,7 Hz, and the AC/DC/AC indirect converters that allow obtaining from a 50 Hz threephase network a single-phase 50 Hz network.

Finally, in the last years an electrification system called Cophase Power Supply has been developed (the first 5 MW installation was installed in 2015 on the Shanxi Line in China) which is a hybrid solution between a STATCOM and a converter of the total power of the substation. There are different configurations but in all cases they use a single-phase converter that generates on the three-phase side the current necessary to balance the load and compensate for the imbalance generated by a single-phase or three-phase transformer from which the power is supplied to the single-phase rail network. The converter provides active power to the single-phase network and also allows filtering of the harmonics generated by the loads. As the load to the three-phase grid is balanced, the collateral transformers can be connected to the same phase so it is possible to eliminate the neutral sections.

The inclusion of power electronic converters implies the existence of a system of regulation of control its operation, and to a greater or lesser extent, the behaviour of the rail network: from the SVC that hardly influences the voltage of catenary, STATCOM and Cophase Power Supply that can vary the voltage in catenary, and finally, the total power converter that can modify the voltage and frequency. The last two systems have the possibility of operating without neutral sections which allows them to distribute the load between collateral substations, both at active and reactive power levels. These converters may have implemented control methods to function independently of each other (such as Droop Control which is a method for interconnecting multiple voltage sources for microgrids operating in isolation) or by methods that exchange information, for example, to avoid recirculation by the rail network which is in parallel with the three-phase transport network.

---

[2] *Data obtained after the tests carried out by ADIF on HSL 050 (Madrid to Barcelona) and HSL 010 (Madrid to Sevilla). An average speed of 270 kph is considered.*

Power Supply	Converter Power Rating (PerUnit Load)	Neutral Sections	Catenary Voltage and Frequency Regulation	Harmonic Compensation	Catenary Short Circuit Protection
Normal	-	Compulsory	No	No	Passive
SVC	0.58	Compulsory	No	No	Passive
STATCOM	0.58	Compulsory	No	Yes	Passive
Full-Power					
Converter	1.0	Can be omitted	Yes	Yes	Active
Cophase Supply	0.5	Can be omitted	No	Yes	Passive

Table 1. A comparison of different feeding systems of AC railways. Source: [1].

4. Another use of Power Electronics: introduction of a static switching system for the operation of neutral sections

Currently some administrators[3] are starting to try static switching systems to allow the training making the transition between electrical phases directly, without being affected in any way and with the aim of improving the operating capacity of the exploitation.

These types of systems use switches on each side of the neutral section, which are represented by semiconductor equipment that allow carrying out a switching in a very short space of time. A very important characteristic of this type of switching is that the exact position of the train should be born in mind.

The operating principle of the system would be as follows (Figure 8):

- The train that is coming along catenary 1 is detected by a detection system that informs the switching system of the next entry of the train into the neutral section.
- At this moment, a close command is established to switch 1 so that the de-energised section is powered at voltage V1. The train does not receive the open command from the traction disconnector at any time.
- Again, the detection system should detect that the train is totally situated inside the deenergised section. When this condition occurs, an open command is issued on switch 1 and a close command on switch 2 so that the de-energised section is powered at voltage V2. This process is short enough for the train not to detect a lack of voltage in the catenary (which would involve opening the disconnector) and there is therefore no loss of traction. The train leaves the neutral section fed by electrical phase 2.
- In a given point of the exit of the neutral section, the detection system identifies the total passing of the train and starts to normalise the neutral section (opening of switch 2), and is ready for the next train to pass.

As indicated above, the switching system should be designed to act in a given time that is determined by the characteristics of the material that circulates on the High-Speed Line. Thus, if a detection system based on electromagnetic pedals is used, the main characteristic that determines the system’s reaction time is the distance between the first wheel of the train and the nearest pantograph that it may carry in service.

---

[3] It should be pointed out that the detection system may be represented by a track circuit specially designed for this function (as is the case of the Japanese system), or by a series of electromagnetic pedals that detect the position of the train wheel (system used in test by Adif).

The length of the de-energised catenary is also a fundamental parameter for analysing the viability of installing this type of systems. Specifically it has been able to conclude that if this length is lower than established for an interoperable neutral section (402 m), installing switches will not be viable. The reason lies in considering the circulation in double composition of a 200 m train. In this case if the distance existing between the train's first wheel and the farthest away pantograph from the second composition the train may carry in service (a small auxiliary distance derived from the coupling of the two trains should also be considered), it is concluded that this distance is the one that determines the maximum length of the deenergised section to be used. Therefore, the greater the no-voltage section the simpler it will be to install a system with these characteristics.

5. AC feeding systems

TPS can be considered to be composed of two main subsystems: electric traction substations and the railway distribution line (catenary).

The railway distribution line can also be considered composed of the overhead contact line, the return circuit and other equipment that we could call complementary elements. The overhead contact line is formed by several conductors (mainly contact wire, catenary wire and droppers) and their elements of support, cable and insulation (cantilever, tie bar, steady arm, etc.). Its design is characteristic of each technology company or of each infrastructure manager and is specially designed to guarantee the transmission of a certain maximum current to each train (the maximum recommended current in EN 50388 is 680 A for 25 kV 50 Hz) an adequate electrical contact with the pantograph at a certain maximum speed. The return circuit is composed of the rails and other components that they group in the set of complementary elements. These complementary elements depend on each feeding system used.

The configurations of the TPS networks are considered from the point of view of the existing type of AC power systems and from the point of view of the unilateral or bilateral connection of the substations. The different types of AC systems of feeding to the High-Speed networks are the following:

- Simple feeding with rail return.
- Autotransformer (AT system).
- Coaxial.

The fundamental characteristics that determine the use of each one of them is its capacity to transmit power according to the length of the line and the electromagnetic disturbances that generate to its surroundings.

5.1 Simple feeding with rail return or additional return conductor

This connection is the configuration adopted when there are sufficient connection points to the transport line. In this case the distance between collateral substations is usually between 35 and 40 kilometers. If a nominal voltage value V is considered in the catenary, this connection is often referred to as the 1 × V kV system. Considering the standard value of electrical voltage in catenary (25 kV), it is called 1 × 25 kV system. The operating diagram is shown in Figure 9.

Generally the substation transformers are connected to the same electrical phases. In this way the section of the substation is connected to the same electrical phase, and it is not necessary to install a neutral zone in the substation. The primary winding of each transformer is connected to the grid and converts this voltage to that of the train (25 kV). The secondary winding is therefore connected between the catenary and earth.

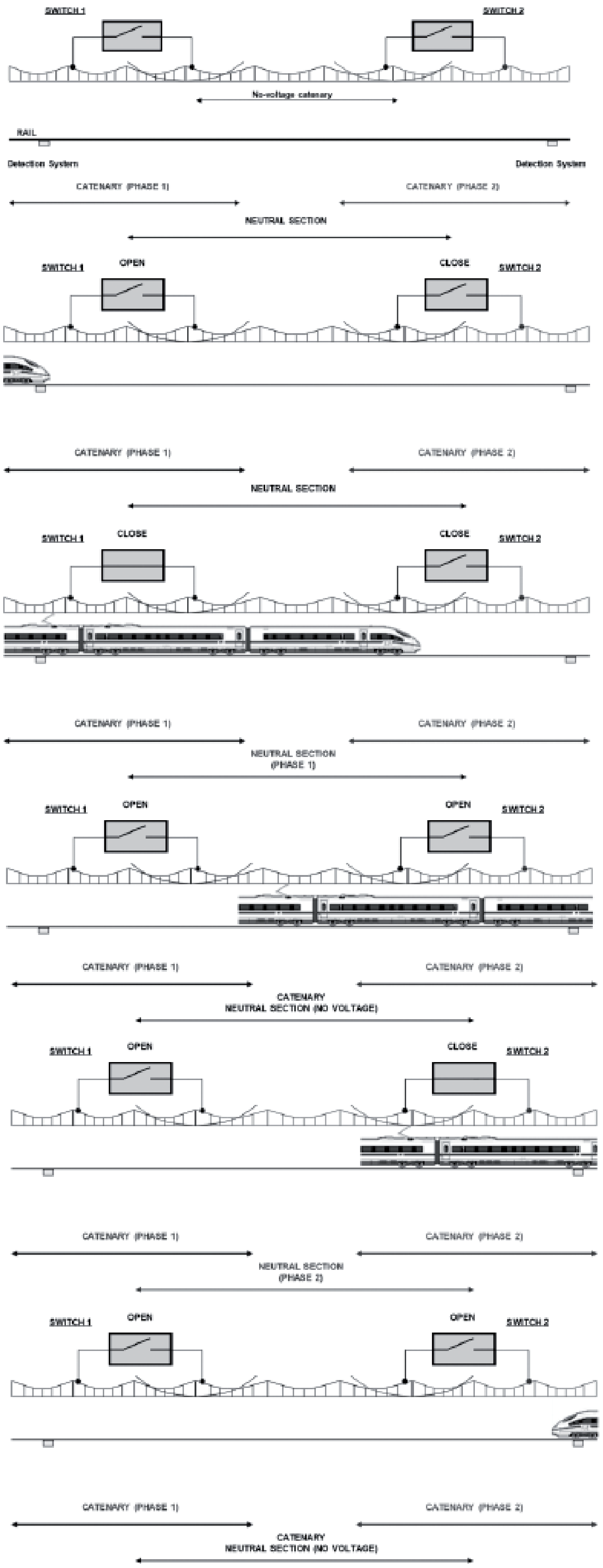


Figure 8. Shunting sequence of the switching system. (Source: Adif).

This connection is economical and easy to exploit and maintain, although there must be sufficient connection points to the grid. As can be seen in Figure 9, energy is transmitted from the substation to the train through the catenary and a reinforcing conductor if it existed. The current return is made by the track rails, the return conductor and the ground.

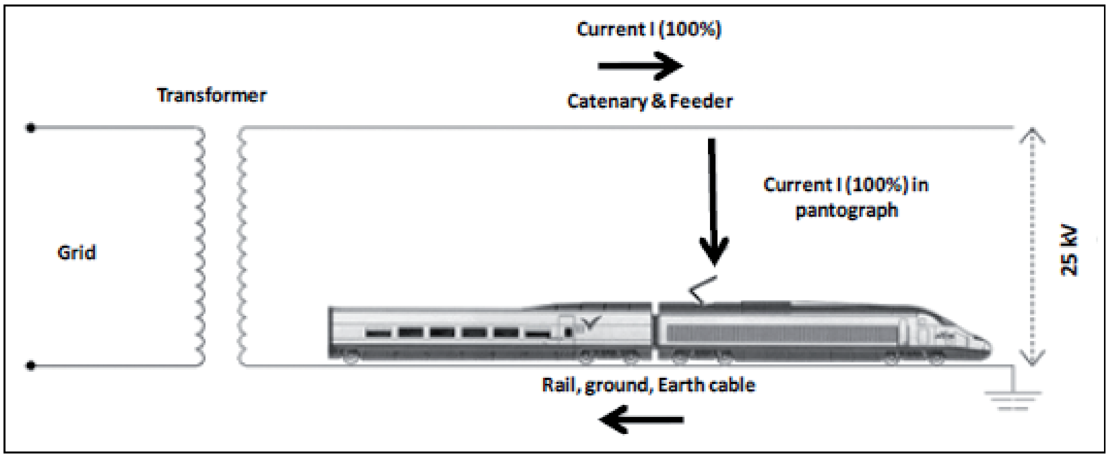


Figure 9. Simple feeding with rail return. (Source: Author).

5.2 Autotransformer (AT)

The AT system is the configuration adopted when there are not enough connection points to the grid. In this case the distance between collateral substations is usually between 60 and 80 kilometers.

If a nominal voltage value V is considered in the catenary, this connection is often referred to as the 2 × V kV system, because as will be seen below, there are actually two electrical circuits in phase opposition. Again, considering the normalized value of electric voltage in catenary, it is called 2 × 25 kV system. This double circuit requires the use of an additional conductor (negative feeder) as well as autotransformers every few kilometers. The main objective is to ensure a voltage drop in adequate catenary considering the greater distance between collateral substations. The operating diagram is shown in Figure 10.

Substation transformers can be connected to the same or different electrical phases. In the latter case each transformer will feed to a semi-section of the substation, being necessary to install a neutral zone in the substation. The primary winding of each transformer is connected to the grid and converts the voltage to twice the train operating voltage (50 kV). The secondary of the transformer has two windings and a central connection connected to ground. A secondary winding is connected between the catenary and earth (25 kV). The other secondary winding is connected between the negative feeder and ground (-25 kV).

Every 10-15 kilometers, autotransformers are connected between the catenary and the negative feeder with its mid point connected to ground. If a train is located at a point in the section, the autotransformers, which have the property of distributing the current that arrives through the central point in two almost equal parts, force the flow of currents indicated in Figure 10. Considering an ideal hypothesis of operation, the distribution of current consumed by the train is made 100% by the catenary to the pantograph. The currents consumed the train are provided depending on the situation of the train. Thus, in the scheme of Figure 10, 50% of the current flows from the substation to the train through the catenary. The remaining 50% is provided by the autotransformers between which the train is located.

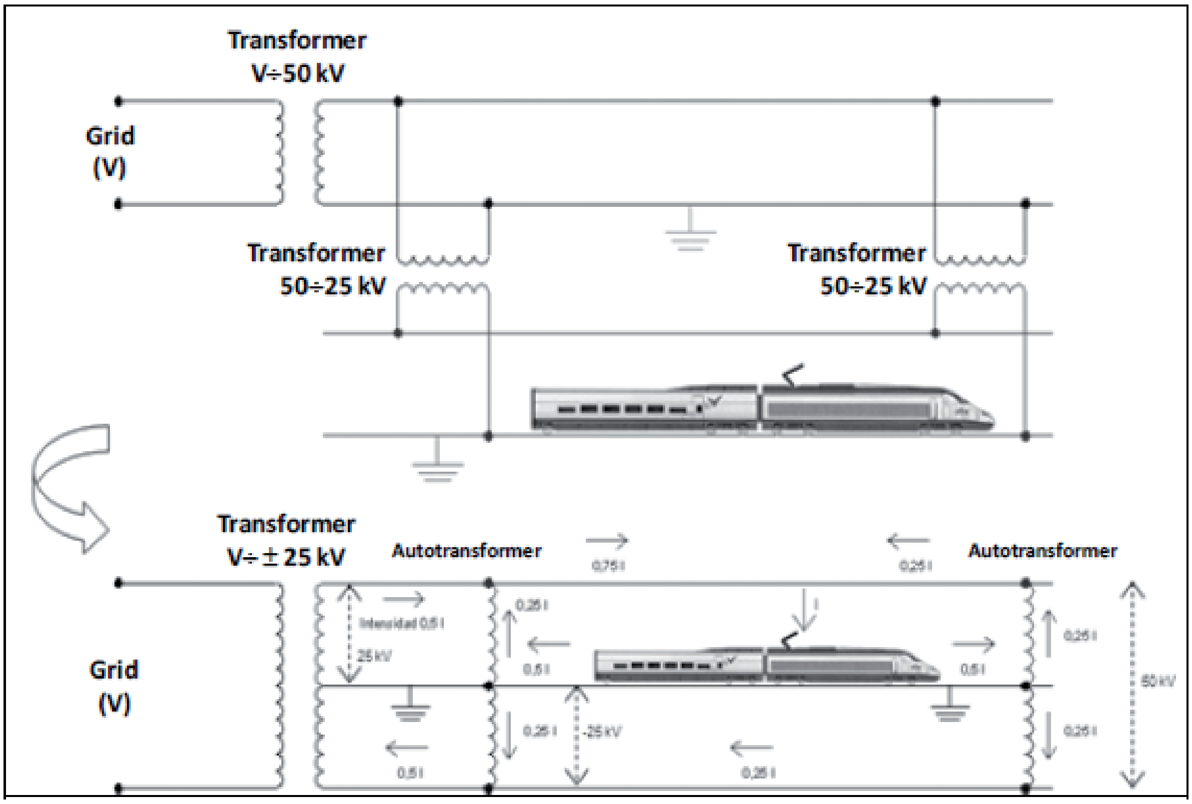


Figure 10. AT system. (Source: Author).

5.3 Coaxial cable

As shown in Figure 11, the coaxial cable feeding system features a coaxial cable laid along the track. Every several kilometers, the inner conductor is connected to the contact wire and the outer conductor is connected to the rail. The cable itself is very expensive but the conductor layout is simple, making it ideal for use where space is limited. Japan is the only country that uses this system (the Tokyo sections of the Tohoku Shinkansen and Tokaido Shinkansen).

In comparison to the overhead line, the coaxial power cable has extremely small loop impedance. Therefore, the load current is boosted in the coaxial power cable from the connection with the catenary. This results in a rail current distribution similar to that of AT system, significantly reducing the inductive interference.

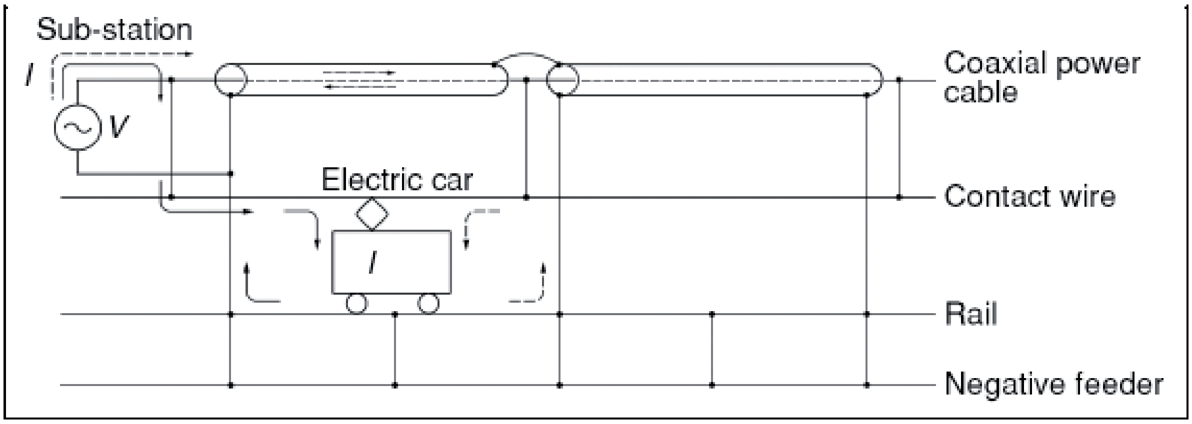


Figure 11. Coaxial cable feeding system. (Source: [2]).

6. TPS of the Madrid-Seville High-speed line: current situation and possible future

At the moment Adif is analyzing the operational advantages that the implantation of new power electronics (Full-Power Converter) in the electric traction substations of the High-Speed line Madrid-Seville would have. It must be considered that these substations have not been modified since the opening of the line (1992) and the current high traffic density recommends analyzing the impact that this new technology could have on the operation of the line.

The main action will be the determination of the design of the converter that allows the flow of energy between the public three-phase network and the single-phase rail network. The objectives are two: to analyze a possible improvement in the affection to the grid; to analyze possible improvements in the power supply to trains, including possible elimination of neutral zones in the catenary.

This High-Speed line has a 1 x 25 kV (simple feeding with rail return system). There are 11 substations with the average distance of 40 km between substations. Each substation has two single-phase 20 MVA transformers connected to the same two phases and three power outputs (one for each direction and another for a section of track in front of the substation itself of approximately 2 km). There are only neutral sections between substations, and there is no neutral section in front of the substation. The tracks are connected in parallel approximately every 10 km by use of disconnectors. The auxiliary systems of the line (technical building, radio stations for mobile communications, tunnel lighting and switch heater) are powered by transformers connected to the catenary.

The substations of this line are connected to the grid with nominal voltages between 132 kV to 220 kV, which have a comparably lower short-circuit power availability than very high voltage networks with up to 220 kV to 400 kV.

To determine the load of the electrification system, simulation software has been used to determine the power demanded or returned by each train at each time point and its position according to the parameters that influence the energy consumption. After determining the power and its position in the track of each train for each instant, and therefore the locations of the loads in the electric circuit, the behavior of the electrification system has been determined. Depending on the time step used in the simulator, the static or dynamic behavior of the system can be analyzed.

Figure 12 gives a schematic supply layout for three substations. The neutral sections at each substation are bridged with bypass switches, the ones between substations are in use.

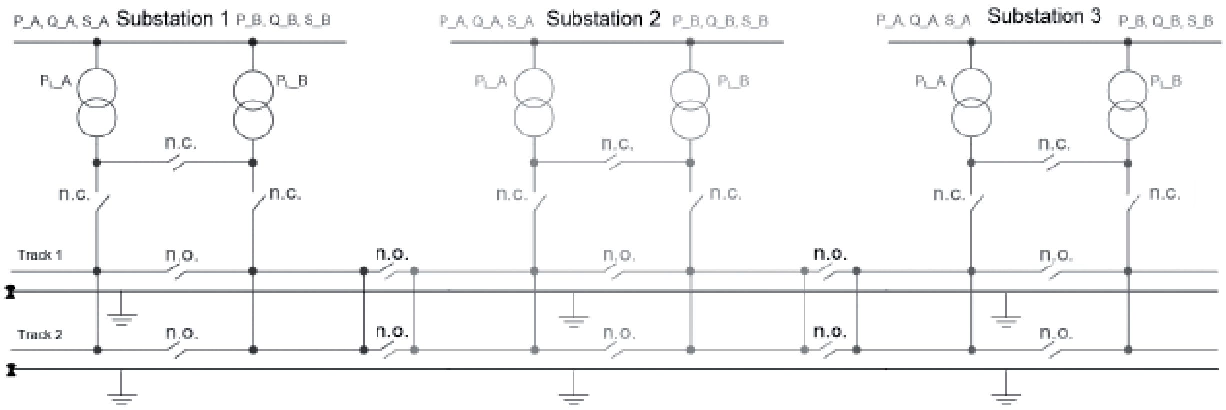


Figure 12. Normal scenario (now). (Source: Adif).

By using converters in each substation, interconnection of all supply sections is possible. Figure 13 gives a schematic supply layout for three substations with converters. All neutral sections at each substation and between all substations are bridged with bypass switches.

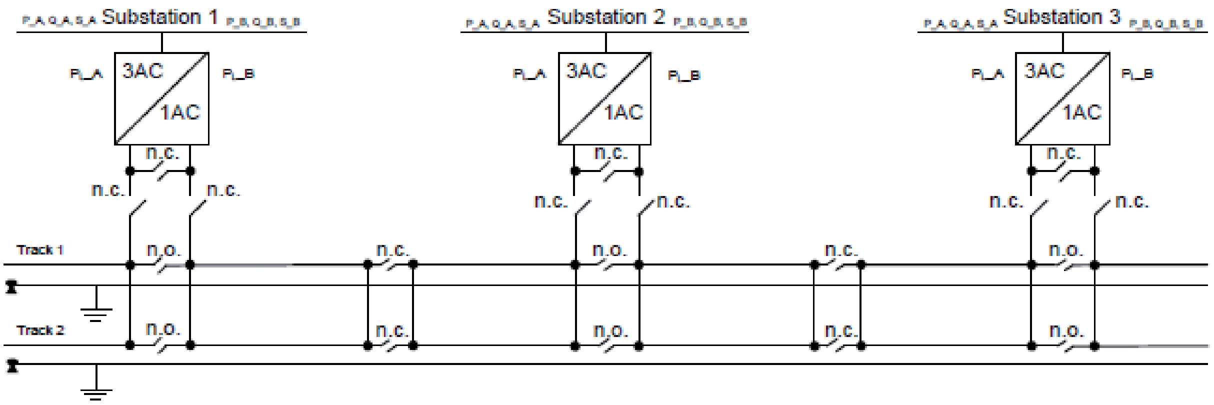


Figure 13. Possible scenario (future). (Source: Adif).

The simulations carried out conclude that the use of this type of converters in a line like the High-Speed line Madrid-Seville has different operative advantages:

- The imbalance factor on the grid is almost completely reduced. This aspect is important because a railway line like this one, with a high traffic density, connected to a transmission line with a not very high short-circuit power, can affect to a greater extent the electric network.
- In case of a degraded scenario (failure of a substation), the voltage drop in the line is optimally regulated (Figure 14).
- Disappearance of the neutral zones, which means that trains do not have to interrupt the traction bus through the route.

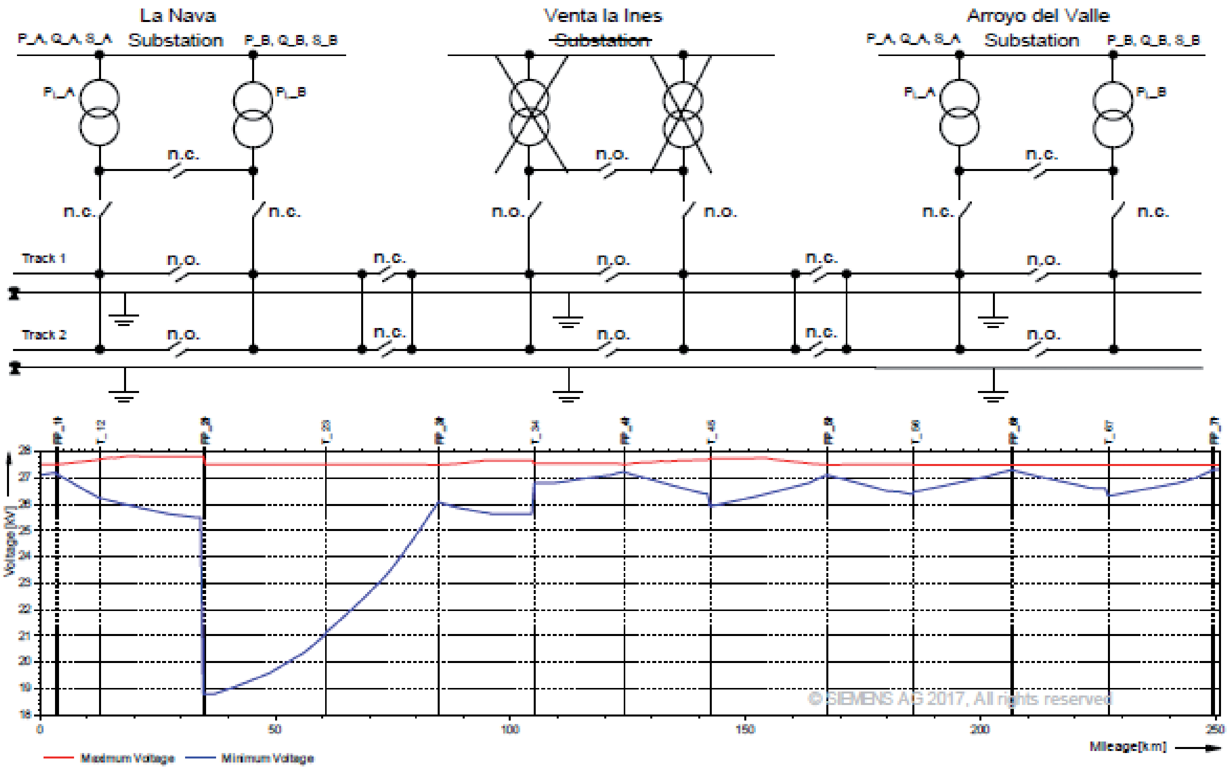


Figure 14A. Degraded operation mode and drop voltage. Conventional system. (Source: Adif).

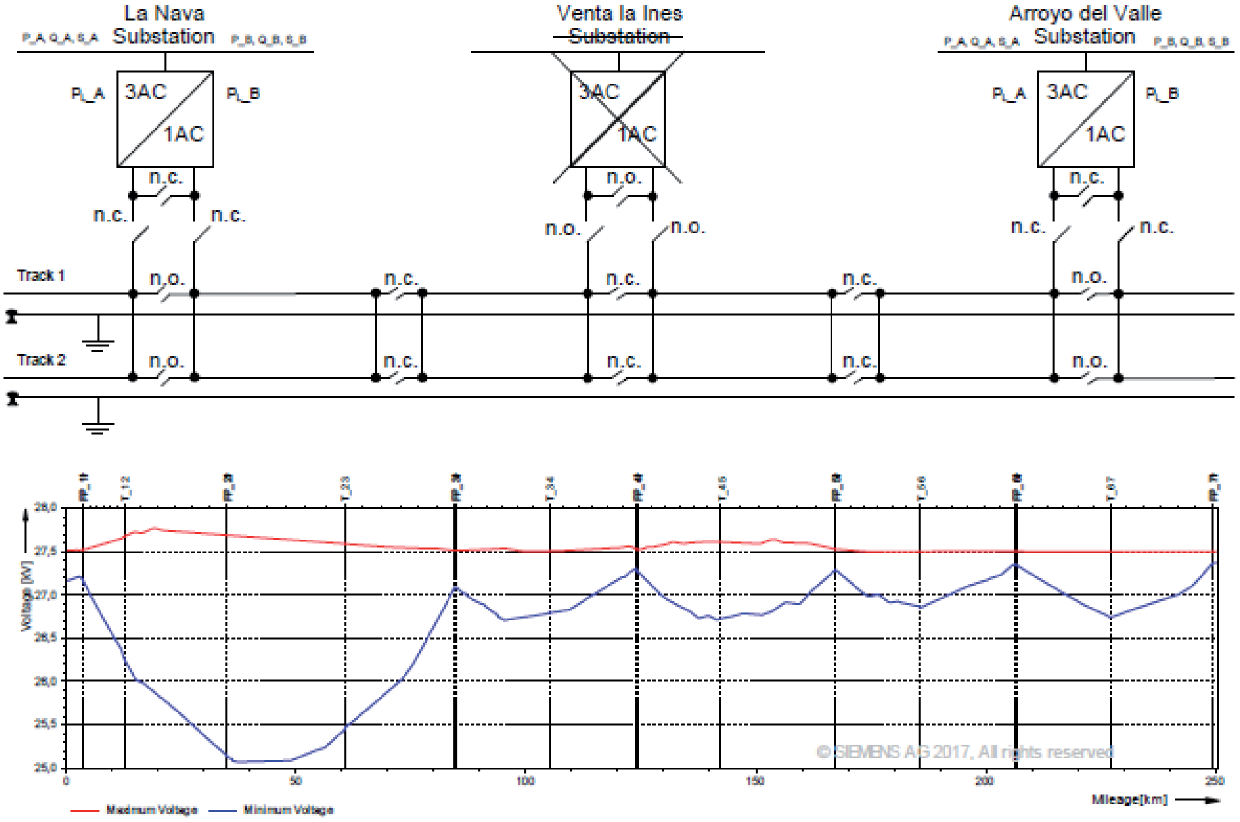


Figure 14B. Degraded operation mode and drop voltage. Converter system. (Source: Adif).

7. References

- 1) Abrahamsson, S.; Schütte, Thorsten; Östlund, S. (2012). Use of Converters for Feeding of AC Railways for All Frequencies. Energy for Sustainable Development.
- 2) Behmann, Uwe; Schütte, Thorsten (2012). Cost effectiveness of using static converters in power supply for 50 Hz railways. RTR.
- 3) Bhargava, Bharat (1999). Railway Electrification Systems and Configurations. IEEE.
- 4) Bin-Kwie, Chen; Bing-Song, Guo (1996). Three phase models of specially connected transformers. IEEE Transactions on Power Delivery.
- 5) Kiessling, Friedrich; Puschmann, Rainer; Schmieder, Axel; Vega, Tomás (2007). Contact lines for Electric Railways. Siemens (Second edition).
- 6) Krastev, Ivan; Tricoli, Pietro; Hillmansen, Stuart; Chen, Minwu (2016). Advanced electrification systems with static converters for ac railways. IEEE Electrification Magazine.
- 7) Laury, J.; Abrahamsson, S.; Östlund, S (2013). OPF for an HVDC Feeder Solution for Railway Power Supply Systems. KTH, Royal Institute of Technology.
- 8) Linhofer, Gerhard; Maibach, Philippe; Umbricht, Niklaus (2010). Static converters, dynamic perfomance. ABB.
- 9) Martinez Acevedo, Jose Conrado (2014). Analysis of disturbances produced by the electric power system of high speed railways fed at industrial frequency. Vía Libre Técnica.

- 10) Mingli, W.; Chengshan, X.; Fan, Y.; Zheng, T.Q. (2006). Performance and mathematical model of threephase three-winding transformer used in 225kV electric railway. IEE Proc.-Electr. Power Appl.
- 11) Mochinaga, Yoshifumi (2000). Recent Trend of Power Feeding Technique for Electric Railways. QR of RTRI.
- 12) Oura, Yasu; Mochinaga, Yoshifumi; Nagasawa, Hiroki (1998). Railway Electric Power Feeding Systems – Japan Railways. Transport Review.
- 13) Olofsson, Magnus (1996). Optimal operation of the Swedish Railway Electrical System. Electric Power Systems, Department of Electric Power Engineering, Royal Institute of Technology.
- 14) Pilo de la Fuente, Eduardo (2005). Optimization of the electrification of railways in alternating current (original in Spanish). ICAI.
- 15) Pilo, Eduardo; Mazumder, Sudip K.; González-Franco, Ignacio (2013). Smart Electrical Infraestructure for AC-Fed Railways with Neutral Zones. IEEE Electrification Magazine.
- 16) Plakhova, Mariia (2015). Development and implementation of a 2x25kV AC bivoltage traction system simulator for high speed railways. University of Oviedo.
- 17) Qunzhan Li (2014). New generation traction power supply system and its key technologies for electrified railway. Journal of Southwest Jiaotong University.
- 18) Usuka, Tetsuo; Nagasawa, Hiroki (2009). AC Power Supply system for railways in Japan. Railway Technical Research Institute, Japan.
- 19) Varju, György (2004). Comparison of the booster transformer and auto transformer railway feeding systems, Feeding features and induction to telecom lines. EMC York.
- 20) White, Roger D. (2007). AC Electrification Supply Design. IEE REIS AC 2007.

High speed rail superstructure

360.revista de alta velocidad
número 5 - junio - 2018. Pág 71 - 79

Automatic gauge changing for freight. The OGI project

Piqueras, Anibal

Saura, Joan

Paños, Francisco

TRIA, SA[1]

Abstract

Having different gauges on a railway network is a difficult challenge for planners and operators. There are a set of tools that can be used to minimize the connection problems at gauge frontiers. For HS, the solution has come through automatic gauge changing trains. For freight there are other solutions working now such as static wheel changing at frontier and third rail, but each has its disadvantages and there is need for enhancing the array of tools at the disposal of the infrastructure planners and decision makers at the transport ministry.

The Spanish rail authorities have decided it is time to develop an automatic gauge changing technology for freight trains, and they have put forward a program that aims at having a certified variable gauge wheelset by the end of this year. ADIF awarded the consortium formed by the Spanish companies TRIA and AZVI a contract to develop and homologate two variable gauge wheelsets for wheel diameters of 920 mm and 760 mm, that mount on Y21 bogies in the case of the 920 mm wheel and on vehicles carriages in the case of the 760 mm wheel.

At this moment, the prototype train has done 20 thousand kilometres on the HS and conventional networks of the 250 thousand required for final approval. Previously, the wheelsets have passed bench tests and the gauge changing tests.

Keywords: *track gauge, automatic gauge changing mechanism, changeover facilities, rail freight transport*

---

[1] *Piqueras, Anibal. Ingeniería y Técnica del Transporte. apiqueras@triaingenieria.com*
Saura, Joan. Ingeniería y Técnica del Transporte.jsaura@triaingenieria.com (corresponding author)
Paños, Francisco. Ingeniería y Técnica del Transporte TRIA, SA. fpanos@triaingenieria.com (corresponding author)

1. Introduction

The decision of constructing the main Spanish railway network to a 1668 mm track gauge continues being today one of the major challenges for the railway operation in Spain. Both freight and passenger trains find significant difficulties when it comes to operating in two different track gauges, mainly from/to the border between France and Spain.

In the case of passenger trains, automatic gauge changing trains and changeover facilities have been highly developed due to the decision taken by the Spanish Government of using the standard 1435 mm track gauge in HSR network. The need of facing the challenge of the coexistence of two different track gauges in the main passenger network (HSR lines and conventional lines) gave priority to the stakeholders to find the solution to this problem.

However, as almost the whole new HSR lines are not designed for mixed use and freight traffic, the need to solve the gauge problem has been forgotten in time for the freight trains. This brings us today to the situation that Spain has one of the lowest percentage rate of rail freight modal share[1] (both inland and international freight). This lack of competiveness of rail freight transport is due, among others, to the different track gauge and the fact of not having developed variable gauge systems for freight wagons.

The Spanish infrastructure manager ADIF, being aware of the problem, published a call for tenders in 2015 to develop an automatic gauge technology for freight trains. The state of the art and future benefits of this new technology will be analysed in this paper.

2. State of the art

2.1 The different track gauge and existing solutions to the problem

It is known that the challenge of different track gauges has existed since the first constructions of railway lines. The main problem of the break of gauge is that a train cannot pass from one track gauge to another and this implies, therefore, a railway border. We can find one of the best examples between the track gauge of the Spanish conventional network (1668 mm) and the track gauge of France and most European countries (1435 mm).

Different solutions have been implemented historically to solve this problem, to eliminate these railway borders, we could list the next types of solutions that are valid for freight trains:

a. Interchange of wagon axles ad wheels, as found in Cerbère/Hendaye borders
b. Transfer of goods, helped by large cranes that transfer the containers from one wagon to another, as found in Portbou/Irun borders
c. Using tracks of three or more rails so that it exists continuity for one of the track gauges, as found in some railway sections in Spain
d. Using automatic gauge changing trains (a solution yet not very extended for freight traffic)

The two first solutions (a and b) require big areas to place the necessary tracks, factories and cranes to carry out the interchange. Apart from that, these two solutions entail an inconvenient in terms of time-consuming process.

---

[1] *European Court of Auditors (2016). Rail freight transport in the EU: still not on the right track.*

The third rail (c) implies solutions of great technical and operational complexity, that affect the existent services in the modified line and requires a complete upgrade of signalling, catenary and operational systems.

The last solution (d) is very extended in passenger trains and services but it still has not been successfully developed in freight trains. One of the main reasons why it has not already been implemented in freight trains deals with the axle load and, more accurately, with the total weigh of the train.

2.2 The axle load and the weight of the train - different automatic variable gauge systems for passenger trains and freight trains

The existent technology of automatic gauge changing mechanism for passenger trains entails to release the weight of the wheels to allow them to adapt the new position. That means that the weight of the train is not supported by the wheels during the change operation and it is the changeover platform that support the weight of the train while it passes through the changeover facility.

The axle load of most passenger automatic variable gauge trains in Spain is about 17-18 t/axle (e.g.Talgo 250 is 18 t/axle[2]). Consequently, the total weigh of a passenger train that passes through a variable gauge facility is about 250-280 tonnes for a simple train and no more of 500-600 tonnes for a multiple-unit train. These variable gauge passenger train weigh is much lower than the axle load of a freight train.

Nowadays freight rolling stock is designed for the standard 22,5 t/axle. This means, in the end, 25-32% increase compared with the previous passenger example of 17-18 t/axle. This variation and the fact that a 700m freight train could carry 1500-2000 tonnes or even more, depending on the track conditions, make it difficult to extend the current variable gauge technology to the freight trains.

To develop a variable gauge technology to higher axle load and heavier trains, the first objective is to be able to adapt the wheels to the new position without releasing the weight of the train from wheels. This entails a significant difference between the actual automatic variable gauge technology - valid for passenger trains - and the automatic variable gauge technology valid for freight trains.

2.3 The OGI project

The OGI project that is being developed by the consortium TRIA-AZVI will mean an important step in terms of a new concept for the freight operation rail services and also in the construction and gauge adaptation of new rail lines in Spain.

This project consists on the development of two new automatic variable gauge systems for freight trains and its certification according to the European standards. Both solutions are based on the OGI technology developed in the 70's for which has been necessary an important labour of re-engineering to adapt it to the current time.

The two new variable gauge wheelsets are being certified for two different types of freight wagons and wheel diameters: 920 mm and 760 mm. The 920 mm wheel diameter is the one that mounts the standard bogie Y21 (valid for 1435-1668 mm gauge tracks) which is similar to the European standard bogie Y25 (just valid for 1435 mm gauge tracks). This type of bogie is the most extended all around Europe. The 760 mm wheel diameter is used in wagons for transport of cars.

---

[2] *Corporate catalogue - TALGO*

In the homologation process a wagon series Sgnss (4 axle wagon, 920 mm wheel, bogie Y21, for transport of containers) and a wagon series Laaers (2 axle half wagon, 760 mm wheel, for transport of cars) have been used.

One of the keys of this OGI wheelset refers to the fact that it is easily interchangeable. This means that it does not matter the series of wagon we are interested to equip with the OGI certified wheelset 920 mm or 760 mm.

2.3.1 The homologation process of the OGI variable gauge system

The homologation process for a variable gauge system is highly regulated in the EU and it also exists in Spain a great experience in the national variable gauge regulation due to its different 1668 mm track gauge.

The legislation followed in the OGI project has been the Spanish ETH ("Especificación Técnica de Material Rodante Ferroviario: vagones"). The certification plan was in accordance with ETH and not with TSI because "variable gauge system" was an open point at the moment the homologation process started.

The process to homologate a variable gauge system consists mainly on:

- The design review.

In the design review are analysed the functionality of the system including calculations of the components according to the EU standards (e.g. wheels or axles), the locking mechanism, the thermal effect of braking on the variable-gauge systems, the maintenance plan and the study of functional reliability RAMS.

- Laboratory tests - fatigue tests to validate the locking system.

The laboratory tests have consisted on the determination of the fatigue strength of a variablegauge running gear. In our case, the bench consisted on a rotational bending test where they were applied dynamic and static forces at a speed of 50km/h. The 10 million of cycles where divided in 3 stages, in accordance with the next table:

Table 1. Phases, cycles and forces of the variable gauge system laboratory test

Phase	Number of cycles	Vertical force [kN]		Transverse force [kN]	
		Static	Dynamic	Static	Dynamic
I	6·106	P	± 0,5 P at 4 Hz	0	± 0,3 P at 2 Hz
II	2·106	P	± 0,6 P at 4 Hz	0	± 0,36 P at 2 Hz
III	2·106	P	± 0,7 P at 4 Hz	0	± 0,42 P at 2 Hz

* Prepresents the static axle load corresponding to the maximum payload

* Frequencies reflected is for a 50 km/h speed.

- On-track tests

It is mandatory to test at least four different variable gauge wheelsets for the on-track and inservice tests which means that two-axle wagons are not admitted. In the OGI project, as we have seen before, a bogie Y-21 Sgnss wagon series was equipped with four 920 mm wheelsets and a Laaers wagon series (composed by two half-wagons) was equipped with four 760 mm wheelsets.

On-track tests consisted in 500 automatic gauge changeover processes run without inspection at the nominal passage speed and under the intended operational conditions. In the end of these tests the locking system had to be checked in terms of back-to-back dimensions, forces during automatic gauge changeover process and measurements of wear of components. These tests are called "on-track" tests in the ETH regulation. Once these "on-track" tests were successfully finished, "in-service" tests were allowed to be performed.

- In-service tests to validate the system in operational and real conditions

In-service tests may cover the following distances in three different phases:

- Phase I: 50.000km in 1668 mm or in 1435 mm track gauge. No automatic gauge changeover processes are allowed in this phase.
- Phase II: 50.000km in both 1435 mm and 1668 mm track gauges (at least a 20% of this distance may be covered in any of the track gauges). In this phase 50 automatic gauge changeover processes will be performed.
- Phase III: 150.000km in both 1435 mm and 1668 mm track gauges gauges (not less than 25% of this distance will be covered in any of the track gauges).

If results of phase III are positive, the technical approval will be extended to the variable gauge system. However, It exists a phase IV one 400.000km or 4 years have passed since the start of in-service tests. By the end of this phase IV, the technical approval of the system may be extended without any limitations.

2.3.2 The OGI changeover facility

The changeover facility consists of a platform whose components may be classified in three different groups:

Table 2. Components of the OGI changeover facility

Group	• Elements
Foundations	• Structural concrete • Steel anchor plates • Plastic pods (for rail screws)
Track and guided components	• Base plates • Rails • Check-rails • Check-rails bearings
Unlocking system	• Unlocking central bearings

The total length of the OGI changeover facility is about 33 metres. The reason why it is required such a large length deals with the automatic gauge changeover operation. In passenger variable gauge trains the length of the platform changeover facility is usually between 14-15 meters because the changeover process is done in just one step for both right and left wheels. Wheels change simultaneously to its new position and new track gauge.

In the case of the OGI operation changeover facility, the process is carried out in two steps. The first step is the changeover operation of one wheel of the wheelset and the second step is the changeover operation of the other wheel that belongs to the same wheelset. Due to this operation, the exerted axial forces to the axle is minimized, what brings more stability to the process.

3. Experimental analysis

3.1 European Rail freight Corridors - the problem of the different track gauge in Spain

Regulation (EU) 913/2010, adopted by the European Parliament and the Council of Europe, enacted the "establishment of international rail corridors for a European rail network for competitive freight, with the overall purpose of increasing international rail freight attractiveness and efficiency".

A list of 9 corridors are included in the Regulation, and Spain has 2 Core Network Corridors established by the Trans-European transport Networks (TEN-T): the Mediterranean Corridor (nº6) and the Atlantic Corridor (nº4).

One of the main challenges of both Corridors are the different track gauges between all the European countries integrated in the Corridors that feature the 1435 mm standard UIC gauge, and Spain, where coexist the 1435 mm track gauge (used on high-speed lines) and the 1668 mm iberian gauge used on the conventional main network (freight transport lines).

As reflects the Implementation Plan of the Mediterranean Corridor RFC 6[3] , "the lack of standard gauge in most of the Spanish sections of RFC6, prevents from dispatching international direct rail freight trains, and forces to car load changing manoeuvres, which penalizes rail transportation".

Several projects have been proposed in order to provide the standard gauge to Spain to the conventional lines. Projects such as the third rail and establishing new standard UIC gauge lines try to solve the different gauge issue. In this sense, the global projects of "The implementation of the standard track gauge between Castellbisball (Barcelona) and Almería" and the "Bobadilla - Villaverde Bajo - Implementation of UIC track gauge" are the greatest challenges faced by the Spanish infrastructure manager ADIF and the Transport Ministry to achieve this purpose.

3.2 The intended purpose of having a freight automatic variable gauge system

3.2.1 Track gauge migration in railway networks

In Spain, the migration to the standard UIC gauge 1435 mm is a priority for the core European network[4] . The final report of the study on the Mediterranean TEN-T corridor reflects that the rail share of the freight flows in the corridor's market today is about 14,6% and that implementing the Corridor, the rail share will grow from 14,6% to a 27,1% by 2030. This shows the strong potential for international rail traffic development on the Mediterranean Corridor,

---

[3] *RFC 6 - Mediterranean Corridor Implementation Plan TT 2015/2016*

[4] *Mediterranean Core Network Corridor Study. EU. Dec 2014*

especially in Spain. Furthermore, it is said in the report that "implementing the Corridor could potentially shift about 33-41 million tons per year from road to rail (about 2,3-3 million trucks/equivalent)".

The variable gauge system will permit the gradual adoption of the standard 1435 mm gauge in conventional lines in Spain. The idea to extend the standard gauge to the conventional lines requires a transition between the actual 1668 mm track gauge and de final 1435 mm track gauge. To achieve this objective it is necessary to implement the automatic variable gauge technology both in passenger trains and in freight trains to avoid services to be affected during the implementation.

This forecast data for different scenarios of the Rail Freight Corridor implementation means that disposing the standard UIC gauge in Spain will permit to promote freight traffic from/to Europe and that OGI project is one of the keys to achieve this objective.

3.2.2 Interoperability

Having a variable gauge system for freight trains will also promote the interoperability of rail networks of different track gauges in Europe. The European railway network is integrated by the different national networks with different track gauges. We find the most important breaks of gauge in the 1668-1435 mm track gauge border and the 1435-1520 mm track gauge border. The freight automatic variable gauge system will permit to improve the conditions for the transport of goods by rail between the different track gauge national networks in Europe.

Furthermore, as the track gauge problem is not only found in Europe, this new interoperable system could also be implemented in other border sections such as the China-Russia border for freight trains from Europe to China. The new trans-Eurasian rail freight services are growing very fast in the latest years and several break of gauge are found in their routes. A variable gauge system, such as the OGI system, will permit to reduce the transit time in the border sections.

3.2.3 New rail network planning tool

The development of this R&D project will provide the infrastructure planners and decision makers a new tool for the railroad and service planning.

The new freight automatic variable gauge system in Spain would permit, as an example, that freight trains run in both the conventional 1668 mm track gauge network and the 1435 mm HS track gauge network. This would mean a new concept of rail freight traffic in Spain and a good opportunity to develop a new mode of transport where freight trains could benefit from the existent HS network.

3.3 Examples of the possible OGI system implementation in Spain and long-term impacts

In this section, two possible scenarios will be analyzed where freight variable gauge system services could be implemented in Spain and its long-term impacts: the Mediterranean Corridor in the Vandellós-Castellon section and the Pajares Pass.

3.3.1 Example 1: The Mediterranean Corridor in the Vandellós-Castellón section

In the current year 2017, the Ministry of Public Works and Transport has decided that the Castellón - Vandellós section will be migrated to the standard gauge[5]. This decision implies a big effort between the operators (Renfe and the private companies) in terms of adaptation

---

[5] *Ministerio de Fomento press release (02/06/2017)*

of its current rolling stock material, along with the rail infrastructure manager ADIF and its implementation plan.

The most important issues during and after the implementation of the Castellon-Vandellós section in standard 1435 mm uic gauge are:

a. The continuity for the long-distance passenger rail services in standard gauge
b. The continuity for the medium-distance and regional passenger rail services in standard gauge
c. The continuity for freight rail services in standard gauge

Due to the non liberalisation of the rail passenger market in Spain, both a) and b) issues may be tackled by the incumbent operator Renfe. On the other hand, as rail freight services are liberated from 2007, it is not an easy task to solve the third issue c) as many stakeholders are now involved.

Furthermore, the continuity for the passenger long-distance trains is also assured at a significant proportion. This is because most long-distance Renfe services in the BarcelonaValencia corridor are already provided by automatic variable gauge trains. However, regional services and freight services may be analysed more carefully.

In the next Table 3 the operational services are presented in the Castellón-Vandellos section.

Table 4. Operational services in the Castellón-Vandellós section.

Source: Adif (freight operations), Renfe (passenger operations)

N.	Section	Daily Mediumdistance/ regional passenger services	Daily freight trains from Monday to Friday (average)	Daily Longdistance passenger services	Total services/day	km
1.1	Vandellós - L'Aldea/Amposta/ Tortosa	32 (16/direction)	12 (6/direction)	28 (14/direction)	72 (36/direction)	37
1.2	L'Aldea/Amposta/ Tortosa - Castellón	8 (4/direction)	12 (6/direction)	28 (14/direction)	48 (24/direction)	116
					Total length	**153**

As we have said before, long-distance passenger services are already provided by automatic variable gauge trains (both Renfe series 130 and 121). Thus, there would not be impediments to give continuity to the 28 long-distance services showed in Table 4 .

In reference to the medium-distance/regional passenger services we find two subsections with large differences in terms of the total amount of services. On the one hand, in the VandellósTortosa we find 32 regional services; on the other hand, in the Tortosa-Castellón we just find 8 regional and medium-distance services. It exists such a difference of services between these subsections because Tortosa is the last stop of the medium-distance R16 line which connects Barcelona and Tortosa. In these two cases, a migration in the existent rolling stock material will be mandatory to allow services in standard track gauge. It is noteworthy to say that it already exists some Renfe series such as the S449 that already has the standard gauge "preinstalled" in the coach.

For these S449 trains the adaptation to the new 1435 mm track gauge will be easier and cost-effective.

To achieve the continuity of the 12 daily freight services in the standard gauge scenario and/ or during the implementation plan, there is a need for an automatic variable gauge freight system, that is to say, the OGI project.

The OGI system will permit the migration from the 1668 mm track gauge to standard 1435 mm track gauge by implementing the variable gauge technology in those wagons that run in the rail section that is being migrated from one gauge to another. Moreover, changeover facilities should be installed in the borders of the section that is being migrated, so that traffic could be running during the migration process.

This variable gauge system permits to replace in wagons the fixed wheelset for the OGI variable gauge wheelset without much difficulties. Thanks for this "interchangeable" feature, the issue of finding different typology and series of wagon running in the Castellon - Vandellós section (or any other) is almost solved.

3.3.2 Example 2: The Pajares bypass section

The project of the new Madrid-León-Asturias mixed traffic high-speed line includes the construction of a new bypass in the Pajares Pass in the section León - Asturias, between La Robla (León) and Pola de Lena (Asturias).

This bypass aims to eliminate the main bottleneck in railway communications between Asturias and the rest of Spain (except the communities of Cantabria and País Vasco). Nearly 2/3 of all the railway freight traffic from/to Asturias run through the current Pajares line with strong limitations to capacity due to the existing 83km single track, the curves with low values of radius and the steep gradients.

The new 49,7 km bypass (30 km less distance than the actual track) includes two single 25-km tunnels that will permit to reduce the transit time, and consequently increase the frequency of both freight and passenger trains.

By adopting the standard track gauge in this section, the same issues would be found in relation with the continuity of the actual traffics, as we have seen in the previous Castellón - Vandellós example:

a. The continuity for the actual long-distance passenger rail services in standard gauge

b. The continuity for the actual medium-distance and regional passenger rail services in standard gauge

c. The continuity for the actual freight rail services in standard gauge

The following table shows the actual railway services in the section Oviedo-León:

Table 5. Operational services in the Oviedo - Leon section.

Source: Adif (freight operations), Renfe (passenger operations)

N.	Section	Daily Mediumdistance/ regional passenger services	Daily freight trains from Monday to Friday (average)	Daily Longdistance passenger services	Total services/day	km
1.1	Oviedo - Pola de Lena	76 (38/direction)	42 (21/direction)	8 (4/direction)	126 (63/direction)	**31**
1.2	Pola de Lena - Puente Los Fierros	22 (11/direction)	42 (21/direction)	8 (4/direction	72 (36/direction)	**12**
1.3	Puente Los Fierros - La Robla	8 (4/direction)	42 (21/direction)	8 (4/direction	58 (29/direction)	**71**
1.4	La Robla - León	4 (2/direction)	42 (21/direction)	8 (4/direction	54 (27/direction)	**26**
					Total length	**140**

The three issues could be solved in the same way that in the Castellón-Vandellós section. The actual long-distance passenger services run under automatic variable gauge systems. Therefore, the first a) issue may be easily solved by constructing a new changeover facility in the end of the new 1435 mm track gauge section.

In the case of the medium-distance and regional passenger rail services we find the main challenge between Oviedo and Pola de Lena and also between Pola de Lena and Puente Los Fierros. These regional services and freight services may be analysed carefully.

For freight services, considering that freight traffic moves from its current running section through Pajares Pass to the new double-tunnel track section in 1435 mm track gauge, the new tool of the automatic variable gauge system would permit to give continuity to these trains.

4. Conclusions

The main conclusions that can be drawn from this paper are listed below:

- To solve the existent connection problems due to the different track gauge in Spain, there is a need to have an automatic gauge changing technology to be implemented in freight wagons. This solution already exists in passenger trains, but it is of the utmost importance to develop it in freight trains.
- An automatic variable gauge system, such as the OGI system, will permit to increase in Spain the percentage of rail freight services from/to Europe. This will entail a change in the modal share to the rail mode, in the search of more sustainable mobility and transport.
- The new automatic variable freight system R&D project will provide the infrastructure planners and decision makers a new tool for the railroad planning. This new tool would permit to enhance the interoperability between different track gauge networks.

5. References

- European Court of Auditors (2016). Rail freight transport in the EU: still not on the right track. http://www.eca.europa.eu/Lists/ECADocuments/SR16_08/SR_RAIL_FREIGHT_EN.pdf
- TALGO - Corporative catalogue.

  https://www.talgo.com/upload/corporativo_es.pdf
- Rail FreightCorridor6 - Mediterranean Corridor Implementation Plan TT 2015/2016.

  https://www.railfreightcorridor6.eu/RFC6/Public/RFC6_CID_Book5_2016-17_24-02-2017.pdf
- European Comission (2014). Mediterranean Core Network Corridor Study

  https://ec.europa.eu/transport/sites/transport/files/mediterranean_study_0.pdf
- European Commission (2016).Second Work Plan of the European Coordinator

  https://ec.europa.eu/transport/sites/transport/files/2nd_workplan_med.pdf
- Ministerio de Fomento (2017). Press release.

  http://www.fomento.gob.es/NR/rdonlyres/467250EB-2727-48A3-AC84409C-3D205019/143270/170602NPvisitaobrasVandell%C3%B3sTarragonav4.pdf
- prEN 17069. Railway applications - Systems and procedures for change of track gauge - Part 1: Automatic Variable Gauge Systems
- REGULATION (EU) No 913/2010 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 22 September 2010 concerning a European rail network for competitive freight
- Rail Freight Corridor 6 - Annual Report 2016

  https://www.railfreightcorridor6.eu/RFC6/Public/RFC6_Annual_Report_2016.pdf
- Rail Freight Corridor 4 - Annual Report 2016

  http://www.atlantic-corridor.eu/es/corredor-es/ra-2
- Piqueras, A., Saura J., Paños F. (2016). Análisis técnico-económico de la aplicación de tecnologías de ancho variable para la mejora de la interoperabilidad de los tráficos de mercancías entre la red de ancho estándar y la red de ancho ibérico. Repositorio Universidad de Comillas.

  https://repositorio.comillas.edu/xmlui/handle/11531/12527
- Plan Director de Infraestructuras para la movilidad de Asturias 2015-2030. Gobierno del Principado de Asturias. Consejería de Fomento, Ordenación del Territorio y Medio Ambiente.
- Estudios Informativos para el desarrollo de la Red de Alta Velocidad en Asturias. Tramo Pola de Lena-Oviedo (2015). Ministerio de Fomento.

  https://www.fomento.gob.es/MFOM/LANG_CASTELLANO/DIRECCIONES_GENERALES/FE-RROC ARRILES/_INFORMACION/ESTYPRYTRM/EILAV_POLAOVIEDO/doc.htm

6. Acknowledgements

- Authors would like to acknowledge the support of the Spanish infrastructure manager ADIF R&D Department.

Safety and signalling systems

Fundación de los Ferrocarriles Españoles

360.revista de alta velocidad

Safety and signalling systems

International Congress on High-speed Rail: Technologies and Long Term Impacts

Ciudad Real (Spain), October, 2017

Escuela de Caminos Ciudad Real UCLM
Escuela de Ingenieros de Caminos, Canales y Puertos de Ciudad Real
UNIVERSIDAD DE CASTILLA-LA MANCHA

Safety and signalling systems

360.revista de alta velocidad
número 5 - junio - 2018. Pág 85 - 91

KEYNOTE

The role of the ertms users group in the consolidation of the ERTMS technical specification for baselines 2 and 3

Michel Ruesen
Jaime Tamarit

ERTMS Users Group

Abstract

The ERTMS Users Group is a European Economic Interest Group (EEIG) constituted by the Infrastructure Managers from Belgium, Denmark, France, Germany, Holland, Italy, Norway, Spain, Sweden, Switzerland and the United Kingdom. This group, founded in the middle of the nineties, working in tight cooperation with the European Signaling companies, grouped in UNISIG, has played a fundamental role in the consolidation of the ERTMS technical specification and its upgrading to Baseline 3, covering all requirements of the conventional railway system.

In this communication, two Managing Directors of the group show the complete process going from the proof of concept of ERTMS to the specification of ERTMS Baseline 3. This Baseline will be the basis for the deployment on the interoperable European railway network. Along this communication, the two managing directors will address the critical points encountered along the consolidation and the features envisaged along the enhancement of ERTMS with the Baseline 3 new functionalities.

1. The demonstration of the feasibility of the Railway interoperability in the nineties (EMSET)

In the middle of the nineties, Head of DG TREN Mr. Edigio Leonardi was commissioned with full political support for the consolidation of the interoperability of the Trans European Network for Railways. He decided to close the pure paper work launching a project within the Fourth

Framework Program with the aim of demonstrating at real scale the feasibility of Railway

Interoperability by the deployment of ERTMS. Spain offered the new high-speed line MadridSevilla to perform the experimental demonstration. 40 Km of this line between La Sagra and Mora stations and the CEDEX railway laboratories where allocated to this project, of EMSET acronym (European Madrid – Sevilla Eurocab Test). This project carried out under the umbrella of the ERTMS Users Group took place between the years 1.994 and 2.000.

This project was performed under the technical supervision of the European Commission, represented by Mr. Antonio Colaço, Mr. Emmanuel Parent de Curzon and Mr. Claudio Traverso. The main outcomes from this project where the following four:

I. The signalling companies of UNISIG, in a precompetitive phase, found a neutral place, of highest quality, to perform cross tests, fundamental for interoperability consolidation.

II. The UNISIG companies where able to cooperate closely at technical level producing a reference test lab constituted by 37 tools. This lab at CEDEX was later the seed for the reference laboratories, accredited for the certification of ERTMS Constituents.

III. Before testing on the line, the cross test between two companies going to the line was previously debugged in the CEDEX laboratory. Consequently, there were no incidents with the commercial traffic during more than one year of the tests on line. Lesson learned: When putting a commercial project in service, debug the project in the laboratory before starting deployment of equipment on the track.

IV. At the end of this project, the ERTMS Users Group, UIC and RENFE delivered in Madrid to the European Commissioner for Transport, Ms. Loyola del Palacio the first version of the System Requirements Specification. The SRS version delivered was the 2.0.0 .

2. The first consolidation of the technical specification (Pilot Lines): Facing the first difficulties

After the experimental feasibility test carried out in EMSET, the technical specifications were not yet backed by the feedback of the experience of the commercial projects. The first commercial project was the Spanish High Speed Line Madrid – Lérida (460 Km), awarded to the French Compagnie de Signaux, later acquired by ANSALDO.

To get feedback from first commercial operation, the ERTMS Users Group, with the support of the European Commission, started the deployment of ERTMS pilot lines in the six countries of the group plus Austria, Switzerland and Vienna.

Mr. Antonio Colaço supervised technically the pilot experiences. The main difficulties faced along this phase were the following:

I. UNISIG delivered a version of the System Requirement Specification containing hundreds of the so-called “Designer Choices”. A designer choice was a requirement not affecting interoperability, of free election by the project designer. This approach introduced a great uncertainty in the certification process. The Steering Committee of UNISIG named them also “Don’t care” requirements.

II. The feedback from the commercial experience produced a great number of "Change Requests". These were proposals for the improvement of the System Requirement Specification resulting from the commercial exploitation. Commercial lines in exploitation had to migrate continuously to get them aligned with the consecutive versions of the technical specification: SRS 2.0.0 -> SRS 2.2.2 -> SRS 2.3.0 -> SRS 2.3.0 D. This has been the case of the Spanish Railway Administration, forced to migrate more than 1.000 Km of lines and hundreds of trains.

III. The feedback from the exploitation of the Madrid – Lérida line showed that the harmonisation of human factors was an open issue. For the first time, different suppliers of ERTMS On-board equipment, operated on tracks equipped by different ERTMS suppliers. A comparative study of the different DMIs, performed by CEDEX, forced the ERTMS Users Group to launch a new working group on Human Factors. Three different approaches were needed to solve this open point with experts on human factors, not linked to the technical world (HEROE Project).

It is worthy to say that the operational interoperability is 50 % of the interoperability, the other 50% being the technical interoperability. Due to this unexpected issue, the ERTMS Users Group could not accomplish the operational harmonisation within the period of the pilot lines. It is only with the launch of the ERTMS Baseline 3, that the DMI specification becomes harmonised and its specification becomes mandatory. The Human Factors Working Group, transferred to ERA from the ERTMS Users Group, accomplished the operational harmonisation.

IV. The handover between the RBCs at the border between Holland and Belgium in the Dutch High Speed Line South caused a major problem. The trains ought to cross the border in Level 1, disconnecting from the handing over RBCs to solve the problem. This patch, also used in other countries, had a big impact on the exploitation of the High Speed line due to the speed limitation in Level 1.

UNISIG came with two proposals for the specification of the RBC handover, asking the ERTMS Users Group to choose the most convenient one. The election made was the basis for Subset-098 specification.

V. The starting of the commercial exploitation of the Spanish High Speed Line between Madrid and Lérida in ERTMS Level 1 was quite problematic due to the lack of reliability of the Eurobalises. This critical situation was solved thanks to an illegal duplication of fixed balises, placing contiguously two identical balises. If one balise was failing once every one thousand of readings, the two contiguous balises had a failure estimated in one every one million of readings, solving miraculously the problem. It is worthy to mention the role played by Mr. Jorge Iglesias in solving this crisis that could have strong political consequences.

VI. The Spanish High-Speed network includes the Madrid-Sevilla line, equipped with the German system LZB. A fleet of hundred high speed trains purchased by RENFE ought to operate on the ERTMS new lines and also on the Madrid – Sevilla line. This was forcing the Spanish Administration to request the STM for LZB, as proposed by UNISIG. At the end, major part of the train fleet was equipped with LZB onboard equipment operating independently from ERTMS because the LZB STM was not available. Today the use of STMs is optional and only trains equipped by Bombardier incorporate the LZB STM.

VII. The continuous evolution of the ERTMS Technical Specifications forced the Spanish Infrastructure Manager ADIF to undertake a continuous migration of their high-speed lines. Today all lines are aligned to the SRS version 2.3.0 D. New trains purchased by the Spanish operator RENFE will be equipped with EVCs of Baseline 3, the backwards compatibility between both consecutive versions is a challenge that will be verified at commercial level within the following years.

3. From Baseline 2 to Baseline 3 and onwards

Finalised the consolidation of the Baseline 2 with the pilot projects, The ERTMS Users Group (EUG) centred its activity on the deployment of the Trans-European Network based on the implementation of the new functionalities required, incorporated to Baseline 3 and on the specification of enhanced performances envisaged for future developments.

The Group management was adapted to the new mission profiles with new Managing Directors and additional members that joined the group.

The mission profile can be summarised into two main aspects:

1. To help the railways in applying ERTMS/ETCS in a harmonised and interoperable way, to enable the free flow of trains and a competitive railway.
2. To offer a platform for railway peers to share experiences and to consolidate their views.

Apart from the six initial members of the group (ADIF, DB, Network Rail, Prorail, SNCF Réseau and RFI), new infrastructure managers joined the group to share this mission: BaneNor, Banedanmark, Infrabel, SBB and Trafikverket).

3.1 Global approach: Common platforms with users, manufacturers and authorities

3.1.1 EUG RU ERTMS/ETCS Platform

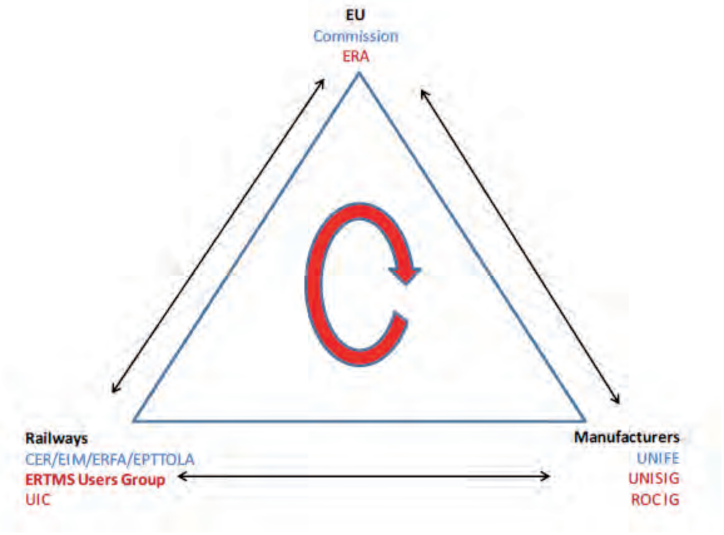
In order to collect the return of commercial experience, EUG invited the Railway Undertakings to join in a common platform set up in 2013 with the aim of sharing experiences and ideas in practical issues concerning specification, design, certification, installation, authorisation, performance, reliability, operation, tendering and maintenance of ETCS On-Board Units in rolling stock.

The partnership:

- Nine Railway operators are providing feedback from commercial exploitation: DB Fernverkehr, DB Cargo, MRCE, NS, ÖBB, SBB, SNCB, SNCF and Trenitalia. Renfe is also invited for the platform, but has so far not participated.
- Four European institutions linked to Railway exploitation are also engaged: CER (Community of European Railway and Infrastructure Companies), ERFA (European Rail Freight Association), EPTTOLA (European Passenger Train and Traction Operating Lessors' Association) and the UK based Rail Delivery Group (formerly called ATOC)
- Other stakeholders including European authorities like European Commission (EC) and European Railway Agency (ERA), suppliers, Notified Bodies (NoBos) and Reference Laboratories are invited on a case-by-case basis.

3.1.2 Role and Level

The mission results are circulated between the institutions having key roles in the consolidation and enhancement of Railway Interoperability through the use of ERTMS: European Authorities, (EC & ERA), Manufacturers of rolling stock (UNIFE) and signalling (UNISIG) and final users (CER, EIM, ERFA, EPTTOLA,ERTMS Users Group (EUG), Union Internationale of Railways (UIC).



3.1.3 The activities of the ERTMS Users Group within the global environment

The EUG focus its activity within this global frame on providing support within the different phases of the process of putting lines into service, considering aspects like:

- Analysis and design
- Testing
- Commissioning
- Operation
- Assessment

Specifically, this support provides help in areas related with the assurance of the interoperability of the projects, like Engineering guidelines, TSI clarifications, Mitigation measures, Correction of application guides and Correction of TSI trough the CCS. All these activities are fundamental for the processes of tendering, deployment and putting in service railway lines where the sharing of experience is a fundamental benefit.

Apart from this assessment, the EUG plays a central role in the enhancement of ERTMS with new functionalities making the system more attractive by its efficiency and performances. The innovation activities of the Shift2Rail programme (in which EUG also participates) are aligned with this enhancement.

3.2 Baseline 2: suited for its purpose

The specifications of Baseline 2 are the reference for the first railway lines put in service in almost all European countries. All these lines have meanwhile converted to Baseline 2 SRS 2.3.0d. They have reached very satisfactory performance after several migration processes with the lines in operation, before reaching the final consolidated specification. Baseline 2 specifications (SRS 2.3.0d) are included in TSI CCS since 2008.

Just to mention some of the most significant projects in service based on SRS 2.3.0d

- France: Paris – Strasbourg - Mulhouse - Basel
- Italy: Torino – Milano – Padua and Milano – Roma – Napoli,
- Spain: Madrid – Lérida – Barcelona – Figueras (French border) and Madrid – Valencia – Albacete - Alicante
- Switzerland: Mattstetten - Rothrist, Lötschberg Tunnel
- The Netherlands: High Speed Line South, Betuweline
- It must be underlined that some of these corridors have several infrastructure suppliers and different ERTMS On-board suppliers. This commercial experience is the confirmation of the railway interoperability at commercial level.

4. Development of Baseline 3

Baseline 2 covers the needs for new (high-speed) lines; however, additional functionality is needed for the conventional network, e.g. level crossings and shunting.

Conventional network puts high requirements on capacity, which Baseline 2 does not cover like Braking curves, enhanced capacity of the radio link and additional functionality needed for migration from class B to ERTMS (Level 1 Limited Supervision). Therefore, the European Commission took a decision in 2008 to develop Baseline 3.

Baseline 3 not only contains all functionality of Baseline 2, but also tackles around 300 errors/ambiguities of Baseline 2 and assures backwards compatibility, i.e. Baseline 3 trains can run on Baseline 2 (2.3.0d) tracks.

Apart from the enhancement of Base line 2 with the cleaning process and the assurance of the backwards compatibility, the Base line 3 incorporates new functionalities required for deployment of the Trans – European Network, mainly:

- Level 1 Limited Supervision
- Level crossings
- Train categories
- Improved braking performance
- Packet switching technology for radio communication
- Online key management.

4.1 Evolution of Baseline 3

As stated before, the specification phase began in 2008 with SRS 3.0.0 and the first release incorporated to the European Technical Specification of Interoperability of the Control Command System (TSI CCS) was the release 3.3.0, incorporated in 2013.

A maintenance release of Base line 3 (SRS 3.4.0) was incorporated to the TSI in 2014, replacing the previous SRS 3.3.0. In 2016, a second release (SRS 3.6.0) was added to TSI CCS, adding the possibility to use packet switching technology (GPRS) and on-line key management.

4.2 Present and future developments

At present, the ERTMS Users Group is working in coordination with the European working groups and authorities in applied research areas included in the global approach of the Shift2Rail programme:

- Use of satellite navigation provided by Galileo to the incorporation of Level 3 applications (with hybrid Level 3 as a first step) and virtual balises
- Automatic Train Operation, (with and without train driver)
- Alternative for GSM-R
- Improved IT-Security
- Enhancement Braking Curves harmonisation

Facilitating the solutions to difficulties encountered along the deployment of Baseline 2 and Baseline 3, the EUG pays special attention to crucial aspects like:

- Backwards compatibility, assuring the compatibility of Baseline 3 trains with Baseline 2 lines
- Migration from conventional class B systems to ERTMS and migration within ERTMS
- Improved performance (RAMS) and
- Lower Life Cycle Costs, making ERTMS more competitive with conventional class B systems.

Safety and signalling systems

360.revista de alta velocidad
número 5 - junio - 2018. Pág 93 - 103

Test and Certification of Railway automation and digitalization approaches (Rail 4.0)

Meyer zu Hörste, Michael
Asbach, Lennart
Hardi, Hungar
Lemmer, Karsten

German Aerospace Centre (Deutsches Zentrum für- Luft- und Raumfahrt – DLR)[1]

Abstract

The current industry mega-trend of digitalization –sometimes called “Rail 4.0”, too - will lead to a major change in railway automation. The automation of railway operations will increase significantly in the coming years and will help to meet the challenges of the railways. Which of the approaches discussed currently will come into operation is depending on technical, operational and legal requirements. Many projects in Germany and Europe are currently dealing with the automation of railway operations. Therefore pilot projects can be expected in the near future. Improved concepts for railway operation e.g. the level 3 of the European Train Control System (ETCS) using moving block, on-board integrity supervision or virtual train sets will result in even higher requirements to the accuracy and reliability of the localization. Hence some major issues are still to be solved. It is among others: Precise and reliable localization, e.g. based on sensor data fusion based on Global Navigation Satellite Systems (GNSS), reliable communication technologies robust against obsolescence, obstacle detection systems fulfilling civil and criminal law as well as cost requirements and cyber security. The contribution will focus on the last aspect: Tests need to be developed to ensure the safe and reliable behaviour of automated trains under normal and disturbed conditions. In current practice, highly automated, mechanically evaluated tests need strictly formal test cases which are unambiguously defined. If these preconditions are not met, the automation effort is high, so that it is often better to only aim for a partial solution. To get test cases of high quality, the choice of tools used for test defining them is crucial: The tools must restrict the engineer to prevent unclear definitions being introduced. They must enable the domain expert to write test cases which can readily be translated for use in the test bench.

Keywords: Train Control, Testing, Conformity, ERTMS, ETCS, ATO

---

[1] *Meyer zu Hörste, Michael. Institute of Transportation Systems. Email: Michael.MeyerzuHoerste@dlr.de. (Corresponding author)*
Asbach, Lennart. Institute of Transportation Systems. Email: Lennart.Asbach@dlr.de
Hardi Hungar. Institute of Transportation Systems. Email: Hardi.Hungar@dlr.de
Lemmer, Karsten. Transport and Energy. Email: Karsten.Lemmer@dlr.de

1. Introduction

Today functional testing becomes more and more important, as complex train control systems (e.g. ETCS) offer a lot of functionalities but specifications, written in natural language. The tests are necessary to reach real interoperability but also offer a chance to reduce the field tests. DLRs research aims to reduce time and costs of test runs and increase the scope of the tests to make the final European admission as easy as possible.

Main focus of the research is testing for ETCS and the standardized interfaces for field components and interlocking (e.g. Eulynx, DB NeuPro). In the field of ETCS especially the onboard (OBU) and radio-block centre communication is focused. Due to the increasing demand for ATO this gets more and more into the focus of DLR s research as well.

With its laboratory RailSiTe. DLR is researching tomorrows testing methods in the railway domain. The paper will show the current approach of conformity testing in a test bench for the entire system. It includes interfaces for OBU, RBC, Interlocking and field elements to allow a comprehensive test execution and evaluation. By now the test bench is able to connect to almost every RBC in the world using internet protocols and special tools for GSM-R adaption.

The goal of the research work on OBU testing is attained when the following scenario is reality: All ERTMS tracks are virtually rebuild in a test bench. All European operational scenarios have been created and formalized. Assuming this, there cannot be any requirement in the system specification, which is not tested in one of the operational scenarios. If there is any, it can be removed from the specification, because it is obviously not needed. By testing every new onboard system on the entire European network virtually, in all operational situations, interoperability is proven and inherits conformity.

Today s testing process of conformity tests and interoperability tests offers high potential for implementation. A first step has been made by formalizing some selected tracks in the European network. By creating an RBC-Proxy, DLRs laboratory is able to behave exactly like certain RBCs, which allows testing of tracks without a real RBC connection. This is important to accelerate the process massively.

The expected result is the reduction of field and manual parts of the laboratory tests to a minimum. At least on the functional layer the field tests can be replaced by laboratory tests more or less easily. This contribution shows first ideas to reuse existing tests for testing ATO systems.

Right now there are already many ATO systems, or at least driverless systems, available and in service. The real challenge will be the implementation of those systems on main-line tracks. Having a look to the current train control systems, on the functional level there is only a very small gap between driver-operated and automatic trains.

The track-train communication is safe and ready for automatic train operation. Thus we can assume that the ATO for main line tracks will be an add-on module for train control systems (red box in figure 1). This includes track-side and on-board systems. This contribution shall focus on the testing of the on-board part, the driver-replacement.

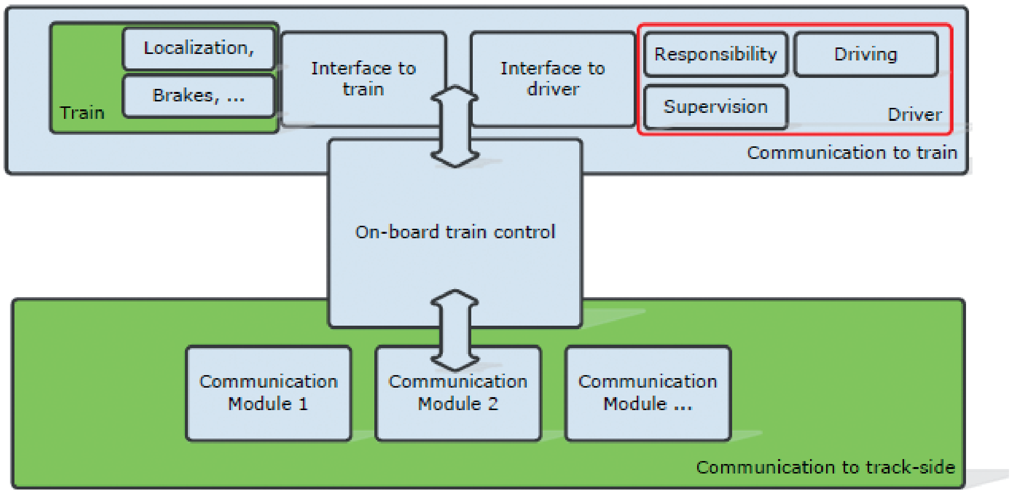


Figure 1: Overview of current setup of on-board train control

Figure 1 shows a basic setup of todays on-board train control. The communication to the track-side is available by many technologies, in case of ETCS one of the communication modules could be the GSM-R channel. The train interface is also already connected to the train control system. Information and commands are exchanged, e.g. braking, doors, pantograph or localization information. In most cases the driver has to control the traction lever according to the information of the train control system. From a technical perspective the main task of the driver is supervision and responsibility for the train. The first part of this contribution will focus the influence of ATO systems to the current testing methodologies.

2. Testing on-board ATO modules

2.1 Changes in testing due to ATO

In the current implementations the on-board conformity tests are running 100% automatically. Since the driver machine interface (DMI) has to be part of the test, because of the current law, it is controlled by a robot. The train interface is controlled by the laboratory. In case of ATO functions of the system-under-test these laboratory modules (robot and train interface) can be replaced by the ATO module. At least if the ATO module is tested together with the on-board train control (in case of ETCS this is called European Vital Computer, EVC) the test could be performed in the same manner but with less effort on the laboratory side. The interesting question is what additional descriptions in the test specification are necessary to ensure the proper testing of the ATO module. Since we assume a black-box test for the ATO module, the first inspection has to be done on the interfaces of the module. Today there are many ideas for these interfaces; one could be the reuse of the interface between the DMI and the EVC, in case of ETCS. As long as there is a radio connection available between the track-side and the onboard, all necessary information will be provided. For example a new movement authority is transmitted to the train control system and the ATO simply has to follow this information. Today it is shown with permitted speed and many more values on the DMI. I.e. for all information, which is available via the current DMI, the functional test specification of ETCS

can be reused. Of course, the evaluation methods have to be adapted, because now not only the train control behaviour has to checked but also the ATO behaviour (train interface unit, driver machine interface etc.). If there are any additional inputs for the ATO-module, maybe cameras to replace the driver s eyes or diagnostic systems to ensure the proper function of the train, they have to be included by adding new test cases in the test specification. Since this should be relatively easy for all pure digital -interfaces, like diagnostic information of the train, it will be more complex for kind of analogue information like the camera image. If, maybe due to legal restrictions, there is a certain reaction required based on the output of image recognition, this will lead to very complex tests. The challenge is the number of possible inputs. It gets even worse, if state-of-the-art methods like machine learning are used for the image recognition. In the current authorization process the certification of a machine learning systems seems to be impossible. Current research approaches are trying to solve certification issues for machine learning algorithms by using watch-dogs, which assure that the machine cannot leave certain boundaries. But right now there is no efficient methodology found.

Even if there is no approach with artificial intelligence and self-learning algorithms the certification and testing of obstacle detection systems will be challenging. Spot checks are possible, but their results are questionable. The number of possible inputs is simply too high (32*10^12 possibilities for a full HD image with 24bits colour depth), to ensure a proper test coverage. Two solutions are conceivable at the moment. Maybe the easiest approach is to ensure no obstacles by fencing the track completely. If a safe fencing is not productive, another possibility would be the acquisition of real data from real train journeys for testing obstacle detection. I.e. a proper way would be to equip all current trains with cameras and use this material (ideally commented) for testing the obstacle detection systems. By this approach, at least a very real test can be assured and the result is more resilient. This procedure can be used for different sensors and is not limited to video based sensors. The following section will exclude the testing of video based obstacle detection and will focus on functional interfaces.

2.2 Solution Approach

Components and systems for railway applications, especially for safe applications, need to be tested comprehensively before taken into operation. These tests have different aims: they can be used to show that a system fulfils the relevant specification, the foreseen operational profile or safety requirements. All these tests need to be described to be performed in the field or to be formalized to be executed in a lab. Both need a formal definition and description to prove the correctness of the results.

The approach used for the conformity tests for ETCS can be extended for operational and safety lab tests, operational field tests and fits very well for testing ATO-systems, too. It may, as demonstrated in the section on interface conformity for digital track side systems, even applied to partial standardisations. The principle method of the generation of the test sequences can be used for the different types of tests. The optimization criteria as well as the rules for the parameterization differ for the different kinds of tests. If the same approach for the formalisation and parameterization is used, the lab environment can be used for any type of test, with, however, some adaptations to be made to achieve sufficient flexibility.

Test case generation and the test sequence construction profit substantially from the application of formal approaches. This field features a variety of languages, methods and tools. Present-day solutions cover only part of the needs of practice, but show potential to be much more useful if applied in a carefully designed process employing adequate formalisations.

3. Test Generation

3.1 Conformity Test Sequences

The group of eight suppliers of train control systems called UNISIG (Union of European

Signalling Companies) have specified the ETCS by writing the SRS (System Requirements Specification SRS [1]) and produce and deliver ETCS components for different railway undertakings and infrastructure manager like Deutsche Bahn AG (German Railways) in Europe. Before these may come into operation, their conformity with the SRS and their interoperability has to be proven with lab tests as stated in previous section.

The issue 3.1.0 of the conformity test was specified by the Subset-076 Working Group (CEDEX, DLR, INECO, MULTITEL, RINA). This specification includes the Hardware-in-the-loop tests of the ERTMS/ETCS on-board equipment [1], the method of their creation [2] and the method of test [3]. The specification of the reference lab architecture [4] completes the set of specifications.

Some of the UNISIG companies are implementing products or just finalising them. Some test trips (field tests) have been performed successfully on the German pilot line Jueterborg-HalleLeipzig. Currently the last steps of implementation of the tests are performed and the latest version will be published very soon. The test target has been defined, generic test cases have been specified and the method for the generation of the final test sequences has been developed and realised. For the better understanding of the following sections these steps will be discussed shortly in the following.

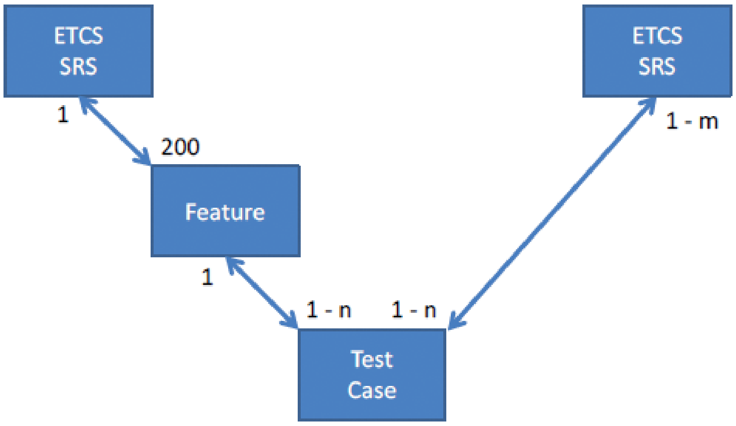


Figure 2: Relation between SRS and Test Sequence

The defined target of test sequences is to test each requirement of the SRS at least once. Firstly, to organise and reduce the amount of requirements of the SRS, more than 200 features have been identified. After this, the required positive and negative test cases have been created for each feature. Totally more than 1700 test cases have been generated. Equivalent test cases for different ETCS modes or levels have been merged to reach a first optimisation and reduction of number of test cases. This means that test cases which are applicable for different mode-level combinations are described as only one test case, if the feature was not dedicated to a specific mode-level combination. Important is the testing of the feature itself. Just for clarification, each ETCS mode is an operational state of the on board equipment and the ETCS level is an overall degree of the usable functionality of ETCS. For the execution of the tests the test cases are concatenated to 775 test sequences, which all start at the powering on of the on-board equipment and end with the no power mode. The test subsequences are concatenated due to their start- and end-conditions to reach a consistent sequence of system states. The test sequences have been optimised to reach the lowest degree of redundancy of testing. Parts of a test are only executed twice or more if they are needed to reach a state which has not been tested yet. Up to the test sequences the specification is completely generic. At this stage the variables and parameters are filled with values, though some remain to be set dynamically during execution of the sequence. So the test sequences could be understood as operational test trips. The relation from the SRS up to the test sequence is shown in Fig. 2.

Finally the fundamental structure of the test sequence should be clarified. Each test sequence simulates a test trip by stimulating the on-board equipment via the black-box-interfaces. In addition, the SRS-conformant reactions of the on-board equipment are defined in each test sequence. The reactions and the stimulating events are bound to the interface where they should be observed and evaluated or raised. Essentially the test sequences consist of the stimuli and the expected reactions of the on-board equipment. In the test sequence one stimulus or reaction is represented by a test step. Fig. 3 shows the structure of a test sequence.

The 775 test sequences contain up to several hundred test steps and their execution in realtime in the labs need up to several hours. A time rafting testing is not possible due to the fact that the real time behaviour of the ETCS component is tested. As mentioned above the test sequences are implemented in the reference labs. Some of the test sequences have been executed successfully, but the stated problems of duration and unstable inputs on the user interface by the human being show that automation is needed. As soon as an input is missed or incorrect the complete test sequence must be repeated.

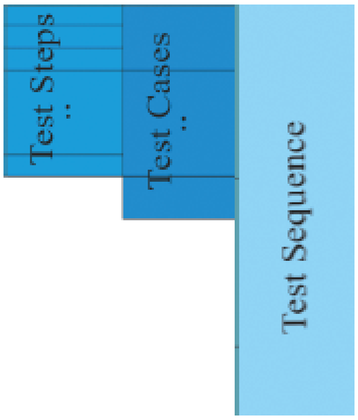


Figure 3: Symbolic structure of a test sequence for conformity testing

The tests defined by this method are documented in the ETCS subset 076. These are used to proof the conformity of the constituent European Vital Computer (EVC) which is the core of the on-board unit.

3.2 Operational Test Sequences

The conformity test sequences which have been discussed in the previous section fulfil the purpose to show that an application is realising the specification sufficiently complete. They do not claim to be operationally reasonable. Thus, a railway undertaking tendering ERTMS/ETCS systems need to check whether these fulfil their operational requirements. These tests are a separate set of test sequences at the moment. They need to be defined by a similar methodical approach as shown above, but they have need to fulfil more requirements: The test sequences must represent the most typical or important scenarios of the operation of the railway. They need to show the fulfilment of the European requirements as well as the national add-ons.

The approach is to use the same test cases for the operational test sequences as well. Some specific test steps and test cases are added to represent operational aspects which are not represented in the technical tests. The method for the generation of the test sequences is used, too. So the test cases consist out of technical and operational test steps. The test cases are concatenated to a realistic test sequence. This operational test sequence is formalised and filled with parameters according to the same rules as the technical test sequences.

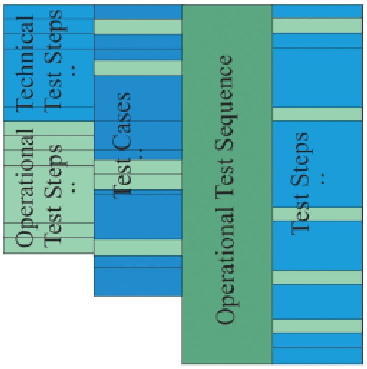


Figure 4: Structure of a test sequence for operational testing

The main difference is that the definition of the test sequences is not optimized to fulfil the specification requirements using the shortest possible sequences. The optimisation criteria are here to find as much relevant or important regular or disturbed operational scenarios to be tested.

The parameterization of the test sequences is done according to the operational environment. Average or standard parameters are typically used for this purpose. Extreme or rare parameters are to be avoided.

The main advantage is here, that the basic database of the test cases as well as the testing environment can be used for both as well as the execution environment in the lab.

3.3 Perspective: ATO Test Sequences

No matter, if the implementation of the ATO is done by integrating new features in the onboard train control or adding a module replacing the driver, both approaches can be tested on functional level by the approaches described above. Due to the experience with the ETCS test specification, the conformity approach is recommended. In contradiction to the current implementation of the conformity tests a closer relationship to the real operation is useful to avoid too artificial scenarios in the lab tests. This is even more important if a field test shall be executed with the same scenarios.

Looking a little bit into the details this means that for example the correct driving behaviour (acceleration/deceleration) of the ATO module can be tested similar to the current implementation of the tests for the braking curve behaviour of the ETCS EVC. The laboratory would stimulate the driving action by sending a movement authority (or whatever signals necessary) and the virtual position of the traction lever can be evaluated by the laboratory. This methodology can be transferred to all functional interfaces of an ATO-module or an integrated ATO system easily, as long as there is digital information available. Beside the already available methodology there is another advantage of this lab-testing approach: Assuming the number of tests will increase massively to reach a certification for an ATO module, the high efficiency of laboratory tests will reduce the effort to a productive level.

During a migration phase there will be a lot of data available from the non-automatic operated trains. These data can be easily reused for testing the ATO systems. I.e. huge number of tests can be defined by using the real data and the trust in the systems can be maximized and a comprehensive testing for ATO on functional interfaces can be assured easily.

4. Interface Conformity of digitalized track-side equipment

4.1 Digitalization of track-side equipment

The general approach of constructing executable test sequences from generic test cases is equally well suited for the conformity test of track-side equipment as it is for testing on-board components. What is different with track-side equipment is the lesser degree of standardisation: Interfaces and even functional architectures may differ, depending on the manufacturer. To improve compatibility, an approach currently employed by the German Railways is to incrementally specify the interface behaviour of equipment components. I.e., only some of the interfaces of an interlocking system are specified (and shall be tested), while others remain to be considered somewhen in the future. The reason is that it is easier to specify and implement a standard version of the focus interface, and not having to come up with a formalisation and re-implementation of the full system.

The downside is that this approach faces an inherent difficulty when it comes to testing. To drive the focus interface (and observe the correct interpretation of messages received over it), it is usually necessary to have access to (all the) other interfaces. Specification is easier by far--one can “internalize” the uncontrolled interfaces by subsuming everything in internal behaviour of an automaton.

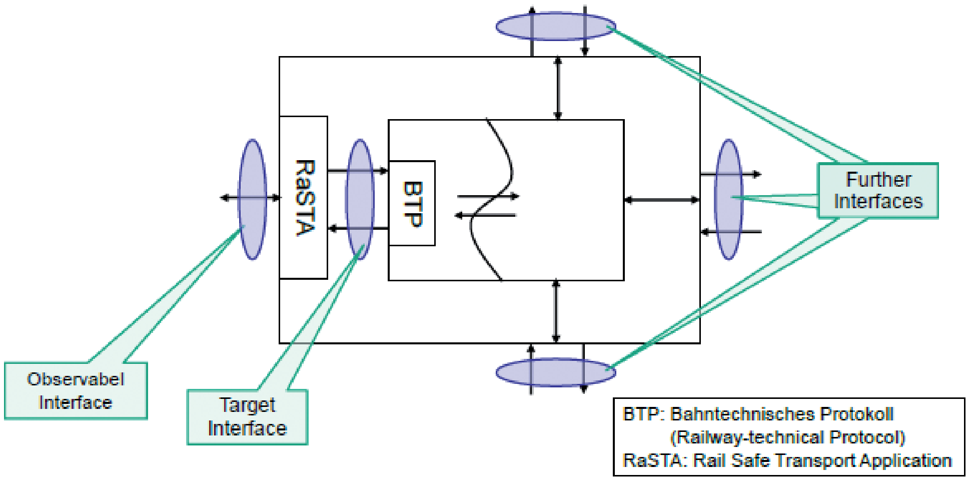


Figure 5: Schema of a system with four interfaces, of which one is to be specified

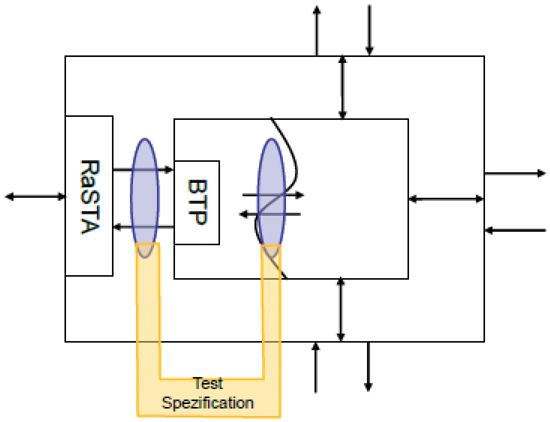


Figure 6:

Schema of a system with additional virtual internal interface as the specification view on the specific interface to be tested

Fig. 5 shows a schematic view of a system where the interface in focus is on the left, and is shown with some detail. The specification addresses the functional level of the Rail Technical Protocol (RTP) and abstract from the concrete implementation of communication through the Rail Safe Transport Application (RaSTA) which utilizes an Ethernet connection. Fig. 6 gives the specification view, where the additional virtual internal interface is added. Telegrams on the focus interface are related to messages and observations on this virtual internal one. Technically, these messages and observations are just actions of UML state machines which make up the specification. They reflect actions happening on the other (masked) interfaces, but are not formally related to them.

This works for the specification, but for testing it can of course not be done in terms of the internal specification interface but needs the real behaviour on the masked interfaces. I.e., test cases and operation have to take the view of Fig. 5, while their derivation must refer to Fig. 6. The problem is acerbated by the unavailability of a precise relation between internal and masked interfaces. In current practice, such a relation does not even exist: There are considerable differences between the masked interfaces (whose standardisation is yet to be initiated) in the different manufacturers' implementations of the devices as already mentioned above.

4.2 Solution Approach

Differences between the solutions of different manufacturers call for integrating them into the test process in some way. Our solution relies on the assumed ability of the manufacturers of bridging the gap between (virtual) internal messages and commands and externals. The envisaged test architecture is depicted in Fig. 7.

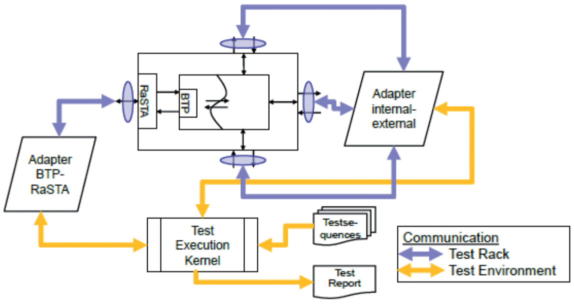


Figure 7: Components of the test architecture.

The test rack adds two components to the test object:

- Adapter internal-external: The manufacturer shall provide a module which translates between internal and masked interfaces. For its realization, interface drivers, simulators, or existing test interfaces accessing internals of the device may be used. Even a test engineer performing manual steps may be integrated via a suitable interface component.
- Adapter RaSTA-RTP: This module must be provided by the test laboratory.

The test rack serves to provide the test object with an interface which is on the same level of abstraction as the specification. The remaining components of the test architecture are rather standard:

- Test Execution Kernel: The kernel controls the test execution, i.e., it initializes the test objects, starts test sequences (including parameter completion in advanced scenarios), protocols the results, performs corrective actions (breaks and restarts if necessary) and generally monitors the execution. The kernel will be partially automatized.
- Test Sequences: A data base with test sequences sharing the characteristics with those of the on-board tests.
- Test Report: A data base for detailed result data and accumulated reports.

5. Validation

To be able to make qualified assertions of standard conformance, several arguments have to spelled out. On the one hand, the correctness and completeness, resp. sufficient coverage, of the test cases with respect to with respect to the specification has to be checked. This involves techniques and methods form the domain of model based testing. Currently, manually derived test suites are evaluated for their suitability. In future enhancements of the overall approach, also test case generation from the specification models is intended to be considered.

Adapter design and validation will have to cope with the common problems of crossing abstraction levels (namely atomicity and timing issues as well as value concretizations). For the internal-external adapter a monitoring concept which observes its operation dynamically is envisioned. The user interface of interlocking systems provides many information about internal states and thus qualifies as an adequate point of observation.

6. The Role of Formal Methods

Formal methods are used increasingly in the testing process of rail equipment, albeit slowly. They make their entry in one of two ways.

The first is via a formalisation of specification. This has a benefit in itself, as ambiguities, omissions and inconsistencies are reduced when a specification is formulated in a more or less rigorous notation. Examples from practice are the specifications of track-side equipment mentioned in the previous section. While currently limited to single interfaces, it is intended to specify the functional behaviour of, e.g., interlocking systems completely. Another example is the development of formalising the ERTMS/ETCS SUBSET 026 in the project openETCS.

Besides the effect of improving the quality, a formalised specification is of course a necessary basis for many further process enhancements. With respect to testing, formalised requirements permit at the very least a systematic derivation and coverage analysis of test cases, and in favourable scenarios even the automation of test case generation. For the latter, commercial tools are already available (rttester, rhapsodyATG, etc.), though these tools are not yet widely used in the rail domain.

The second way formal methods enter the testing process is in the automation of test construction. The elaborations in the preceding sections illustrate the highly nontrivial task of constructing a test set. In arranging test cases for the ETCS OBU, the Chinese postman algorithm has been used to find a first solution, which led to a stable test suite only after extensive work on parameter instantiation and calibration. Getting the timing, position and velocity parameters right proved to be rather difficult.

The construction of test sequences for track-side equipment is less difficult with respect to real-valued parameters, but on the other hand has to solve issues with testing generic systems which are to be instantiated to control a particular local arrangement of track elements. Testing is done on one or few sample layouts. To allow a flexible arrangement in test sequences, test cases should be formulated in a parametric style, permitting their instantiation depending on the needs of other test cases.

In both cases, formal languages for parametric test case specification are used, and algorithms for optimizing the arrangement into sequences are applied. These formal based solutions are still to be improved upon, as the currently available techniques have a limited scope and achieve non-optimal results.

7. Conclusion

The approach used for the conformity tests and operational tests for ETCS is proven in use and can therefore be extended for testing ATO functionalities and interfaces, too. As long as the interfaces are functional and digital the methodology can be reused easily. It may, like for interface conformity, even applied to partial standardisations. The principle method of the generation of the test sequences can be used for the different types of tests. The optimization criteria as well as the rules for the parameterization differ for the different kinds of tests. If the same approach for the formalisation and parameterization is used, the lab environment can be used for any type of test, with, however, some adaptations to be made to achieve sufficient flexibility.

Test case generation as well as test sequence construction profit substantially from the application of formal approaches. This field features a variety of languages, methods and tools. Present-day solutions cover only part of the needs of practice, but show potential to be much more useful if embedded applied in a carefully designed process employing adequate formalisations.

Future approaches for digitalisation of rail traffic management systems, train control systems or interlocking focus on a stronger modularisation and standardisation. Therefor the same approach should be applied to ensure conformity and interoperability of the modules of such digitalized components. By applying this proven-in-use approach the cost and time required for the tests can be reduced up to the final goal of zero field test .

8. References

- [1] ERTMS -- Test Sequences, SUBSET-076-6-3, Version: 3.0.0, 1/06/2016
- [2] ERTMS -- System Requirements Specification. SUBSET-026, Version 3.4.0, 15/06/2016
- [3] ERTMS -- Test Sequence Generation: Methodology and Rules. SUBSET-076-4-1, Version 1.0.2, 1/07/2016
- [4] ERTMS -- Methodology of testing. SUBSET-076-3, Version 2.3.1, 1/07/2016
- [5] ERTMS -- Functional requirements for an on-board reference test facility, Version:3.0.0, 15/06/2016

Safety and signalling systems

360.revista de alta velocidad
número 5 - junio - 2018. Pág 105 - 115

ATLAS: The road to Baseline3

Fernández Suárez, Enrique
Rodríguez, Antonio

Abstract

European Rail Traffic Management System (ERTMS) development started several years ago, with some rail manufacturers, infrastructure managers and trains operators, and the European Union.

Alstom was present from the beginning in this process, being the first to put in service the ATLAS solution for ERTMS in a High Speed Line (HSL) in Italy, the first freight line in The Netherlands, and the first high-density line with short headway in Switzerland.

The first set of ERTMS specification, stable and complete, is based in subset 026 with version called "2.3.0 D", also known as Baseline 2. This Baseline 2 was used in most of the European projects today in commercial operation, and in Spain in all the Adif (Administrador de Infraestructuras Ferroviarias), and Renfe (Public train operator) projects.

About Spain, there are several Adif lines in operation (Madrid-Barcelona, Madrid-Valencia, Madrid-Valladolid, etc…), and also Albacete-Alicante, the first to enter in commercial operation with ERTMS level 2, without fall back of level 1, in May 2014.

Also Renfe has installed in some of their trains the ERTMS baseline 2, in High Speed, regional and commuter train types.

Alstom has installed its ATLAS solution in Renfe S-100, S-104, S-114 and different type of CIVIA trains for commuter lines. Asltom has demonstrated experience in the ERTMS onboard equipment, with more than 5450 trainsets of 115 different types running with ATLAS solution.

Keywords: atlas, ERTMS, infraestructures,lines

1. ATLAS: The road to Baseline3 - (artículo de 10 páginas – versión 170913a)

European Rail Traffic Management System (ERTMS) development started several years ago, with some rail manufacturers, infrastructure managers and trains operators, and the European Union.

From that time, several projects have been deployed across Europe, and there are several high speed, regional, and freight traffic lines in operation, and hundreds of trains running with ERTMS in commercial operation. Alstom was present from the beginning in this process.

Alstom was the first to put in service the ATLAS solution for ERTMS in a High Speed Line (HSL) in Italy. After Alstom did it again for the first freight line in The Netherlands, and the first high-density line with short headway in Switzerland:

Date	Country	Line	Operational Speed	ERTMS Baseline	Line (km)
2005	Italy	Roma-Napoli	300 km/h	2.3.0D	220 D
2006	Switzerland	Maastesten Rothrist	200 km/h	2.2.2+	40 D + 10 S
2007	The Netherlands	Betuwe Route (Freight only)	120 km/h	2.3.0D	90 D

(S: single line; D: Double line)

The first set of ERTMS specification, stable and complete, is based in subset 026 with version called "2.3.0 D", and it is also known as Baseline 2. The complete list of documents and applicable versions can be found at:

http://www.era.europa.eu/Core-Activities/ERTMS/Pages/Set-of-specifications1.aspx

This Baseline 2 was used in most of the European projects today in commercial operation, and in Spain in all the Adif (Administrador de Infraestructuras Ferroviaras, Infraestructure manager), and Renfe (Public train operator) projects.

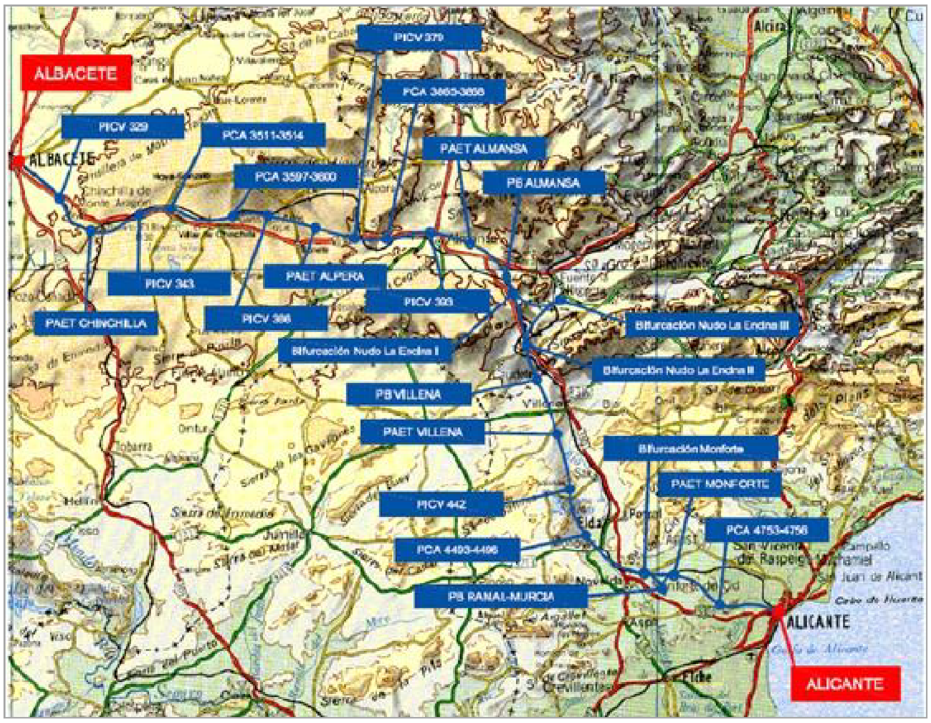
Here is the list of the most relevant European projects with ERTMS level 2, in commercial operation (in **bold** what is ATLAS solution from Alstom):

Date	Country	Line	Operational Speed
2005	**Italy**	**Roma-Napoli**	**300 km/h**
2006	Italy	Torino-Milano	300 km/h
2008	Italy	Milano-Bologna	300 km/h
2009	**Italy**	**Bologna-Firenze**	**300 km/h**
2007	**Belgium**	**L3**	**260 km/h**
2009	**Belgium**	**L4**	**300 km/h**
2011	Spain	Madrid Lleida	350 km/h
2012	Spain	Madrid Valencia	350 km/h
2014	**Spain**	**Albacete Alicante**	**300 km/h**
2009	The Netherlands	HSL Zuid	300 km/h
2012	**The Netherlands**	**Hanzelijn**	**200 km/h**
2006	**Switzerland**	**Maastesten Rothrist**	**200 km/h**
2007	Switzerland	Loetschberg tunnel	250 km/h

About Spain, there are several Adif lines in operation, some included in the table above, but I would like to highlight Albacete-Alicante, the first to enter in commercial operation with ERTMS level 2, without fall back of level 1.

Línea de Alta Velocidad Albacete-Alicante

Albacete – Alicante line belongs to Madrid – Levante HSL, and it is in commercial operation since May 2014. The most relevant characteristics of the line are:



- 160 km length with double track, from KP. 326,261 to KP 486,123 at Alicante terminal station.
- Maximum operational speed, tested with commercial trains is 300 km/h, even if the line is designed for 350km/h. With ASFA the maximum speed is only 200km/h.
- There are 22 technical buildings along the line to locate three main electronic interlockings Smartlock, in Bonete, La Oliva and Monforte del Cid, and two Radio Block Center (RBC), located at Bonete and Monforte del Cid.

Renfe installed in some of their trains the ERTMS baseline 2, in High Speed trains, but also in regional and commuter trains.

Renfe has also ERTMS installed in all the HS trains they have: S-102, S-103, S120, S-121, S-130 and S-730.

Alstom has installed its ATLAS solution in Renfe S-100, S-104, S-114 and different type of CIVIA trains for commuter lines.

ERTMS on board of CIVIA Renfe trains

Project includes the installation and commissioning of 123 trains to be used in Cercanías Madrid commuter lines. First train is in operation with ERTMS level 1 since Feb 2012.

The trains to be equipped were: 8 CIVIA I (in commercial operation), 34 CIVIA II (in commercial operation), 40 CIVIA III (in commercial operation), and 11 CIVIA III and 30 CIVIA IV under manufacturing at the time of the project.

Central car with new rack installed, cabling between cars, cabin and driver desks were modified:

Today, all units are ready for ERTMS commercial operation.

At the end, Alstom has demonstrated a big experience in the ERTMS onboard equipment, the integration in several train types, and the management of other national ATPs when required. In total there are more than 5450 trainsets of 115 different types running with an Alstom ERTMS onboard.

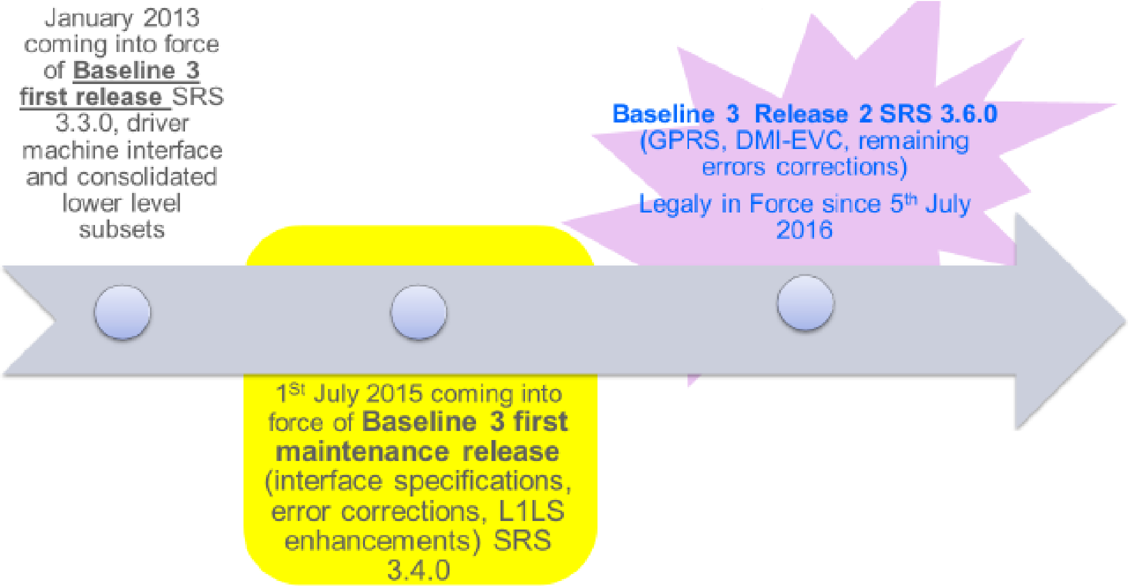
In the first implementations, Adif, Renfe and the Ministerio de Fomento created some functions (called National Functions) to cover Spanish specificities not included in the ERTMS standard. I mean, new functions not against the standard, but needed considering operational rules. This was a must because Spain was (and is) one of the European countries with more km of line and trains in commercial operation with ERTMS.

A short list of these National Functions is (first version of this specification was created in 2003):

- Función 20: Gestión separada de limitaciones temporales de velocidad según el nivel
- Función 24: Gestión del mensaje por defecto de eurobaliza

- Función 27: Gestión ERTMS del equipamiento ASFA independiente
- Función 77: Transición degradada desde nivel 2 a nivel 1 por pérdida de comunicación.
- Función 121: Inhibición de niveles disponibles (via external switch)
- Función 124: Gestión de la reacción de enlace
- Función 125: Operación en áreas ERTMS fuera de servicio

From those first implementations, the ERTMS specification has evolved to improve defects or misunderstandings found, and to implement new functionalities requested by Infrastructure Managers and Train Operators. The release 2.3.0 D evolved to version 3.0.0, 3.1.0, 3.2.0 and 3.3.0, and after:



January 2013 coming into force of **Baseline 3 first release** SRS 3.3.0, driver machine interface and consolidated lower level subsets

1st July 2015 coming into force of **Baseline 3 first maintenance release** (interface specifications, error corrections, L1LS enhancements) SRS 3.4.0

Baseline 3 Release 2 SRS 3.6.0 (GPRS, DMI-EVC, remaining errors corrections)

Legaly in Force since 5th July 2016

The Baseline 3 compared with version 2.3.0 D includes:

- 22 new functions and 32 Errors Change Requests (CRs) in SRS 3.0.0 (due to finalisation of Subset 108)
- 46 Errors CRs with SRS 3.1.0
- 73 CRs with SRS 3.2.0
- 30 CRs added and 2 removed with SRS 3.3.0

The major evolutions are the system version management, harmonised braking curves, the Limited Supervision, and the Level crossings management.

But other functions have been also included like:

- Speed restriction to ensure a given permitted braking distance
- Inhibition of revocable TSRs from balises in Level 2 or 3
- Redefinition of international train categories

- Shunting in level NTC areas
- Supervision of the safe radio connection
- Non-Leading input signal
- Cold movement detection
- Door control supervision
- Inhibition of the BG message consistency reaction
- Passive Shunting mode
- Lines under construction (Virtual Balise Cover)
- Changing Train Data from sources different from the driver
- Etc…

From release 3.3.0, the specifications evolved to 3.4.0 in 2015, also called **Baseline 3 Maintenance Release 1** (MR1). This MR1 includes:

- 28 CRs, with 21 error corrections and 7 enhancements
- 25 subsets of specifications affected
- Key changes on: L1 Limited Supervision (CR 1223) with toggle function to Select Swiss / Other countries; Increased SIL requirements on some DMI features

The changes implemented consider the following CRs:

Id number	Headline
0944	Data unit/resolution/size
1088	Subset-039 upgrade to Baseline 3
1104	Subset-094 upgrade to baseline 3
1109	error non-stopping areas (Follow-up CR 1015)
1124	Findings on SRS section 3.13 “Speed and distance monitoring”
1127	Non convergence of the release speed calculated on-board
1147	DMI text message handling
1148	Trigger of specific NTC data entry
1149	Alignment of PBD SR requirements with the new braking curve model
1150	Incomplete V_MRSP definition vs train position
1151	Error in Subset-037 Table 11
1153	Train interface passive shunting input simplification
1154	Train interface - clarification of isolation output
1155	CR712 follow-up: packets sent as non-infill information from infill device
1157	SUBSET-076 upgrade to Baseline 3
1158	SUBSET-074 upgrade to Baseline 3
1159	Missing train-to-track message specification for RBC X=1
1168	Unspecified ACC RBC behavior when receiving new pre-announcement messages in ongoing transaction
1173	Miscellaneous problems with STM specifications

1176	Feedback on SRS chapter 6 from Baselines compatibility assessment
1183	Unclear use of telegram header info when a balise telegram or BG message is ignored/rejected
1185	Miscellaneous editorial findings in SRS&DMI spec 3.3.0
1223	Display in Limited Supervision
1231	Miscellaneous editorial findings in SUBSET-027 v3.0.0

A year after, in 2016, the specifications evolved to 3.6.0, also called **Baseline 3 Release 2** (R2). This R2 includes:

- 55 CRs, with 39 error corrections and 16 enhancements
- 25 subsets of specifications affected
- Key changes on: GPRS; Key management; ATO (finally postponed…), and Low adhesion factor.

The changes implemented consider the following CRs:

ID number	Headline
0239	Train data on TIU
0299	Version compatibility check
0539	Set speed indication for driver
0740	Unclear requirements concerning functions active in L2/L3 only
0741	Packet data transmission for ETCS
0852	Definition of level 2/3 area and level transition border
0933	Storing of RBC contact information
1014	Duplicated balises ambiguities
1033	Disable Start in SR if no safe connection
1084	Target speed masking
1086	Unknown L1 LRBG reported to RBC
1087	Manual network selection
1089	Ack for text messages in NL mode
1091	Insufficient driver information in OS
1094	Unclear stop conditions for display of some DMI objects
1107	Status planning information on the DMI in FS mode
1117	Reception of an order to terminate a communication session while session is being established
1122	Communication session establishment to report change to SL mode
1125	Clarification of human role in ETCS safety analysis
1129	DMI indication of level announcement in SB
1152	Avoid increase of permitted speed and target distance
1163	Train interface - Track conditions related outputs to be harmonized
1164	Ambiguity in assignment of coordinate system
1167	Juridical data for the equivalent brake build-up time
1169	Ambiguity about the variable L_STMPACKET in juridical data STM INFORMATION

1172	Problems related to level crossing supervision
1180	Guard rails and cables in the vicinity of balises
1184	Missing requirement for the number of communication sessions an OBU must be capable to handle simultaneously
1187	Indication marker inconsistency
1188	Balises in Multi-Rail Track
1190	UES text message end condition
1197	Ambiguity regarding the temporary EOAs and SvLs
1213	SUBSET-091 upgrade to Baseline 3 Release 2 (B3R2)
1221	Availability of Override and Start buttons
1222	Inconsistency regarding list of BGs for SH area
1229	Age requirement for estimated speed
1236	Criteria for Levels in train unclear
1237	KMS evolution
1242	Several problems with STM specifications
1245	Display of ETCS override in level NTC
1249	Problems with pre-indication
1250	Incorrect description in gradient profile
1254	Session establishment attempts to report mode change
1255	Impossibility to transmit unknown values in the message “Additional data”
1260	Inconsistent set of clauses regarding the service brake interface in SH mode
1262	Issues related to the initiation of a communication session by an RBC
1265	Miscellaneous editorial findings in B3 MR1
1266	Classification of SRS clauses
1273	Impact of UIC 544-1 new version
1275	Eurobalise transmission susceptibility requirements not linked to interoperability
1277	D7 of SoM procedure is reached while no Mobile Terminal is registered yet
1278	SUBSET-074 upgrade to Baseline 3 Release 2 (B3R2)
1280	System version number increment for B3R2
1283	Inconsistent use of the terms EOA and LOA
1284	SUBSET-092 upgrade to Baseline 3 Release 2 (B3R2)

The complete list of documents and versions can be found at

http://www.era.europa.eu/Core-Activities/ERTMS/Pages/Set-of-specifications3.aspx

Going back to Spanish National functions, some of them have been include in the standard, or can be managed through standard functions and interfaces:

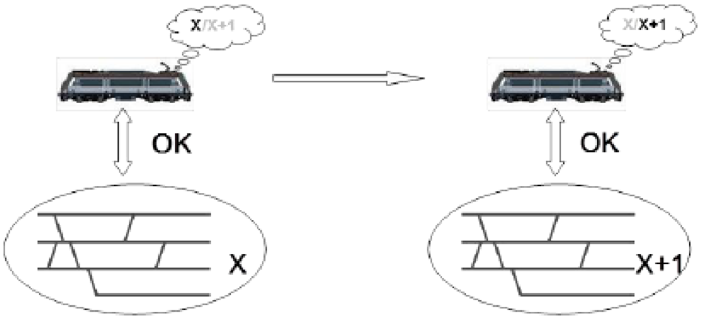
- Función 20: Gestión separada de limitaciones temporales de velocidad según el nivel
- Función 24: Gestión del mensaje por defecto de eurobaliza
- Función 125: Operación en áreas ERTMS fuera de servicio

This three are already included in the Baseline 3 specifications.

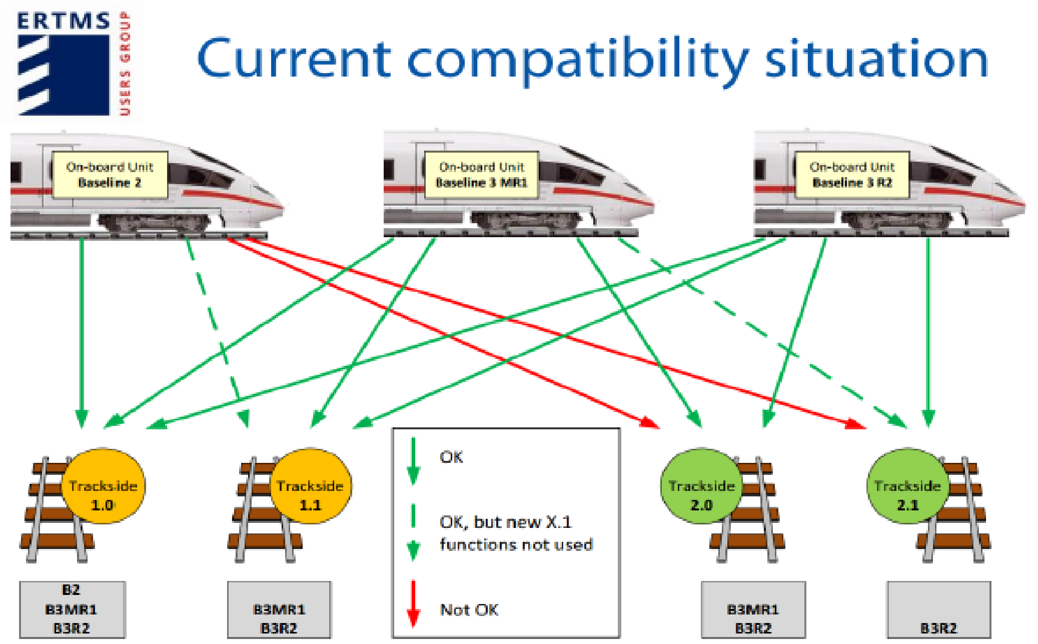
- Función 27: Gestión ERTMS del equipamiento ASFA independiente
- Función 121: Inhibición de niveles disponibles (via external switch)

This two can me managed with Train interface functions integrated in the ERTMS onboard equipment.

And with new Baseline 3 legaly in force, what happen with existing trains /lines in Baseline 2?



The Baseline 3 specifications includes the System version management, to ensure all Baseline 3 trains can operate on lines with at least 2 different versions:


Current compatibility situation

The full picture about compatibility is:

As a summary, Baseline 3:

- offers backward compatibility for baseline 3 trains on baseline 2 infrastructure.
- offers the possibility to upgrade a baseline 2 trackside while keeping most of the existing balises untouched.
- offers the possibility for additional trackside functionalities without losing compatibility with baseline 2 on boards (additional baseline 3 information will be ignored by a baseline 2 train but taken into account by a baseline 3 train)

A detailed document has been released by the European Railway Agency (ERA) together with the Infrastructure Managers and Trains Operators, to identify all the potential interoperability problems between Baseline 2 and 3 subsystems, called Baseline Compatibility Assessment.

The document about Baseline 3 MR1 compatibility can be found in:

http://www.era.europa.eu/Document-Register/Pages/Baseline-Compatibility-Assessment-B3-MR1.aspx

The conclusion, for the evolution from version 2.3.0 D to Baseline 3 MR1, is "vast majority of the 436 analysed Change Requests, the analysis demonstrated that the compatibility objectives for Baseline 3 […]have been achieved and no potential compatibility problems were identified".

Besides, "some other potential compatibility issues were identified, due to shortcomings or 'grey areas' in Baseline 2, for which mitigation measures could be needed to ensure interoperability."

**(extracts from Baseline Compatibility Assessment - Final Report, Reference EUG_UNISIG_BCA, version 1.0.0)*

This document about Baseline 3 R2 compatibility can be found in:

http://www.era.europa.eu/Document-Register/Pages/Baseline-Compatibility-Assessment-for-B3-R2.aspx

The conclusion is for 50 out of the 55 CRs added in MR2, the compatibility between the baselines has been achieved and no potential compatibility problems were identified. For other 5 CRs (CR 933, 1089, 1184, 1249 and 1262) have been identified mitigation measures.

All these analysis have been performed to facilitate migrations to Baseline 3, and coexistence of both Baselines from a legal but also real point of view.

And together with specifications evolution, the products also evolve.

New generation of ERTMS onboard equipments have:

- reduced HW, with less components and less power consumption
- more functions integrated in the European Vital Computer
- more processing capacities
- and open to other standards /equipments used in other sectors.

New EVC included in the Alstom ATLAS solution, includes:

- 19”, 6U rack with all functions integrated (Radio and Balise Transmission Modules integrated, Juridical Recording function, Diagnostic, …)
- Additional communication modules to integrate EVC in central diagnostic systems.
- Capability to use other communication than GSM-R and Profibus.
- ATO and Eco driving
- Etc…

Safety and signalling systems

360.revista de alta velocidad
número 5 - junio - 2018. Pág 117 - 123

High-speed railway and the digital future

González, Ricardo

Siemens Rail Automation[1]

Abstract

This paper deals addresses the current state of High-Speed Railway with respect to signalling and train control technology and how digitization will play a significant role. The analysis includes proposal in the medium term for improvements from current situation.

Different aspects of the current High-Speed Railway are described, including their success factors, the requirements of the railway operators and the technology currently used. It also describes the current portfolio of products offered by Siemens for the signalling of High-Speed Railways.

Finally this paper describes a potential future for High-Speed Railway and how digitalization will play a key factor in its transformation. As an example of the systems that benefit from digitalization, ETCS Level 3 system is briefly described.

Keywords: high-speed railway, digitalization, ETCS, future, smart data

---

[1] *González, Ricardo. R&D Portfolio Manager Main Line and Interlocking. Siemens Rail Automation. Email: ricardo.gonzalez_garcia@siemens.com*

1. Introduction

Digitalization is one of the key factors to ensure an increasing use of railway, especially High-Speed Railway and to achieve three important mobility objectives: efficiency, reliability and sustainability.

The High-Speed Railway has evolved from the initial projects. The evolution of the technology during the last years has produced several techniques and applications that can be used by High-Speed Railway. Although, due to the High-Speed Railway safety requirements, every technological step in railway technology has been taken with great care. It is time to evaluate the current state of the railway and define the functionality required for the medium term.

The paper describes these applications and the development of railway signalling technologies, linked with the digital era. One of them is ETCS Level 3, where Siemens have been working in the last years, and where wireless communication will make part of track infrastructure unnecessary. This will improve the availability and maintenance of infrastructures, facilitating greater occupation of the railways and total safety.

2. State of the High-Speed Railway today

This chapter describe the success factors of the High-Speed Railway, the current requirementsof the different stakeholders and the Siemens signalling products related with High-Speed Railway.

2.1 Success factors

Today High-Speed Railway is worldwide perceived as a transportation system which is:

- Safe
- Comfortable
- Reliable
- Green
- Fast
- Efficient

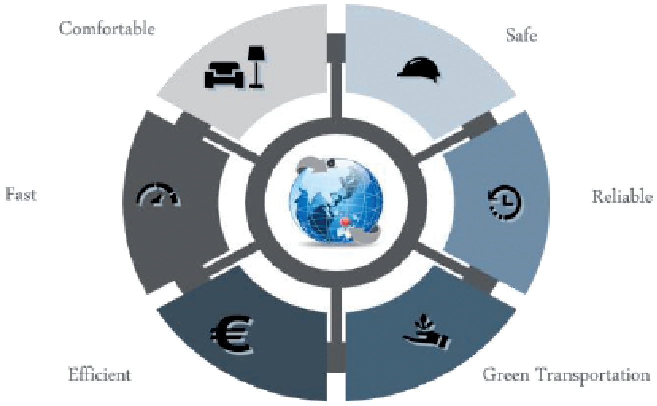


Figure 1: Success factors

Safety is one of the main aspects of a High-Speed Railway. The number of fatalities experienced on High-Speed Railways is much lower than in other transport systems. Signalling systems used for the control of High-Speed Railways meet the highest safety standards, ensuring that the likelihood of a wrong side failure is very low.

Comfortability is provided by modern rolling stock and its associated services on board. Another important element that contributes to comfortability is the reduced time that the passenger needs to be at the station before the departure of the train. This is really an advantage when compared with other transport systems as aviation. Another positive factor is that main train stations are located in the centre of the cities, which facilitates access by public transport, as well the reduced time necessary to reach the main stations.

The reliability of the system ensures that train departure and arrival times are consistently met. Punctuality rates are the best among current transportation systems.

High-Speed Trains are usually powered with electricity. This reduces the CO^2 emissions during operation and in its complete lifecycle. That is why it can be considered as a **"Green transport"** system (although electric power can be generated by several means and some of them emit CO^2 into the atmosphere).

Finally High-Speed Railway is the **fastest** ground transportation system. Due to this, and the characteristics described above, it is perceived as an **efficient** transportation system.

2.2 Requirements from operators

From a High-Speed Railway operator s point of view the main requirements for this type of railway are similar to other railways, transportation systems or industrial businesses. Requirements as Availability, Ease of Maintenance, Efficiency and Low Lifecycle Costs (LCC) can be applicable to different areas. Even so, achieving those requirements requires a clear definition of the targets and their associated costs.

Apart from those requirements, there are other requirements specific to the Railway business. They are: Standardization (interoperability) and Improved transport capacity. Although standardization is also claimed by most industrial business, one of the key aspects linked with standardization is Interoperability. This aspect is closely related with ground transportation.

Interoperability (as the one provided by the ERTMS/ETCS signalling system) is a key element to provide operators the possibility to operate High-Speed trains across borders with the same signalling system and, eventually, with similar if not the same rules. Interoperability saves efforts and time in terms of definition of operational rules, equipment investment, time, Interoperability also contributes to improve the overall safety of the railway, as all equipment behave exactly in the same way in despite of the country they operate, reducing operational procedures that are sometimes more prone to human errors.

Another important requirement for the High-Speed Railway is the increase of the transport capacity and, associated to it, the improvement in headway. We can have very fast trains but, if we cannot adapt the operation to customer demands, this will not be enough. Improving transport capacity allows the High-Speed Railway operator to have more trains per hour, increasing the number of passengers per hour. This is one aspect that is becoming more relevant as the population increases, especially in the main cities where the number of passengers increases. Systems that can best contribute to increasing transport capacity are modern signalling systems such as ETCS level 2.

2.3 Signalling Solutions

Modern High-Speed lines use ERTMS/ETCS as the preferred signalling system. And in most projects ETCS Level 2 is the solution chosen. ETCS Level 2 system reduces the amount of equipment to be installed on the track and provides a better headway capacity than other systems used in previous decades. Siemens TrainGuard system provides a complete ETCS Level 2 solution for ETCS Level 2 applications.

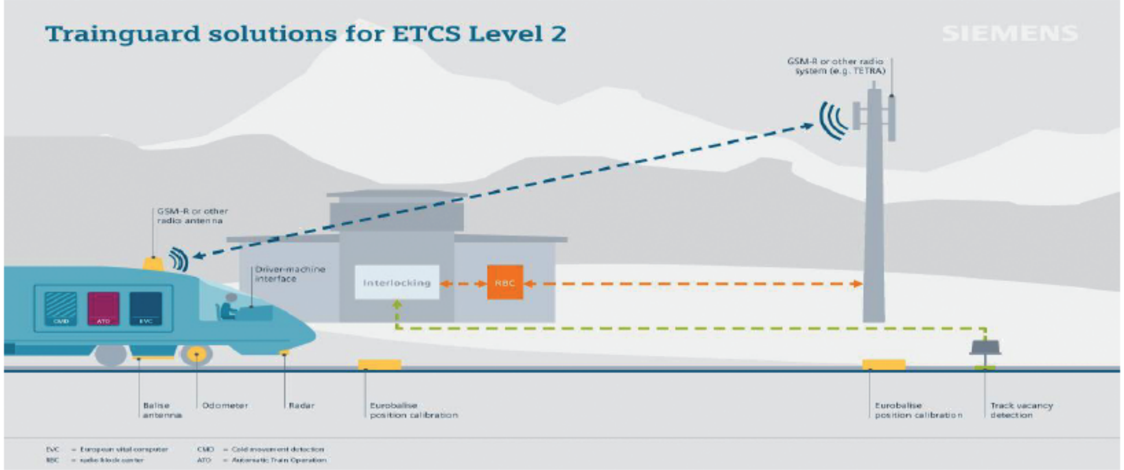


Figure 2: ETCS Level 2

3. High-Speed Railway needs

In the future High-Speed Railway will probably have the following needs:

Better Availability: Even if current High-Speed Railway shows high availability figures, once the infrastructure is used at its limit (i.e. more passenger and freight traffic), availability will become even more important. Use of redundant and distributed systems is the basis for ensuring and increasing availability. The use of cloud computing and virtualization allows having redundant systems and immediate fault recovery in case of disasters.

Reduced Maintenance: Maintenance costs are an important factor in Life Cycle Costs. The use of Smart Data based systems, which can provide predictive maintenance, will reduce maintenance costs and also increase availability. Reducing the amount of equipment also leads to a better maintainability: less probability of faulty elements and less requirements for spares and maintenance personnel.

Increased Efficiency: The usage of Automatic Train Operation and/or Driving Advisory Systems for optimal driving profiles helps to reduce the carbon footprint. The use of cloud computing maximizes the efficiency of investment.

New technology: Digitalization is becoming a must in many businesses. And this also applies to High-Speed Railway. Use of wireless technology, improved Security, Smart Data strategies and implementation of new signalling systems such as ETCS level 3 for High-Speed Railway, will open the door to a digital future.

4. The Future of High-Speed Railway

4.1 Retrospective

Spain has had the High-Speed Railway since 1992. If we look at the progress of technology over the last 25 years, we can provide examples of technology that did not exist at that time, but are essential today:

- Internet (in 1992 it was very primitive)
- MP3 players such as IPod
- GSM (in 1992 it was just started)

- GPS
- DVD
- Social Networks
- WiFi
- Use of Solar Power for end consumers

Some of them (e.g. Ipod, DVD), are now even becoming old technology. They were very successful and disruptive inventions but technological progress and digitalization has been so rapid that in less than 20 years they became obsolete.

However the progress of High-Speed Railway has not been so fast. It is true that there have been several major steps, but High-Speed Railway has several limitations to a rapid progress of technology. One of this limitations is Safety. The High-Speed Railway has to be Safe since any type of incident / malfunctioning can lead to an accident with hundreds of deaths / injuries. This is why the technology adopted by the High-Speed Railway has to be very mature and Proven-By-Use.

4.2 Where we go?

However passengers, operators and maintainers are demanding from High-Speed Railway that provides better performances and services. Requirements such as faster trains, higher train frequencies, improved connectivity, use of open systems... demand from High-Speed Railway technology to move faster. Of course we cannot expect High-Speed Railway to progress at the same pace as IT technology but we can expect broader use of current technology based on the digitalization process. This will include:

- Advanced Centralized Control: including SCADA and concentrate all High-Speed Railway techniques (signalling operation, maintenance, rolling stock,) in one place with the same management systems.
- Smart data and Cloud Computing: This will provide, among other benefits, comprehensive maintenance management. It will also reduce the obsolescence hardware costs and CAPEX and OPEX costs related to computing hardware. The use of open standards will also be beneficial to this step. This technology is also linked to Smart Access technology that will reduce access time to High-Speed Trains and increase Security.
- Virtual reality: that will be applied for maintenance and training purposes. Other existing technologies will be used on High-Speed Railway such as 3D printing to supply spares just in time. The use of Drones and Robots would be beneficial for maintenance and security activities.
- Intelligent Systems: that will be applied to High-Speed Railway to increase frequency of trains through dynamic distancing, conflict detection and resolution and the use of intelligent surveillance to automatically solve any type of situation during the operation of the High-Speed Railway.

4.3 Siemens proposal and the basis for the Digital Future

In all the systems described above there is a common topic: Digitalization. The use of digital technology will radically change the shape of our High-Speed Railways.

Siemens is focused on Digitalization trend in all its activities and businesses: from industry to transportation.

Siemens proposal for the digital future of High-Speed Railways is focused on:

Safety and Security: Future High-Speed Railway systems will become more complex in terms of functionality and interfaces. This will require continuing to ensure the safety of the system as some elements such as signalling require being safety critical. But on top of that, Security is increasingly important, especially in all technological areas (IT Security). Special measures will be required to ensure that High-Speed Railway Security is not violated, as a Security failure may cause a failure in Safety. IT Security implementation will be required for most IT related systems on a future High-Speed Railway.

A sustainable and green system: Reducing carbon footprint can be achieved through the use of advanced Automatic Train Operation (ATO) and Driving Advisory Systems, which are capable of providing optimized driving strategies depending on the demand. This includes from Ecodriving when the requirements in transport capacity are medium-low to the maximization of the transport capacity (with greater energy consumption) when the transport capacity requirements are higher. In addition to this, conflict detection and resolution strategies will optimize the use of tracks and trains avoiding unproductive waiting times.

Improved Traveller experience: By using Smart Data to provide immediate information to all High-Speed Railway stakeholders (operators, maintenance personnel,..) and especially to the passenger in such a way that travelling in a High-Speed Train is an easy and pleasant experience.

Intelligent investment: This will be achieved through the use of current technologies and tools. As an example the use of Cloud Storage and Cloud Computing, so that investment costs related to hardware assets can be reduced. This would also offer the possibility of having distributed computing that offers better availability and the possibility of centralizing, in a single application, all management of High-Speed Railway. Extensive use of wireless technologies will also help to reduce installation and maintenance costs. As an additional intelligent investment initiative, it can be included the ETCS Level 3 application, as described below.

4.4 ETCS Level 3

ETCS Level 3 is the future of signalling systems. It combines most of the advantages of existing signalling systems with few drawbacks.

The main advantages of this system are:

- Based on existing ETCS Level 2 system
- Interoperability (it is an ETCS Level and will become interoperable)
- Very low requirements for track detection systems (low CAPEX and OPEX)
- Based on radio communications
- The usage of Moving Block/Virtual Block technology notably improves the performance of Level 2 as it can potentially provide the best possible headway improving transportation capacity

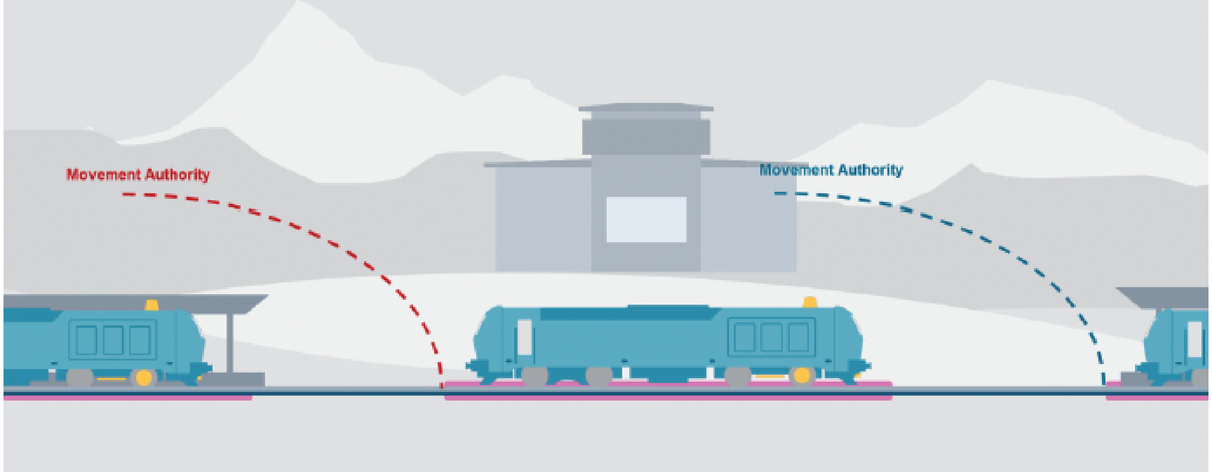


Figure 3: Moving Block Principle

Due to the requirements of High-Speed Railways, now and in the future, ETCS Level 3 may be the optimal signalling solution since it reduces the CAPEX and OPEX linked to these projects, providing flexibility in terms of transport capacity. Interoperability also ensures competition and long Life Cycles.

5. Conclusions

Implementation of new technologies in High-Speed Railways takes more time than in other sectors. One of the main reasons is the need to have Proven-In-Use systems and include requirements for Safety Critical systems. However digitalization is a trend that is also reaching the railway and is expected to accelerate railway digitalization during the next few years.

High-Speed Railways will use Digitalization to improve CAPEX and OPEX, to have better travel experience, to increase the availability figures and reduce carbon footprint. The use of Smart Data, Cloud Computing and new interoperable systems such as ETCS level 3 will be the artefacts for this digital future.

Safety and signalling systems

360.revista de alta velocidad
número 5 - junio - 2018. Pág 125 - 147

Probabilistic Safety Analysis of High Speed and Conventional Railway Lines

Grande, Zacarías
Blanco López, Marta
García Tamames, Alberto
Castillo, Enrique

University of Cantabria[1]
Spanish Royal Academy of Engineering

Abstract:

A probabilistic safety analysis methodology based on Bayesian network model is presented, in which all the elements encountered when travelling along a railway line, such as terrain, infrastructure, tunnels, viaducts, light signals, speed limit signs, curves, switches, rolling stock, and any other element related to its safety are reproduced. Especial attention is given to human error and modelling the driver behaviour variables and their time evolution. The conditional probabilities of variables given their parents are given by means of closed formulas, which facilitate the software implementation of the proposed models. The model provides a probabilistic safety assessment of the line such that its most critical elements can be identified and sorted by importance. This permits improving the line safety and optimizing the maintenance program by concentrating on the most critical elements. To reduce the complexity of the problem, a method is given that divides the Bayesian network into small parts such that the complexity of the problem becomes linear in the number of items. In addition, when an accident occurs, a backward inference process allows us to identify the causes of incidents. The case of the real Palencia–Santander and Vitoria-Zaragoza lines together with some other simple examples are used to illustrate the advantages of the proposed methodology.

Keywords: Bayesian networks, Backward analysis, Partition technique.

---

[1] *Grande, Zacarías. University of Cantabria. Email: zgandrade@gmail.com*
Blanco López, Marta. University of Cantabria. Email: marta.blancolo@alumnos.unican.es
García Tamames, Alberto. University of Cantabria. Email: alberto.garciat@alumnos.unican.es
Castillo, Enrique. Founder Member of the Spanish Royal Academy of Engineering. Email: castie@unican.es

1. Introduction and motivation

Probabilistic safety analyses are mandatory and used regularly in the safety assessment of nuclear power plants because of the serious implications of nuclear accidents. However, this type of analysis is not mandatory in the case of railway lines. In this paper the convenience of performing probabilistic safety assessments of railway lines is analyzed and seriously recommended following the trend of introducing new computational methods in Railway lines initiated by (Amit & Goldfarb, 1971) (Assad, 1980), (Burdett & Kozan, 2010) (Cacchiani & Toth, 2012), (Caprara, et al., 2002) (Carey, 1994) (Carey & Crawford, 2007) (Carey & Lockwood, 1995), (Castillo, et al., 2015) (Castillo, et al., 2011) (Castillo, et al., 2009) (Castillo, et al., 2016) (Cordeau, et al., 1998) (D'Ariano & Pranzo, 2004) (D'Ariano, et al., 2007) (Haghani, 1987) (Haghani, 1987) (Lin & Ku, 2013) (Ouyang, et al., 2009) (Pachl, 2014) (Petersen, et al., 1986) (Sahin, 1999) (Yang & Hayashi, 2002), etc.

The actual protocols to evaluate risk in railway lines start with an evaluation of each possible hazardous event to decide whether or not the associated combination of frequency of occurrence and consequences is important enough to deserve a detailed analysis of such event and decide the required actions if needed. In this context, Table 1 of Risk Assessment, which is taken from the European Standard 50126 and the Spanish UNE-EN-50126, is used.

Table 1 Risk levels considered in the European safety analysis of railway lines

Frequency of occurrence of hazardous event	**Risk level**			
Frequent	Undesirable	Intolerable	Intolerable	Intolerable
Probable	Tolerable	Undesirable	Intolerable	Intolerable
Occasional	Tolerable	Undesirable	Undesirable	Intolerable
Remote	Insignificant	Tolerable	Undesirable	Undesirable
Improbable	Insignificant	Insignificant	Tolerable	Tolerable
Incredible	Insignificant	Insignificant	Insignificant	Insignificant
	Insignificant	**Minimum**	**Critical**	**Cathastrophic**
	Severity levels of hazard consequences			

Table 1 shows the different risk levels considered in the European and Spanish standards. Though the word risk is normally associated with the probability of occurrence of an event, here the concept of risk is actually the expected damage (product of the probability of occurrence of the event by the damage produced). Since both the frequencies and the severity levels are given by a rather imprecise terminology, in Table 2 the frequency levels according to the ADIF methodology are explained. Although each frequency is defined in two different ways, it is still very imprecise and can lead to different interpretations by different experts, which means that two different experts may assign a different level of frequency and risk level to the same undesirable event. As we shall see, this is rather serious because of its negative consequences on safety.

Table 2 Sets of frequencies used in the standards

Frequency	Description
Frequent	It may happen frequently. The threat will be played continuously.
Probable	It will happen several times. The threat can often be expected to occur.
Occasional	It may happen several times. The threat can be expected to occur several times.
Remote	It may occur during the system life cycle. It can be reasonably assumed that the threat will occur.
Improbable	Slight chance but possible. It can be assumed that the threat can occur exceptionally.
Incredible	Extremely unlikely. It can be assumed that the threat will not occur.

For example, when referring to events that occur frequently or several times, it is not indicated if this refers to one day, one month, one year or the life of the system being analyzed. Similarly, the meaning of slight chance or extremely unlikely are not quantified, which is the only way of avoiding a misinterpretation of the codes.

Once frequencies and severity levels of their consequences of given events are determined, Table 1 allows to determine the risk levels, which can be classified as: Intolerable, Undesirable, Tolerable, and Insignificant.

Table 3 shows the required action to be taken under any risk level, which has a relevant role in the safety assessment of hazardous events. Thus, careful attention must be paid to the different actions to be carried out for each risk level: “Intolerable”, “Undesirable”, “Tolerable” and “Insignificant”, resulting from the combinations of the different frequencies (“Frequent”, “Probable”, “Occasional”, “Remote”, “Unlikely”, or “Incredible”) and severity levels (“Insignificant”, “Minimum”, “Critical” or “Catastrophic”).

Table 3 Required actions associated with the different risk levels

Risk level	Required action
Intolerable	It must be removed
Undesirable	It will be accepted only when the risk reduction is impracticable and in agreement with the Railway Authority (ADIF)
Tolerable	Acceptable with proper control and in agreement with the Railway Authority (ADIF)
Insignificant	Acceptable without any agreement

It is surprising that in Table 1 the levels of “Critical” and “Catastrophic” risk associated with an “Incredible” frequency are associated with an “insignificant” level of risk, that is, a risk acceptable without any agreement. Even “Critical” and “Catastrophic” risk levels associated with “Improbable” occurrence frequencies are associated with a “Tolerable” risk level, that is, acceptable with adequate control. Therefore, events that are “Incredible”, that is, extremely unlikely or that it can be assumed that the threat will not occur, are exempted from further verification regardless of their consequences. Events that are “improbable,” that is, unlikely but possible or that the threat can be assumed to occur exceptionally, are tolerable under control even if they have very serious consequences.

This highlights the danger of using not only vague and inaccurate, that is, unquantified or grossly quantified frequencies, but also imprecise boundaries between events requiring a safety analysis and those free of it. In other words, it is clear that this methodology, despite of being a recommendation of the European authorities, is not only the most appropriate but should be corrected as soon as possible.

Fortunately, there are recommendations that already correct these defects, such as, for example, the important work of (Beales, 2002). Given that these recommendations imply a major change of the initial content of the rules, in fact they indirectly recognize their serious deficiencies. Thus, an urgent change of the current code is needed.

For the sake of illustration, Table 4 is the table for risk assessment recommended in the document published by the RSSB (Railway Safety Standards Board), which leads to safety analysis of many events not included in Table 3.

Table 4 Table for risk assessment recommended in the document published by the RSSB (Railway Safety Standards Board).

		Consequence				
		2	**3**		**4**	**5**
		Fatalities/event				
Frequency		**1 fatality in every 125 events**	**1 fatality in every 25 events**	**1 fatality in every 5 events**	**1 fatality in every event**	**5 fatalities in every event**
	No/year	0.008	0.04	0.2	1	5
7	31.25	0.25 8	1.25 9	6.25 10	31.25 11	156.25 12
6	6.25	0.05 7	0.25 8	1.35 9	6.25 10	31.25 11
5	1.25	0.01 6	0.05 7	0.25 8	1.35 9	6.25 10
4	0.25	0.002 5	0.01 6	0.05 7	0.25 8	1.35 9
3	0.05	0.0004 4	0.002 5	0.01 6	0.05 7	0.25 8
2	0.01	0.00008 3	0.0004 4	0.002 5	0.01 6	0.05 7
1	0.002	0.000016 2	0.00008 3	0.0004 4	0.002 5	0.01 6
Units are fatalities per year						

Note that in Table 4 the frequencies and severity levels of consequences are classified as levels 1 to 7 and 1 to 4, respectively.

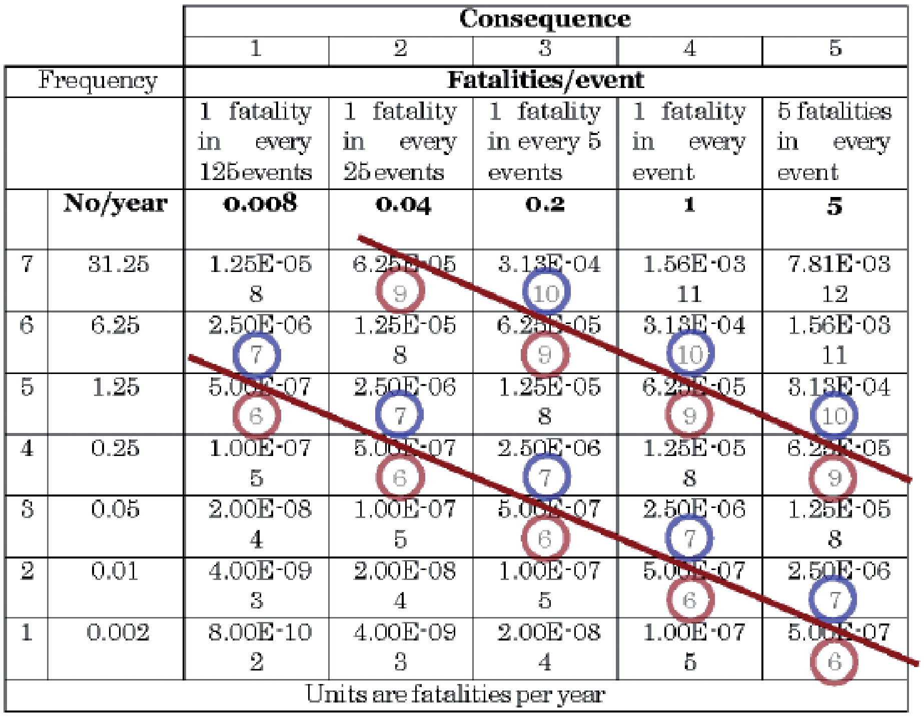
Note that both, frequencies and consequences are quantified in levels related by a factor of 5. In addition, to assign numerical risks of each frequency-consequence combination, instead of multiplying the two row and column entries, they add the levels (exponents of power 5), on the basis that the power product of the same basis is another power with the same basis and the sum of the exponents. In other words, the indices contained within each of the central cells, which take integer values between 2 and 11, are the sum of the levels of their row and column.

Table 4 with the entries multiplied by a factor, which depends on the particular event being analyzed, and two threshold values allows classifying the individual risk, measured in probability of fatality per year, into three regions (see Table 5): the region where risk must be removed, the region where the risk must be mandatory analyzed in detail and the region where no further action is required. One example is Table 5, where the event refers to an accident of a commuter assuming 500 journeys per year and 10E+06 passenger journeys per year. This means that the cell in Table 4, with Frequency 7 and consequence 1, that is, a risk level of 0.25 fatalities per year corresponds:

$$\frac{0.25}{10^6} \cdot 500 = 1.25 \cdot 10^5 \text{ fatalities/passenger/year,}$$

Which is the value in the same cell of Table 5

Table 5 Table for risk assessment recommended in the document published by the RSSB (Railway Safety Standards Board).



		Consequence				
		1	2	3	4	5
Frequency		Fatalities/event				
		1 fatality in every 125events	1 fatality in every 25events	1 fatality in every 5 events	1 fatality in every event	5 fatalities in every event
	No/year	**0.008**	**0.04**	**0.2**	**1**	**5**
7	31.25	1.25E-05 8	6.25E-05 9	3.13E-04 10	1.56E-03 11	7.81E-03 12
6	6.25	2.50E-06 7	1.25E-05 8	6.25E-05 9	3.13E-04 10	1.56E-03 11
5	1.25	5.00E-07 6	2.50E-06 7	1.25E-05 8	6.25E-05 9	3.13E-04 10
4	0.25	1.00E-07 5	5.00E-07 6	2.50E-06 7	1.25E-05 8	6.25E-05 9
3	0.05	2.00E-08 4	1.00E-07 5	5.00E-07 6	2.50E-06 7	1.25E-05 8
2	0.01	4.00E-09 3	2.00E-08 4	1.00E-07 5	5.00E-07 6	2.50E-06 7
1	0.002	8.00E-10 2	4.00E-09 3	2.00E-08 4	1.00E-07 5	5.00E-07 6
Units are fatalities per year						

Note also that the limit between "Tolerable" and "Intolerable", indicated by the line in red, is between 9 and 10 and that the limit between "Acceptable" and "Tolerable", indicated by the bottom line in red, is between 6 and 7.

Consequently, the above Table 4 significantly improves the European standard 50126 and the Spanish standard UNE-EN-50126, since:

1. It uses seven frequency levels instead of six.
2. Quantifies the frequencies accurately.
3. It uses frequency levels (proportional to 5^n where n is the level) that multiply by five the frequencies of the previous ones, which allows levels to be associated with frequencies, that is, qualitative with quantitative information, which does not occur neither in European standard 50126 nor in the Spanish UNE-EN-50126.
4. Quantifies the consequences in terms of deaths and using also a factor 5 to pass level, which avoids ambiguities.
5. Declassifies as "acceptable without any agreement" ("Negligible risk") low frequency cases with serious consequences

Nevertheless, the use of tables, such as Table 5, can also be criticized because they provide a risk per year and this depends on the number of trips per year and its length (the larger the number of trips and its length, the larger the risk). In other words, using this yearly risk discriminates the different events by length and number of trips. In our opinion the risk must be given in fatalities per kilometer. The fact that a user travels more time per year or uses longer or shorter trips should not change the required safety level.

The most important conclusions that can be deduced from the above are:

1. The European standard 50126 and the Spanish UNE-EN-50126 are not the most appropriate, and if used could easily lead to the conclusion that the risks associated with the event, "very unlikely with serious consequences" are "acceptable without any agreement", that is, do not need a safety analysis, if such event had been considered "Incredible."
2. This will not occur in the case of using the recommendations indicated by the RSSB, or more modern methodologies, such as those based on fault trees or in Bayesian networks.
3. The imposition of the ADIF methodology, based on European and Spanish regulations, as the only accepted ones, giving as a reason the difficulty of comparing results if different methods were used, should be eliminated, or at least thoroughly reviewed, by the risks which can be overlooked in its application, which is also mandatory.
4. Finally, the risk should be given by km and not per year.

2. Probabilistic Safety Analysis (PSA)

Once the need of a detailed probabilistic safety analysis (PSA) has been detected we have to identify all possible risks and proceed to evaluate the whole risk of the line. To this end, we can use the above method, which is very cumbersome or use an alternative.

In this paper we present a method to simplify the process based on Bayesian networks, which can be used as the main tool for probabilistic representation of multidimensional variables in order to analyze railway safety.

The methodology of the model can be summarized in different stages described below:

The first step is to identify and reproduce the most relevant variables that play an important role in the safety of a railway line to model its multidimensional random behavior.

The second stage consists of reproducing the elements that the driver and the train observe when they travel along the line. Particularly, we reproduce the evolution of the level of attention of the driver, the speed and occurrence of incidents.

2.1 Proposed Bayesian network model

2.1.1 Introduction

As event trees and fault trees, although very powerful present some important limitations, Bayesian networks have been selected, given the previous experience of the authors in this very important tool of representation of the probabilistic structure of multidimensional random variables.

A Bayesian network consists of two elements: A directed acyclic graph and a set of conditional probability tables. In this way, any joint multidimensional probability can be reproduced with no restriction, which implies being able to treat any multidimensional random variable. To properly perform a Probabilistic Safety Analysis (PSA) it is convenient to use the following stages: The first requires to identify and reproduce all the items or elements which are relevant to the safety of a railway line. The second stage must identify the variables that influence the elements previously described and how they influence them. In particular, the evolution of the driver's level of attention (variables M), speed (V variables), the presence of signals of all kinds, tunnels, viaducts, occurrence of accidents (variables A), etc. are required. Finally, the third step must be devoted to define the structure of the model and quantify the conditional probabilities.

With this aim, a video from the train's cabin recording the railway line becomes an essential tool to identify all these possible hazardous items that the engineer encounters when travelling along the railway line. Some elements that can be considered are crossing switches, tunnel crossings, cutting and embankments, viaducts, infrastructure crossings (overpass and underpass), level crossings, landslides, etc. not forgetting the light signals and signs of speed limitations, since a large distance is required to be able to reduce the speed and errors are not always protected by the automatic protection systems (ATP).

2.1.2 Variables used in the model

From the previous discussion, we can identify the following list of variables that are very important to the line safety:

1. *Tr: Driver's tiredness*. Since the driver is subject to an increase of tiredness with driving time a variable is needed to analyze how it changes along the line when travelling. Since tiredness is known to be one important contribution to human error, it must not be forgotten.
2. *D: Driver's attention.* It refers to the driver's attention level that in our model is simplified to three states: distracted, attentive and alert. We assume that an alert situation always leads to a correct decision and that a distracted situation leads to a no action at all. Contrary, an attentive situation is subject to both correct and incorrect actions with given probabilities, such that the probability of the first is much bigger that the probability of the second.

By *distracted* we understand a situation in which the driver lacks the necessary attention to correctly react when an action is required. By *attentive* we refer to the case in which the driver is able to react adequately to the required actions with a small probability of error. Finally, *alert* refers to the case where the driver is ready to take an action and knows that he/she has to act immediately (for example, after seeing a warning signal or consulting the railway driver’s guide, etc.).

The evolution of driver’s attention when driving progresses and different elements along the line encountered must be considered.

3. *S: Speed* It refers to the train speed at the corresponding location and can take a discrete list of values. In the examples considered in this the- sis we have simplified the model using the following list of values.
4. *A: Accident*. It refers to the accident occurrence at the actual location or before and can take the following values: none, minor, medium and severe.

   The model assumes that once an accident has occurred at a given location, no accidents can take place in any other forward location.
5. *RS: Rolling Stock*. It refers to the rolling stock conditions and includes the damage levels: none, minor, medium and severe.
6. *Inf: infrastructure.* With this variable the infrastructure state (rails, sleepers, ballast, plate, maintenance standards, etc.) is considered, and it includes the following damage levels: none, minor, medium and severe.
7. *T: Terrain.* This variable is used to consider the risk associated with falling stones on the infrastructure or slope sliding in cuttings and embankments and takes values: stable, small instability, medium instability and high instability.
8. *DE: Driver’s decision on speed control*. It refers to the decision made when speed is controlled by the driver, and includes the following levels: correct, error I (speed remains unchanged), error II (selected speed does not coincides with required speed).
9. *DA: Driver’s decision at signal.* It refers to the decision made when the train encounters a signal, and includes the following levels: correct, error (incorrect action of the driver).
10. *ATP: Automatic Train Protection System.* This variable refers to the supervising or driving assistance system operating at the considered point of the line. It takes the values:”ERTMS”, “ERTMS-ASFA”, “ASFA-dig”, “ASFA-AV’’, “ASFA-Conv”, “ASFA- anal”, “SR” (staff responsible).
11. *AS: Light signal Decision.* It refers to the possible errors at a light signal: none, error I (stop announcement signal), error II(signal at red).
12. *DS: Driver’s decision made at a speed limit signal.* It refers to the possible errors made by the driver at a speed limit signal: none or error I (fail to reduce speed).
13. *SS: Light signal state.* It refers to the light signal: free, stop announcement, stop.
14. *TF: Technical failure.* It refers to the possibility of a technical failure: yes or no. For example, if the driver tries to stop the train and the brakes fail.

2.1.3 Markovian Model

To analyze the driver’s attention, we propose a Markovian model that considers only three driver’s attention states (distracted, attentive and alert) (see Figure 1). “Distracted” means that the driver does nothing, “alert” means that the driver makes correct decisions, that is, without error, and, finally, “attentive” implies normally a correct decision but some errors with a very small probability.

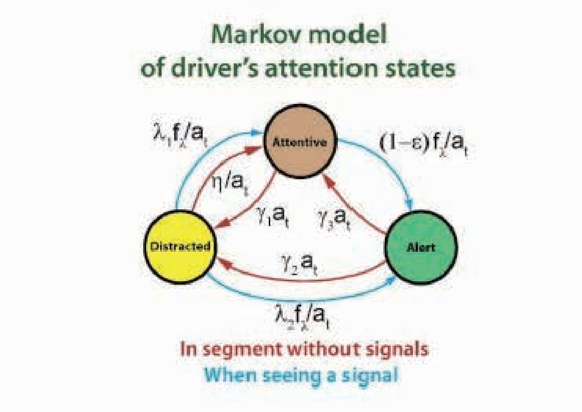


Figure 1. Illustration of the Markovian model used to model the driver’s attention.

Therefore, the changes of driver’s attention state due to the different line incidences; as well as how erroneous decisions are corrected by the supervisor systems must be modeled.

Since travel times of the different trains circulating along the network or line could be very different and it is not the same a delay of five minutes in a one hour trip than in a three hours trip, we use in our model relative travel times. The relative travel time is defined as the quotient between the actual travel time and the minimum possible travel time, that is, at maximum speed. This means that a relative travel time 1 means that the train travels at maximum speed, and a relative travel time of 1.10 menas that the travel time is 10 % above the minimum travel time. In this way we can combine trains with small and large travel times and also freight trains. In addition a train priority is considered as a factor to be applied to the relative travel time of each train.

2.1.4 Bayesian Network

A Bayesian network consists of two elements:

1. A directed acyclic graph, which includes one node per variable and links which determine from which variables (nodes) each variable directly depends on.
2. Tables of conditional probabilities of each variable given its parents, which quantifies the dependence relations among the variables.

This allow us to reproduce any joint multidimensional probability distribution without any restriction, so that we can model any set of multidimensional variables.

To simplify the Bayesian network building process, it can be divided in parts, each part corresponds to one item or element with its variables and links or to the segment between consecutive items and its variables and links. Figure 2 illustrates how these different parts are assembled in order to obtain the whole Bayesian network from the parts. It contains the real line as it is perceived by users and the mathematical model where the left graphs correspond to segments between items and the right graphs, to the items themselves.

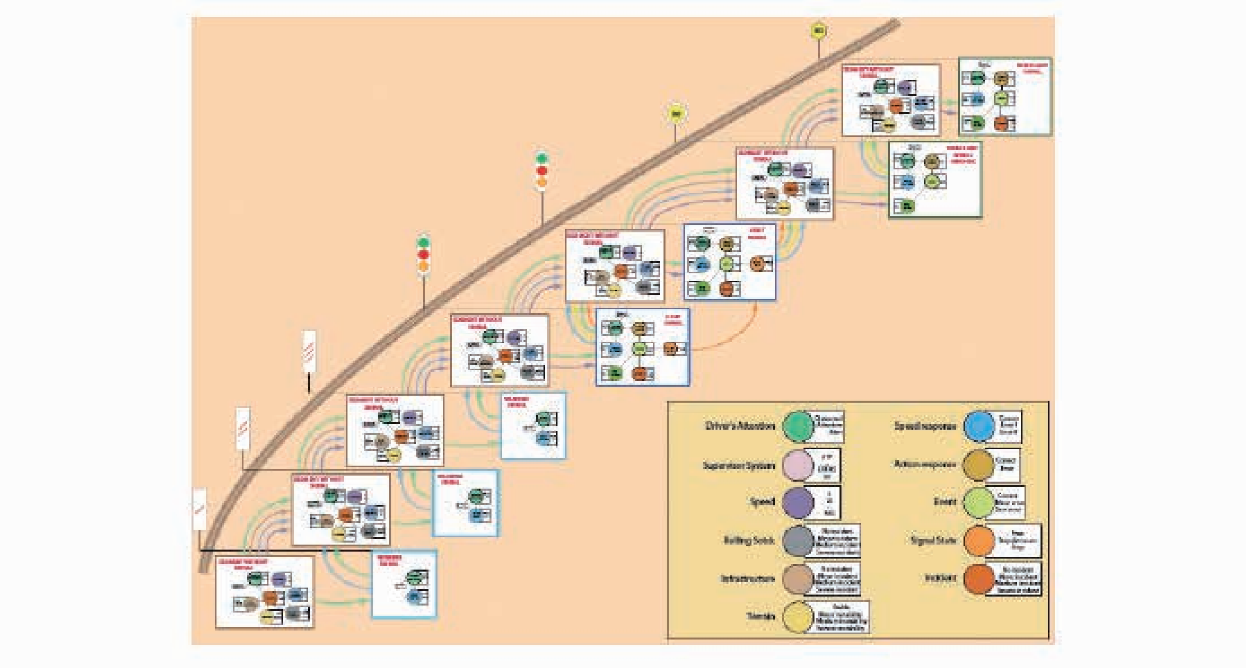


Figure 2. Assembling process of the different parts of the Bayesian network showing the segment between items and the item parts in the left and right locations, respectively.

Figure 3 shows the Bayesian network graph after the assembled process and the corresponding real line.

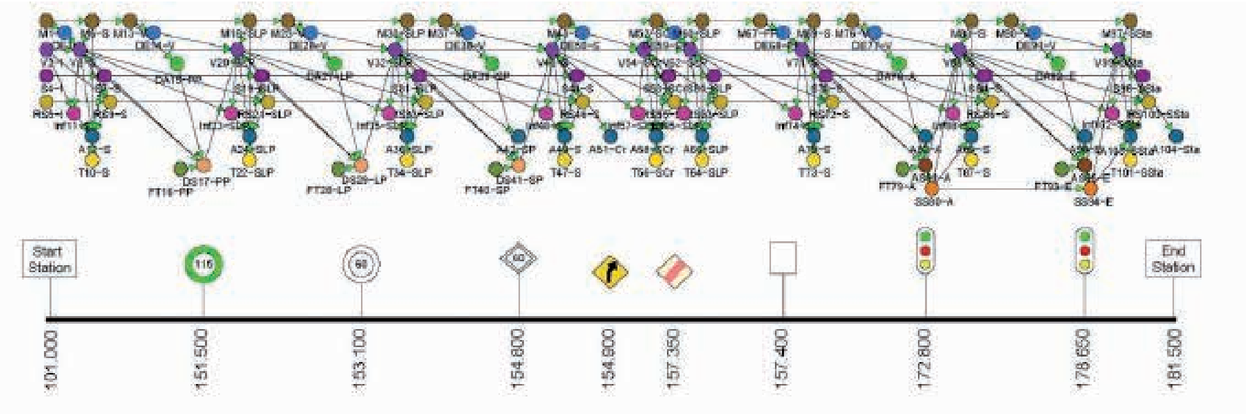


Figure 3. The Bayesian network graph after the assembled process and the corresponding real line.

2.1.5 Sub-Bayesian networks

2.1.5.1 Light signal

In this example a light signal composed of three warning signals, an advanced signal and an entry signal is modeled showing the associated influential variables (see Figure 4):

- Driver's attention state
- Speed
- Accidents
- Rolling stock
- Terrain and infrastructure failures
- Driver's decision
- Technical failure
- Supervisor
- Signal state
- Driver's tiredness

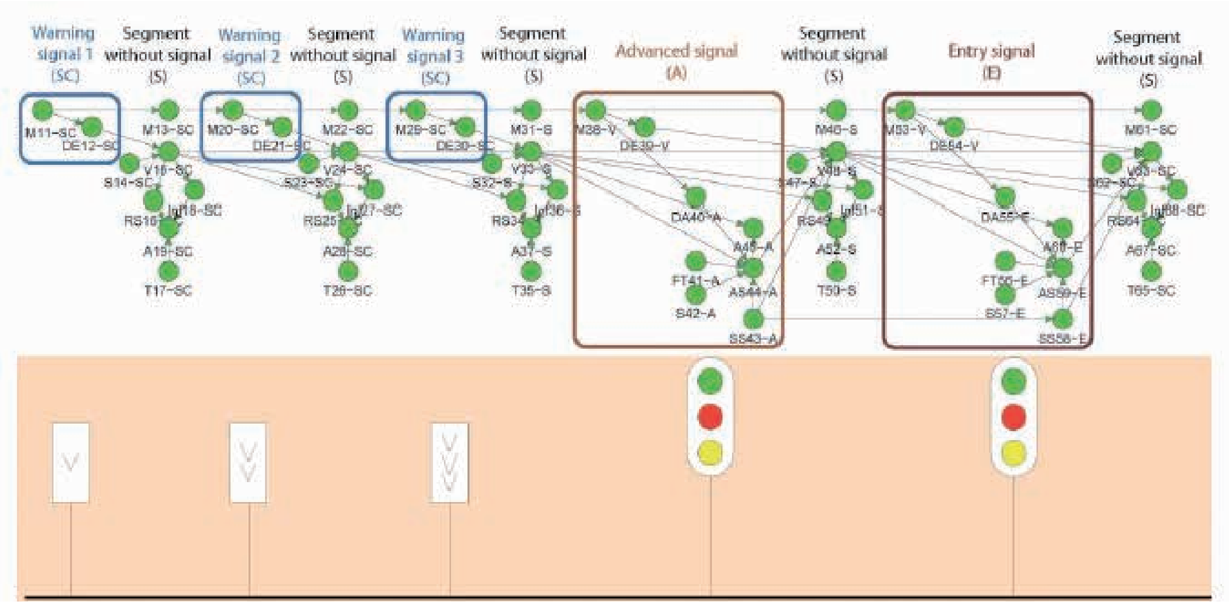


Figure 4. Illustration of the sub-Bayesian networks associated with warning and light signs showing the corresponding variables and links.

2.1.5.2 Temporal speed limit signal

Temporal speed limit signals consist of four signals (preannouncement, announcement, effective speed limit and end of limitation signals) and are related to their influential variables:

- Driver's attention state
- Speeds
- Accidents
- Rolling stock
- Terrain and infrastructure failures
- Driver's decision
- Technical failure
- Supervisor
- Signal state
- Driver's tiredness

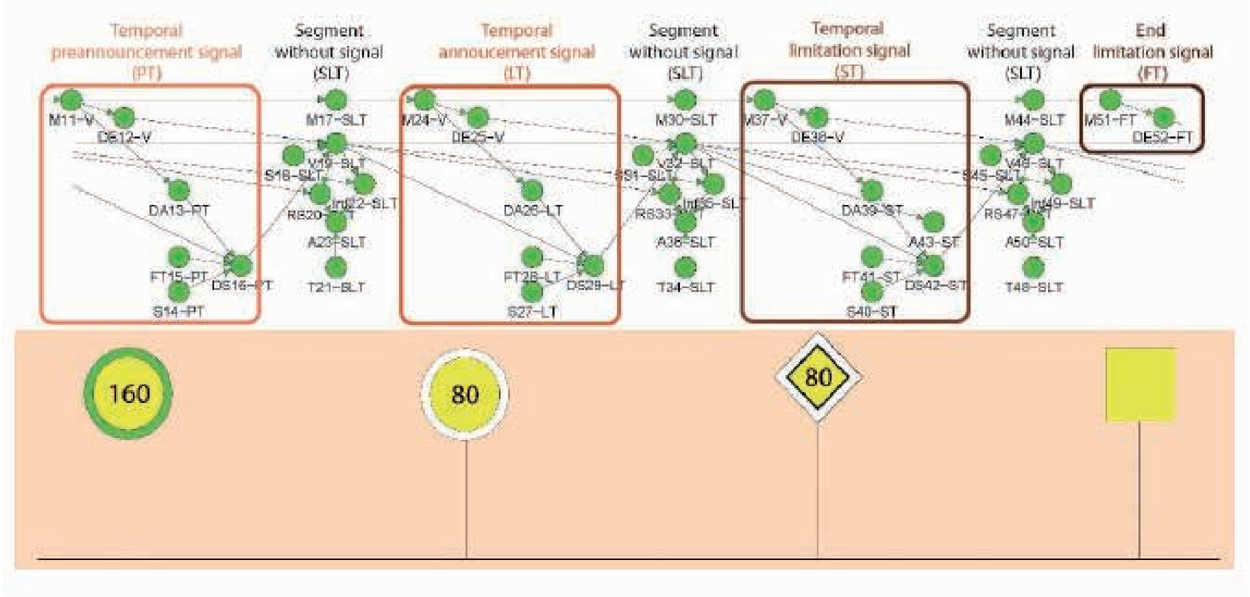


Figure 5. Illustration of the sub-Bayesian networks associated with temporal speed limit signs showing the corresponding variables and links.

2.2 Quantitative information of the Bayesian network. Conditional proba bilities.

Once the qualitative information of the Bayesian network is given, which includes the list of variables with the corresponding values and the lists of parents of each variable, we need to supply the quantitative information, which consists of a list of conditional probability tables, one per node, and contains the probabilities of each node given its parents.

One example is given in Figure 6, where the node Speed V is connected to its parents, nodes DE, Vp, S, AS and SS. This means that we need to give the probability of all combinations of the possible values of the child and the parents. To simplify this process it is convenient to provide a closed formula for the calculation of these probabilities in terms of a given set of parameters.

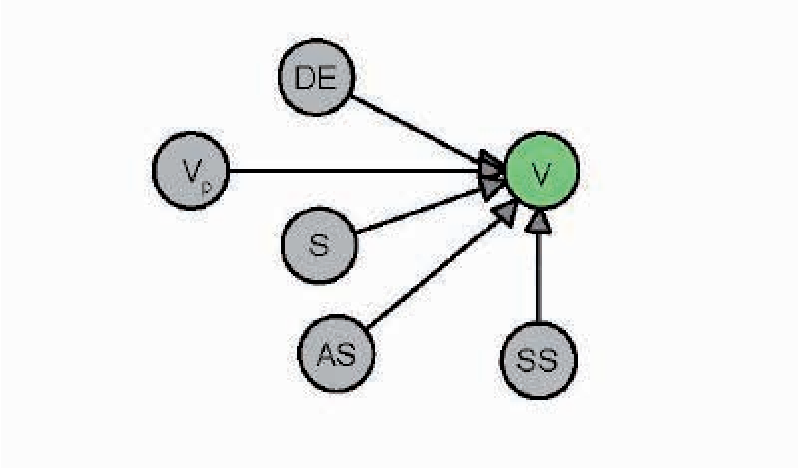


Figure 6. Speed node, V, connected to five parents (DE, Vp, S, AS and SS)

An example of closed formula for the conditional probability

$$p_{\mathrm{a,b,c,d,e,f}}(s) = P(\mathrm{V} = \mathrm{a} | \mathrm{V_p} = \mathrm{b}, \mathrm{DE} = \mathrm{c}, \mathrm{SS} = \mathrm{d}, \mathrm{AS} = \mathrm{e}, \mathrm{S} = \mathrm{f})$$

Is the following:

$$\begin{aligned} p_{\mathrm{a,b,c,d,e,f}}(s) \quad = \quad & \delta_{\mathrm{e},1}\left[\delta_{\mathrm{c},1}\delta_{\mathrm{a},s} + \delta_{\mathrm{c},2}\left((1-\rho_{\mathrm{f}})\delta_{\mathrm{a},s} + \rho_{\mathrm{f}}\delta_{\mathrm{a,b}}\right)\right. \\ & +\delta_{\mathrm{c},3}\left(\kappa_1\rho_{\mathrm{f}}\delta_{\mathrm{a},\max(1,s-1)} + (1-\rho_{\mathrm{f}}(\kappa_1+\kappa_2))\delta_{\mathrm{a},s}\right. \\ & \left.\left.+\kappa_2\rho_{\mathrm{f}}\delta_{\mathrm{a},\min(n,s+1)}\right)\right] + (\delta_{\mathrm{d},2}\delta_{\mathrm{e},2} + \delta_{\mathrm{d},3}\delta_{\mathrm{e},3})\left((1-\rho_{\mathrm{f}})\delta_{\mathrm{a},s}\right. \end{aligned}$$

Where the deltas are Kronecker's deltas and the rest are parameters.

2.3 Partition Technique

As the complexity of the problem grows with the square of the number of variables N and this number is normally above several thousands, it is convenient to partition the Bayesian network into much smaller subnets, so that the complexity grows linearly with N. In order to analyze the possible partitions it is enough to use the acyclic directed graph and to study the conditional independence between subnetworks. Figure 7 Illustrates how the Bayesian network can be partitioned in order to allow for the complexity to become linear in the number of variables and provide results in reasonable CPU times.

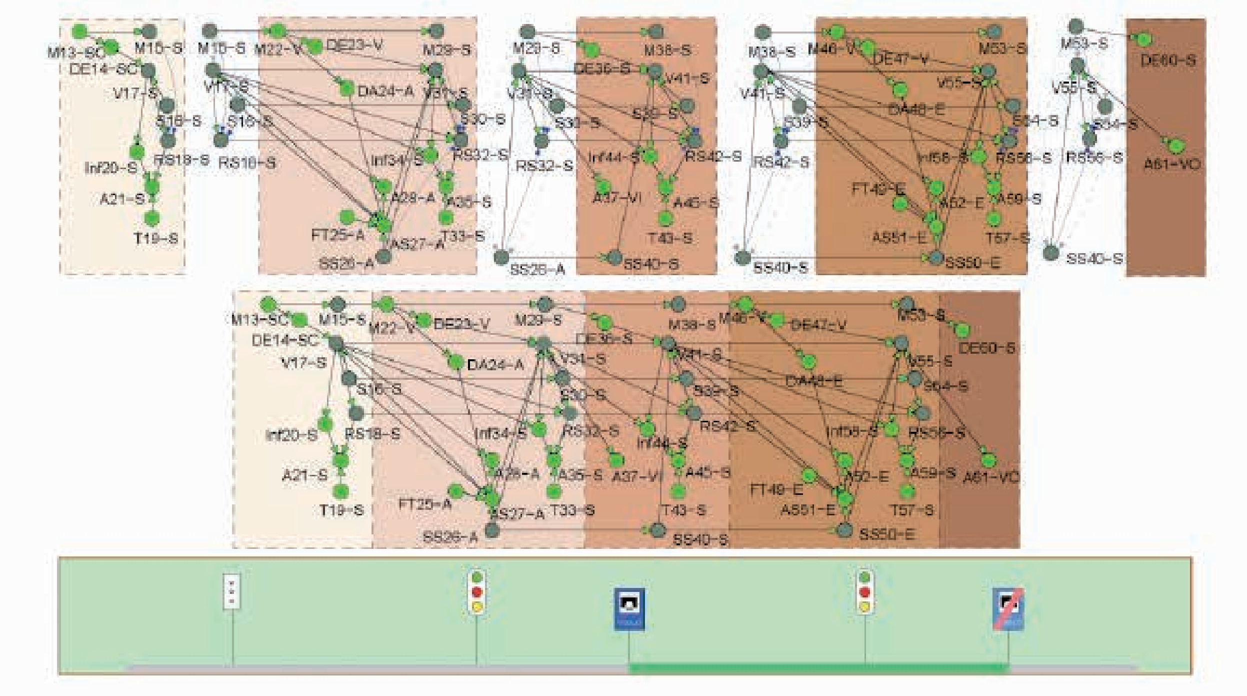


Figure 7. Illustration of how the Bayesian network can be partitioned in order to allow for the complexity to become linear in the number of variables and provide results in reasonable CPU times.

2.4 Backward analysis

The Bayesian network also allows to determine the causes of a given incident by proceeding backwards using available information and modifying the probabilities accordingly. Figure 8 shows the conditional probability tables before any information is supplied. For example, the probability of a severe accident number 33 is as low as 9.01 x 10^{-6}, signal 31 is in red 15% of the times, the probability of the driver to make an erroneous decision 14 is r, not obeying the sign, the speed was 220 km/h, there was a driver's erroneous decision14 "error I" is 6.01 x 10^{-5}, and the probability of the driver to be distracted was 2.84 x 10^{-7}.

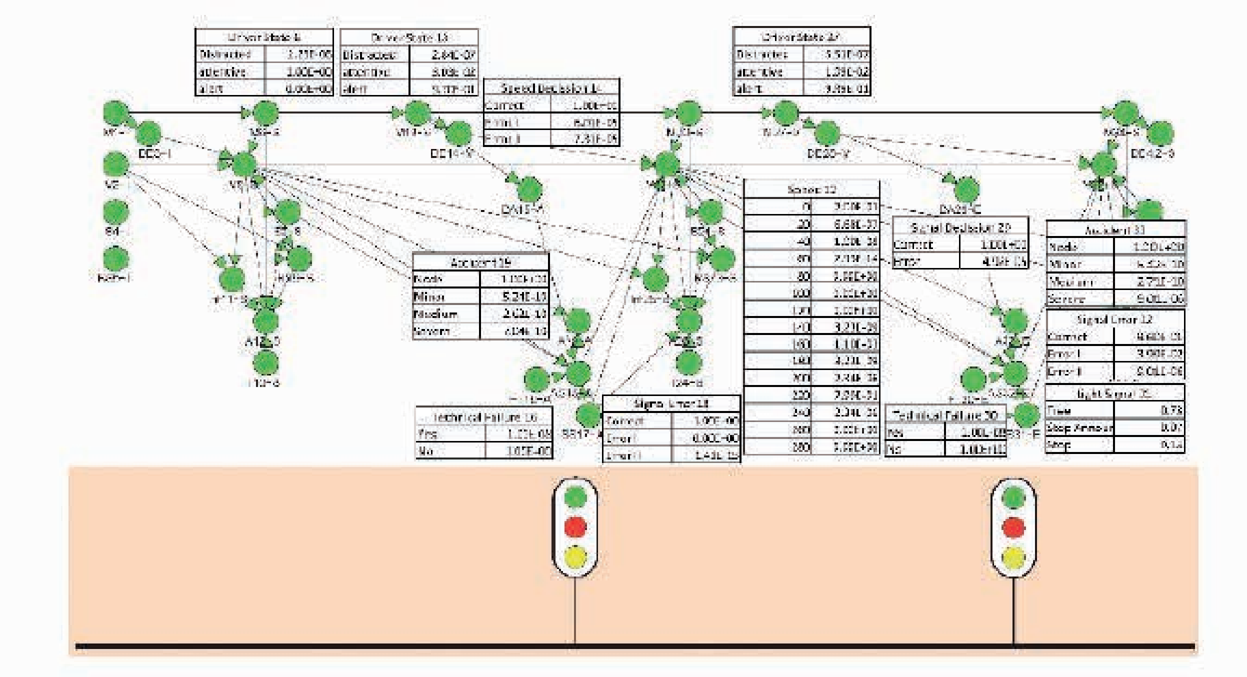


Figure 8. Conditional probability tables before any information is supplied.

However, after an accident 33 has occurred and evidence is obtained about the no occurrence of technical failures 16 and 30, their probabilities become one, as indicated in Figure 9 The Bayesian network techniques permit recalculating the probabilities of all other variables and consequently, provide the probabilities of any other event having information about the causes of accident 33. For example, Figure 9 informs us that signal 31 was in red, the driver made an error, not obeying the sign, the speed was 220 km/h, there was a driver's erroneous decision14 "error I", and this could be due to a distracted state or to an attentive state, but the later seems to be more probable.

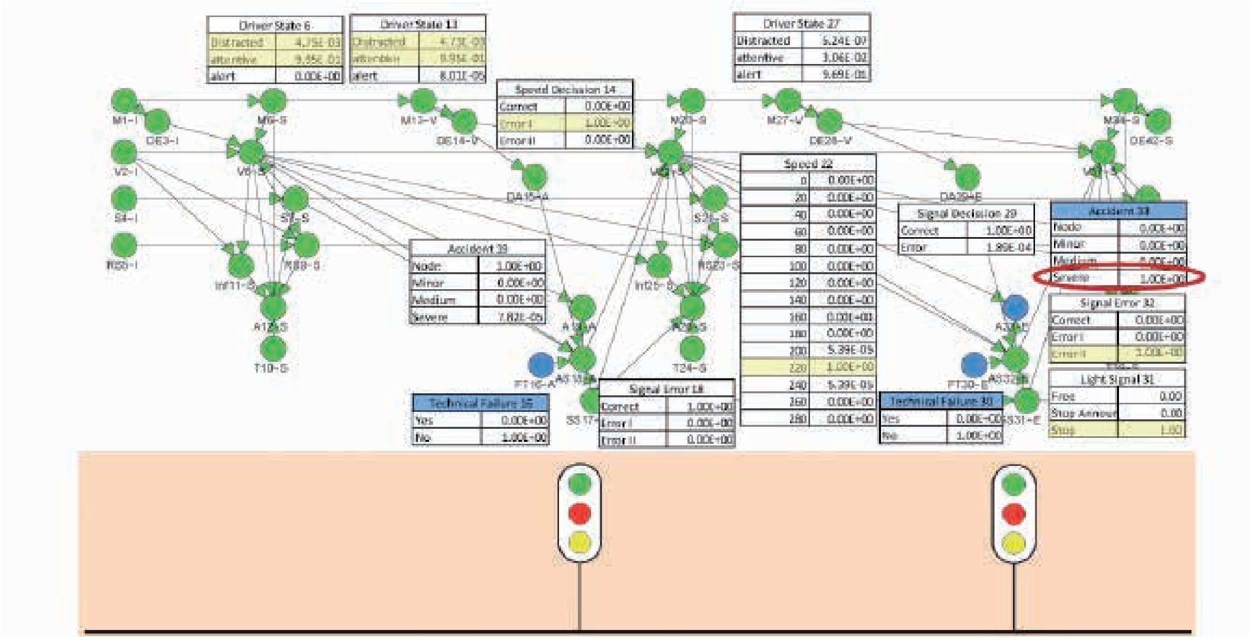


Figure 9. Conditional probability tables after a severe accident has occurred and no technical failures took place at two given locations close to the accident location.

All this shows that the power of Bayesian networks to facilitate a cause analysis study is really important.

2.5 Software development and provided information

The above described model has been implemented into a software package that allow us to perform the following tasks and provides the following information:

1. Check the input data for correctness and completeness.
2. Build the Directed acyclic graph of the Bayesian network including all variables and links.
3. Plot the Bayesian network.
4. Plot a scheme or diagram of the railway line including all its elements and corresponding Pks.
5. Calculate the marginal probabilities of the incident nodes and evaluate the ENSI (expected number of severe equivalent incidents) values of all items.
6. Plot the cumulated graph of ENSI values for identifying the most risky items.
7. Provide tables, sorted under different criteria, with the ENSI values and probabilities of all incidents along the line.

8. Plot the trace of the line with the location of the most risky items.
9. Provide tables with the circumstances that produce the most risky incidents in order to address the corrections to the real circumstances.

2.5.1 Necessary Information

The software for the probabilistic safety analysis (PSA) needs the following elements of information or input data:

1. *Railway regulations to be applied.*
2. *Line description.* A detailed description including the location and characteristics of switches, signals, level crossings, tunnels, viaducts, curves, etc.
3. *Driver's booklets.* With the characteristic safety regulations for each train and line including detailed maximum speeds, timing, etc.
4. *Train characteristics.* Power, maximum speeds, lengths, maximum accelerations and decelerations, etc.
5. *A video taken from the cabin in both directions*. These two videos are very important to identify risks and make decisions about the line safety.

Each element must be modeled properly by providing a line of code with the corresponding information, as shown below:

```
‘Underpass’, 370.35,’T’, ...
‘AnnouncementP’, 375.0, 0, 85, ...
‘AnnouncementGradeCrossing’, 376.0, ‘P’,2,...
‘SignalP’, 376.105, 1, 85, ...
‘CurveIn’, 376.215, 350,’R’, ...
‘GradeCrossing’, 376.350,’P’, ...
‘SignalA’, 378.2, ...
‘AnnouncementGradeCrossing’,379.765, ‘P’,1,..
‘CurveOut’, 379.768, 350, ...
‘SignalFP’, 379.775, 0, 85, ...
‘GradeCrossing’, 380.7,’P’, ...
‘ContinuousOFF’,380.85, ...
```

3. Examples

Next, several examples of application are presented to illustrate the possibilities of the proposed methodology. In particular, the safety problems are identified and resolved with a quantification of the safety level previously and after the corrections have been done. This quantifications makes one of the mean differences of the proposed methods with respect to existing ones.

3.1 Example of the line Zaragoza-Miranda

After performing the probabilistic safety analysis of the Zaragoza-Miranda line, the worst identified location corresponds to a permanent speed limit sign of 30 km/h, without the preannouncement sign, as illustrated in Figure 10, with an ENSI value of 0.244, which is extremely high. Since there is an end of speed limit (corresponding to a speed of 10 km/h) at PK 337.950, the driver will start to increase speed at this location to reach the maximum speed, which in this case is 160 km/h. However, since the distance to the speed limit sign is 1.050 km the train can reach only 60 km/h. Fortunately, the speed is limited to this low value, reducing the possible risk.

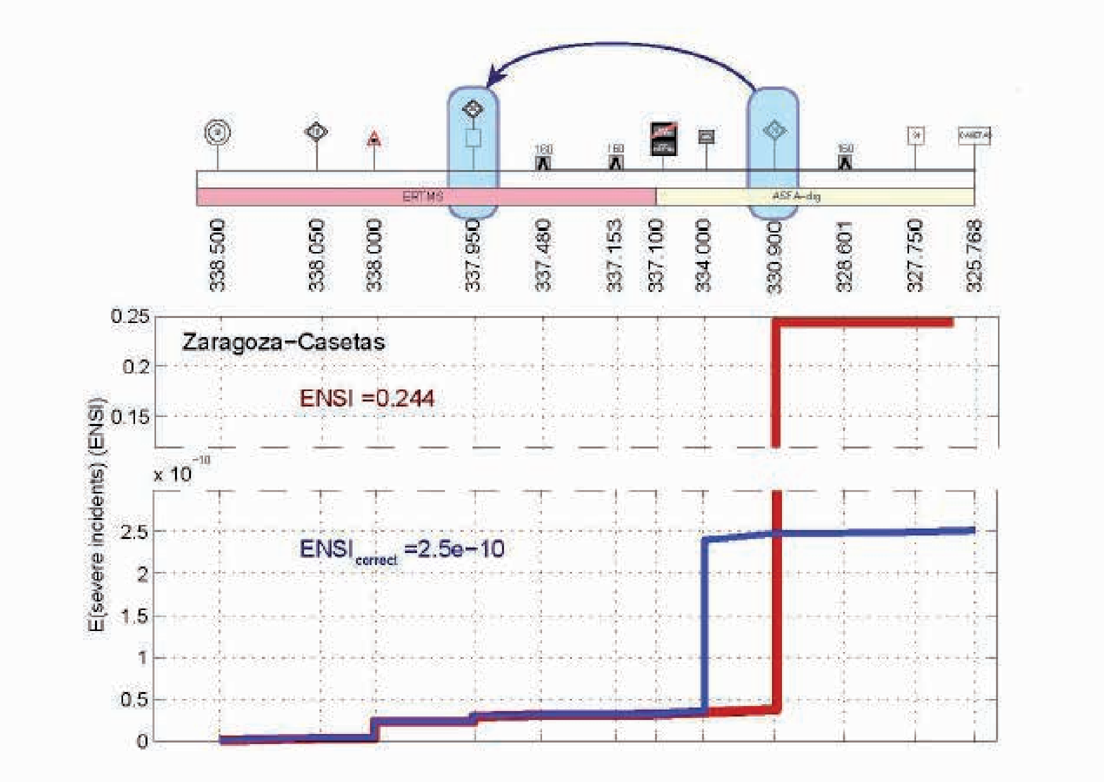


Figure 10. Probabilistic safety analysis of the Zaragoza-Casetas line

To correct this problem we move this speed limit sign 30km/h to PK 337.950, and remove the final speed limit sign at that location. In this way, the speed will be maintained at 30km/h without no increase and the problem is solved, as illustrated in Figure 10

3.2 Example of the line Palencia-Santander

After performing the probabilistic safety analysis of the Palencia-Santander line, one of the worst locations corresponds to a light signal located at PK 384.100 (see Figure 11). The main cause is an incorrect location of the advanced signal, which is too close to this light signal (at 258 m). Thus, the correction consists of moving this advanced signal to PK 382.916 (at 926m). By repeating the PSA the ENSI reduces from 4.77E-08 to 5.84E-12.

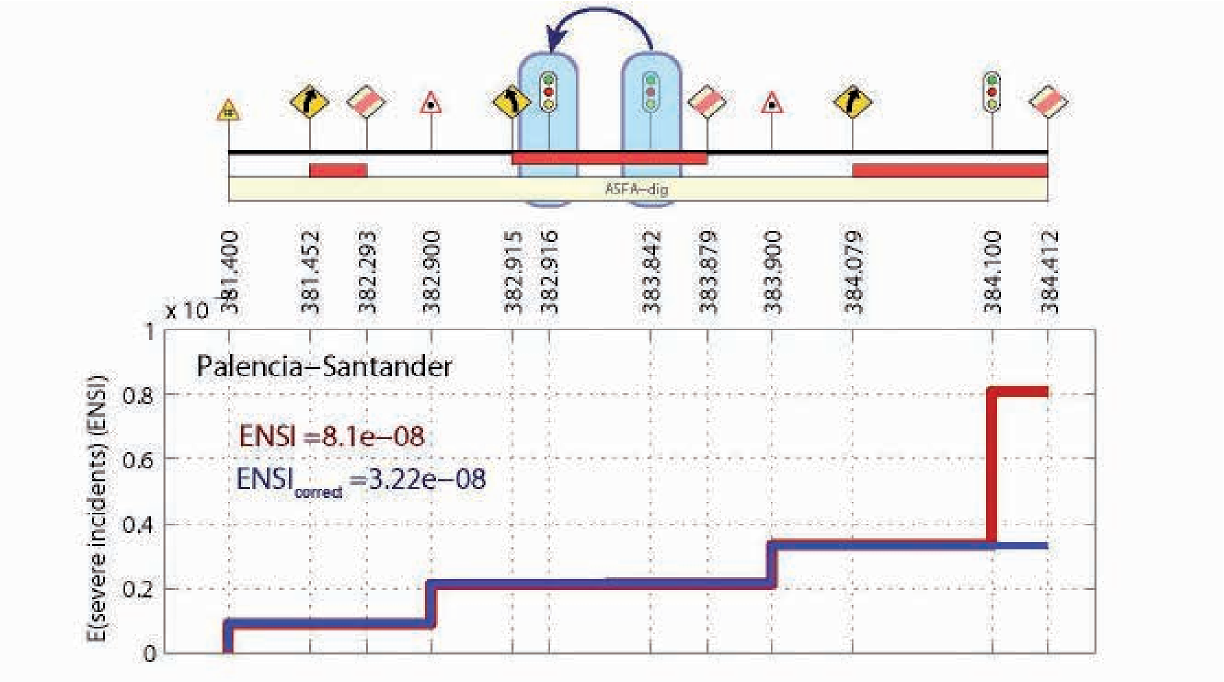


Figure 11. Probabilistic safety analysis of the Palencia-Santander line

3.3 Safety correction at a curve

In this example a safety problem due to an speed excess at a curve is corrected by means of the adequate speed limit signs. It can be seen in Figure 12 that the ENSI reduces from 0.00035 to 3.14 x 10 $^{-16}$. We note that speed excesses at a curve have been the main cause of important accidents in recent years, not only in Spain but in other countries of the EU and the United states of America. In particular, the responsibility cannot be given only to the driver (engineer or conductor) and ATP systems must be incorporated to avoid this type of accidents.

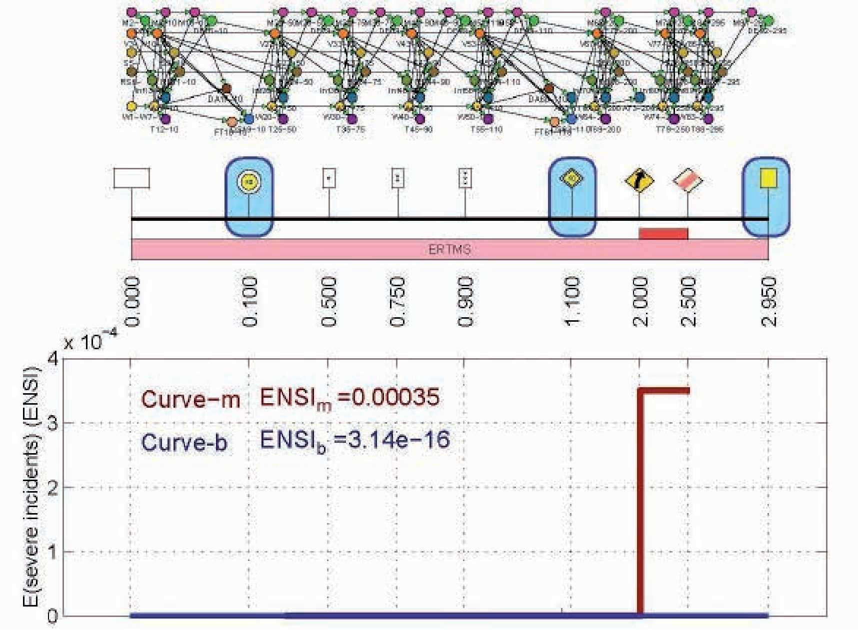


Figure 12. Illustration of how the safety improves at a curve when the speed limit is reduced.

3.4 Safety correction at a grade crossing

In this example a safety problem due to an insufficient announcement of a grade crossing is corrected by means of an adequate announcement sign and corresponding protection. It can be seen in Figure 13 that the ENSI reduces from 3.42 x 10 $^{-9}$ to 4.81 x 10 $^{-10}$. Grade crossing are known to be responsible for a large number of rail accidents and if possible thsy must be eliminated. On the other hand they produce important increases in travel times because of the deceleration and acceleration phases required when approaching grade crossing locations.

3.5 Safety improvement at a light signal

In this example the safety at a light signal is improved first by means of warning signs and later by the ERTMS (an ATP system). It can be seen in Figure 14 that the ENSI reduces from $1.09\cdot10^{-9}$ to $1.24\cdot10^{10}$ due to the warning signs and later to 6.48 x 10 $^{-14}$ due to the ERTMS. The light signal was initially under SR (staff responsible) protection and after after installing the warning signs it improves safety in one order of magnitude. Finally, installing the ERTMS safety improves in more than three orders of magnitude. It is important to enphasize the role played by the three warning signs, which produce an important improvement of the driver attention and thus increase safety substantially.

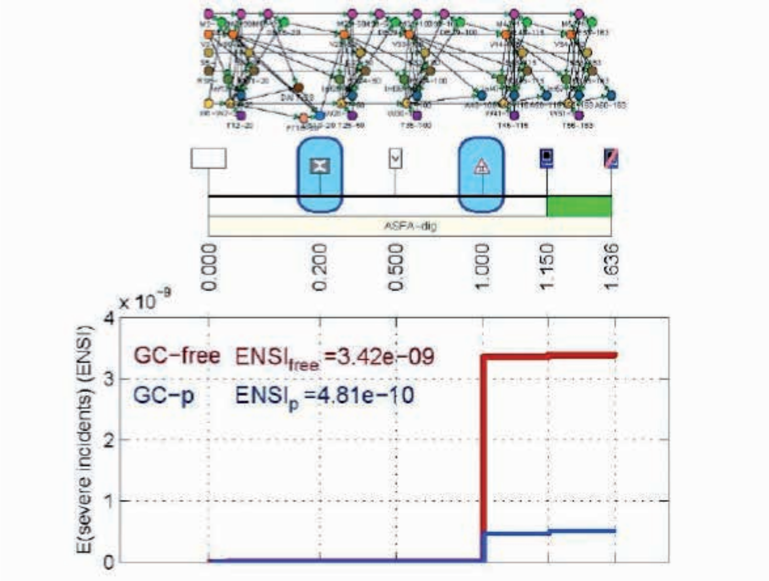


Figure 13. Illustration of how the safety improves at a grade crossing when by protection.

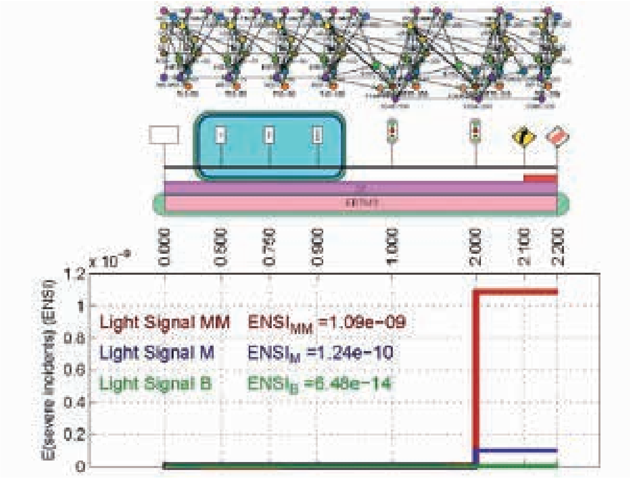


Figure 14. Illustration of how safety improves at a light signal by warning signs and the ERTMS.

3.6 Safety correction of a speed limit signal

In this example the safety at a speed limit signal is improved. Initially an erroneous announcement of a 120 km/h speed instead of 80 km/h causes the safety problem leading to an ENSI value of 0.145, which is extremely high (see Figure 15). This is due to the fact that the ASFAdig does not cover the speed excesses. To correct this sittuation, an additional announcement of 80km/h has been incorporated (see the blue shadowed plot in Figure 15) leading to an ENSI value of 3.52 x 10^{-9}, which is satisfactory. In addition if the ASFA-dig is replaced by the ERTMS, the ENSI reduces further to a value of 1.58 x 10^{-12}, which is three orders of magnitude smaller (see the green shadowed plot in Figure 15). This example illustrates the importance of speed limit signs, which can be permanent or temporal. The fact that permanent signs are not covered by certain ATP systems should be corrected to avoid accidents. In addition, the location of the sequences of speed limit signs must be done very carefully.

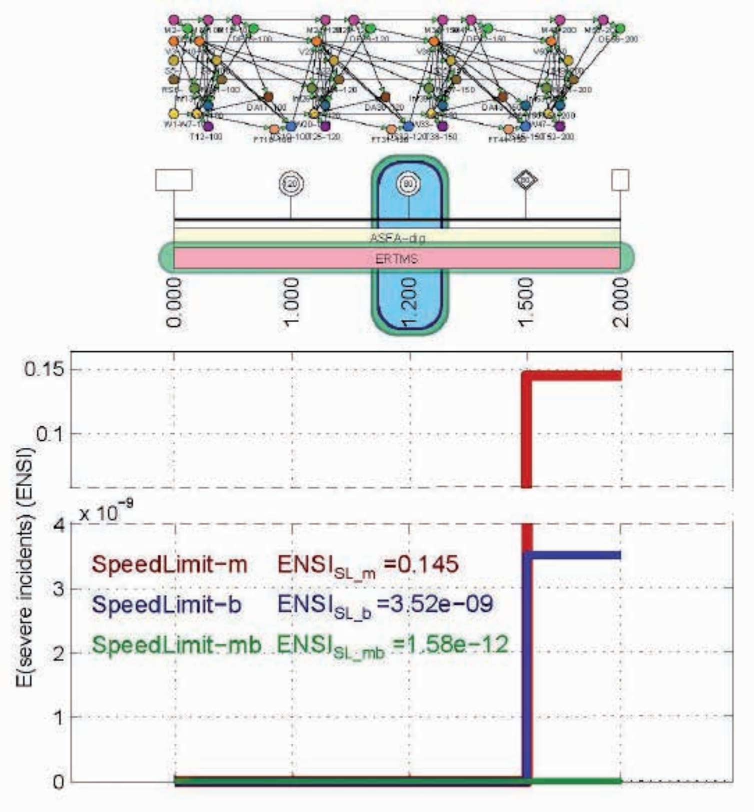
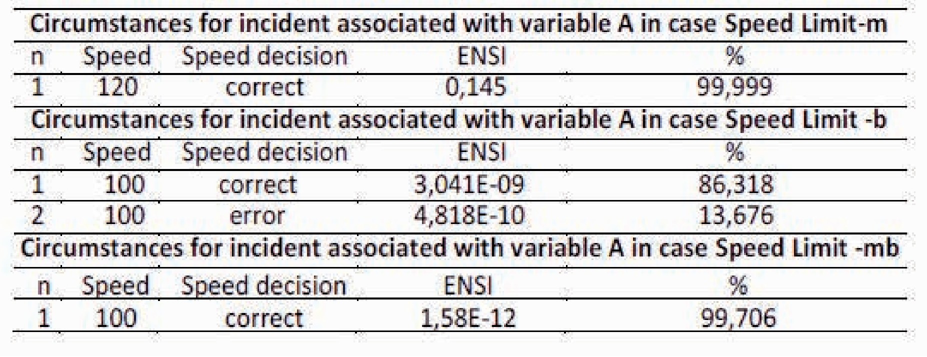


Figure 15. Illustration of how the safety improves at a speed limit signal by correcting errors.

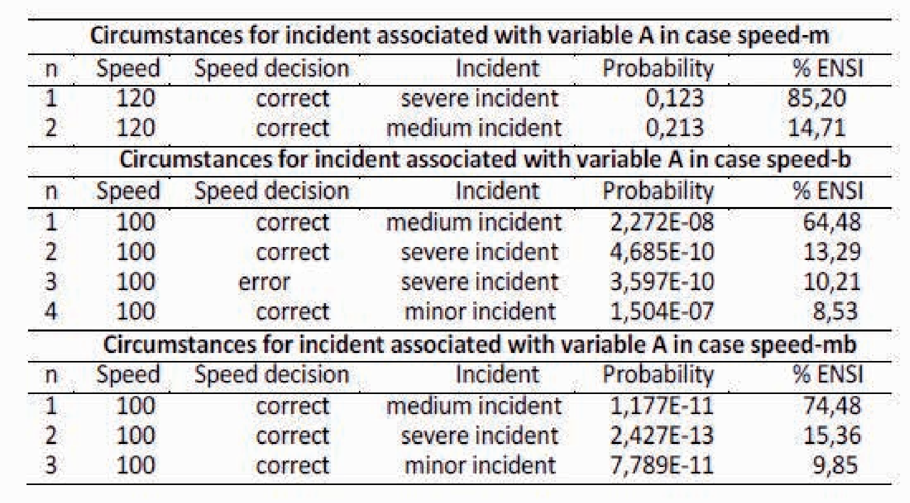
In Table 6 the circumstances under which the incidents at the speed limit signs are shown for the cases speed-m, speed-b and speed-mb, respectively. It is interesting to see that the case speed-m in which an erroneous announcement was located, the incident occurs, as expected, at this speed, with no error in the driver decision at practically all cases. In case speed-b, the speed reduces to 100 km/h and the driver decision remains correct in 86,318 % of the cases, but erroneous in 13,676% of the cases. Finally, in case speed-mb we have again no error and a speed of 100 km/h in almost all cases.

Table 6. Circumstances for incidents associated with variables incident for cases speed-m, speed-b and speed-mb showing the associated ENSI and corresponding percentages.



Circumstances for incident associated with variable A in case Speed Limit-m				
n	Speed	Speed decision	ENSI	%
1	120	correct	0,145	99,999
Circumstances for incident associated with variable A in case Speed Limit -b				
n	Speed	Speed decision	ENSI	%
1	100	correct	3,041E-09	86,318
2	100	error	4,818E-10	13,676
Circumstances for incident associated with variable A in case Speed Limit -mb				
n	Speed	Speed decision	ENSI	%
1	100	correct	1,58E-12	99,706

Table 7. Circumstances for incidents associated with variables incident for cases speed-m, speed-b and speed-mb showing the Probabilities and % ENSI values.



Circumstances for incident associated with variable A in case speed-m					
n	Speed	Speed decision	Incident	Probability	% ENSI
1	120	correct	severe incident	0,123	85,20
2	120	correct	medium incident	0,213	14,71
Circumstances for incident associated with variable A in case speed-b					
n	Speed	Speed decision	Incident	Probability	% ENSI
1	100	correct	medium incident	2,272E-08	64,48
2	100	correct	severe incident	4,685E-10	13,29
3	100	error	severe incident	3,597E-10	10,21
4	100	correct	minor incident	1,504E-07	8,53
Circumstances for incident associated with variable A in case speed-mb					
n	Speed	Speed decision	Incident	Probability	% ENSI
1	100	correct	medium incident	1,177E-11	74,48
2	100	correct	severe incident	2,427E-13	15,36
3	100	correct	minor incident	7,789E-11	9,85

Table 7 is very similar, but now we include a column with the probability for each incident severity and the corresponding ENSI percentage. It is interesting to see the probabilities of minor, medium and severe incidents for each case. As expected minor incidents occur at the speed limit sign more frequent than medium and these more frequently than severe incidents.

4. Conclusions

Bayesian network models provide an important tool to perform a probabilistic safety assessment of railway lines, and are more powerful than the commonly used fault and event trees, especially when common causes are present.

The proposed model reproduces all the variables involved in the problem, including their qualitative dependencies and the quantification of the associated conditional probabilities.

In particular, human error must be carefully considered in an integrated form, that is, considering how it depends on tiredness and attention levels, as being one of the most important factors in the safety of railway networks and lines.

A simple list of items can be given for a computer program to build the acyclic graph associated with the Bayesian network automatically.

The proposed partitioning technique reduces the initial nonlinear complexity to a complexity which is linear with the number of nodes.

The examples analyzed in this article show that the method is able to identify and quantify relevant incidents and their probabilities of occurrence.

The backward possibilities of the Bayesian network permits to analyze the causes of incidents and especially those leading to fatal accidents.

The most critical part of the proposed model is the parameter estimation and calibration, which must be done with the collaboration of various groups of experts.

5. Bibliography

- Amit, I. & Goldfarb, D., 1971. The timetable problem for railways. *Developments in Operations Research,* Volume 2, pp. 379-387.
- Assad, A., 1980. Models for rail transportation. *Transportation Research Part A,* Volume 14, pp. 205-220.
- Beales, L., 2002. *Guidance on the Preparation of Risk Assessments within Railway Safety Cases,* s.l.: RSSB.
- Burdett, R. L. & Kozan, E., 2010. A disjunctive graph model and framework for constructing new train schedules. *European Journal of Operational Research,* Volume 2010, pp. 85-98.
- Cacchiani, V. & Toth, P., 2012. Nominal and Robust Train Timetabling Problems. *European Journal of Operational Research,* Volume 219, pp. 727-737.
- Caprara, A., Fischetti, M. & Toth, P., 2002. Modeling and solving the train timetabling problem. *Operations Research,* Volume 50, pp. 851-861.
- Carey, M., 1994. A model and strategy for train pathing with choice of lines, platforms and routes. *Transportation Research Part B,* Volume 28, pp. 333-353.
- Carey, M. & Crawford, I., 2007. Scheduling trains on a network of busy complex stations. *Transportation Research Part B,* Volume 41, pp. 159-178.
- Carey, M. & Lockwood, D., 1995. A model, algorithms and strategy for train pathing. *Journal of the Operational Research Society,* Volume 46, pp. 988-1005.
- Castillo, E. et al., 2015. An alternate double-single track proposal for high-speed peripheral railway lines. *Computer Aided Civil and Infrastructure Engineering,* Volume 30, pp. 181201.
- Castillo, E., Gallego, I., Ureña, J. M. & Coronado, J. M., 2009. Timetabling optimization of a single railway track line with sensitivity analysis. *TOP,* Volume 17, pp. 256-287.
- Castillo, E., Gallego, I., Ureña, J. M. & Coronado, J. M., 2011. Timetabling optimization of a mixed double- and single-tracked railway network. *Applied Mathematical Modelling,* Volume 35, pp. 859-878.
- Castillo, E., Grande, Z., Moraga, P. & Sánchez Vizcano, J., 2016. A time partitioning technique for railway line design and timetable optimization. *Computer Aided Civil And Infrastructure Engineering,* Volume 31, pp. 599-616.
- Cordeau, J. F., Toth, P. & Vigo, D., 1998. A survey of optimization models for train routing

and scheduling. *Transportation Science,* Volume 32, pp. 380-404.

- D’Ariano, A., Pacciarelli, D. & Pranzo, M., 2007. A branch and bound algorithm for scheduling trains in a railway network. *European Journal of Operational Research,* Volume 183, pp. 643-657.
- D’Ariano, A. & Pranzo, M., 2004. *A real time train dispatching system based on blocking time theory.* Delft, DUP Science, pp. 129-152.
- D’Ariano, A., Pranzo, M. & Hansen, I. A., 2007. Conflict Resolution and Train Speed Coordination for Solving Real-Time Timetable Perturbations. *IEEE Transactions on Intelligent Transportation Systems,* Volume 8.
- Haghani, A. E., 1987. Rail freight transportation: a review of recent optimization models for train routing and empty car distribution. *Journal of Advanced Transportation,* Volume 21, pp. 14-172.
- Hellström, P., 1998. *Analysis and evaluation of systems and algorithms for computer-aided train dispatching,* Sweden: s.n.
- Higgins, A., Kozan, E. & Ferreira, L., 1996. Optimal scheduling of trains on a single line track. *Transportation Research Part B,* Volume 30, pp. 147-161.
- Jia, L. M. & Zhang, X. D., 1993. Distributed intelligent railway traffic control based on fuzzy decision making. *Fuzzy Sets and Systems,* Volume 62, pp. 255-265.
- Kraay, D. R. & Harker, P. T., 1995. Real-time scheduling of freight railroads. *Transportation Research Part B,* Volume 29, pp. 213-229.
- Lin, D. Y. & Ku, Y. H., 2013. Using Genetic Algorithms to Optimize Stopping Patterns for Passenger Rail Transportation. *Computer-Aided Civil and Infrastructure Engineering,* pp. n/a--n/a.
- Ouyang, Y. et al., 2009. Optimal Locations of Railroad Wayside Defect Detection Installations. *Computer-Aided Civil and Infrastructure Engineering,* Volume 24, pp. 309-319.
- Pachl, J., 2014. Railway Timetabling and Operations. Analysis - Modelling - Optimisation - Simulation - Performance Evaluation. In: Hamburg(): Eurailpress, pp. 23-24.
- Petersen, E. R., Taylor, A. J. & Martland, C. D., 1986. An introduction to computer aided train dispatching. *Journal of Advanced Transportation,* Volume 20, pp. 63-72.
- Sahin, I., 1999. Railway traffic control and train scheduling based on inter-train conflict management. *Transportation Research Part B,* Volume 33, pp. 511-534.
- Yang, Z. & Hayashi, Y., 2002. GIS-Based Analysis of Railways Origin/Destination Path-Selecting Behavior. *Computer-Aided Civil and Infrastructure Engineering,* Volume 17, pp. 221-226.

Safety and signalling systems

360.revista de alta velocidad
número 5 - junio - 2018. Pág 149 - 157

Precise and reliable localization as a core of railway automation (Rail 4.0)

Hutchinson, Michael
Marais, Juliette
Masson, Émilie
Mendizabal, Jaizki
Meyer zu Hörste, Michael

NSL, IFSTTAR, RAILENIUM, CEIT, DLR [1]

Abstract:

High Speed Railway services have shown that Railways are a competitive and, at the same time, an environmentally friendly transport system. The next level of improvement will be a higher degree of automation, with partial or complete automatic train operation up to fully automatic unattended driverless operation to reduce energy consumption and noise, as well as improving punctuality and comfort. Based on extensive experience in separated railway systems such as Metros, VAL (Véhicule automatique léger) and subways, today's discussion focuses on the fully automatic train operation on regular railway lines. This introduces some questions that are more complex than in today's systems: the performance, length and weight of the trains spreads in a wider range; there are other and more complex operations; components such as level crossings in different equipment variants are added; requirements are higher and the technological equipment point is more heterogeneous. Many evolutionary or innovative approaches for railway automation rely on precise and reliable localization. Especially for higher levels of automation it is essential. Starting from existing technologies such as the European Train Control System (ETCS), moving on to advanced driver assistance systems for local automatic operation up to full automatic train operation (ATO), all of these solutions depend on knowledge of the position of the train on the network. Improved concepts for railway operation e.g. using moving block, on-board integrity supervision or virtual train sets will result in even higher requirements for accuracy and reliability of the localization. The contribution shows an approach based on Global Navigation Satellite Systems (GNSS) as a priority source of information used in combination with sensor data fusion. Elements such as trackside augmentation as well as the use of digital maps will be discussed. Different approaches are considered, e.g. those currently under development by projects in the innovation program in the joint undertaking Shift2Rail.

Keywords: localisation, odometry, GALILEO, train control, digital route map

---

[1] *Hutchinson, Michael. Nottingham Scientific Limited (NSL). Email: michael.hutchinson@nsl.eu.com*
Marais, Juliette. French institute of science and technology for transport, spatial planning, development and networks. (IFSTTAR) Email: juliette.marais@ifsttar.fr
Masson, Émilie. Innovation R&D Project coordinator RAILENIUM. Email: emilie.masson@railenium.eu
Mendizabal, Jaizki. Centre for Technical Research and Studies - CEIT). Email: jmendizabal@ceit.es
Meyer zu Hörste, Michael. German Aerospace Centre (DLR). Email: Michael.MeyerzuHoerste@dlr.de. (corresponding author)

1. Introduction

The precise as well as reliable detection of the position and their safe communication play a key role for the train operation control system. In recent years, the tendency to shift more and more positioning equipment functionality for safety-critical applications from the track side to the vehicle on-board-side could be observed. Train-borne positioning systems enable a costeffective operation technology for railway lines with a low traffic density and a more flexible operation than track-side positioning systems can offer for lines with a high traffic density. For on-board positioning systems, such as satellite-based train positioning systems, the mapmatching process is of particular importance, since then the data provided by different positioning sensors, such as GALILEO receivers, Doppler radars, odometers and inertial sensors are related to a digital map of the respective railway network. This implies that the digital map is a key element in a positioning system for safety-critical applications.

2. State of the art

Train location is of main use in signalling applications. In Europe, the train position is traditionally processed with the help of equipment on tracks. The typical equipment is a track circuit, i.e. a simple electrical device used to detect the presence of a train on rail tracks. This equipment is thus not devoted to locating the train specifically but to locating it indirectly on a track portion. The location can also be determined with the help of detectors placed along the track, on which the train protection relies. These sensors can be transponders (Balises), which communicate with the train on-board equipment when the train runs over them. In order to harmonize European solutions, Europe has developed the European Train Control System (ETCS) for the signalling, control and train protection. In levels 1 and 2 of ETCS, a Balise or a group of Balises is installed on tracks to give a passing train a position reference. The Balise initializes the odometer, and the train position is computed by the odometer as a distance run since the last relevant Balise group. The system composed of a Balise and odometer computes the position of the train interfaced with the EVC (European Vital Computer) train-borne subsystem (cf. figure 1). In level 3 of ETCS, the train location shall be sent by the train itself to the ground [UNISIG SUB026]. No line side signals will be required for delivering movement authorities. All information will be exchanged between the ETCS on-board system and the Radio Block Center (RBC) trackside system through mobile networks. Two main pieces of information are communicated by the train to the RBC: its location and the confirmation that the train did not lose any wagon (train integrity). Fig. 1 shows the generic structure of the onboard train control system.

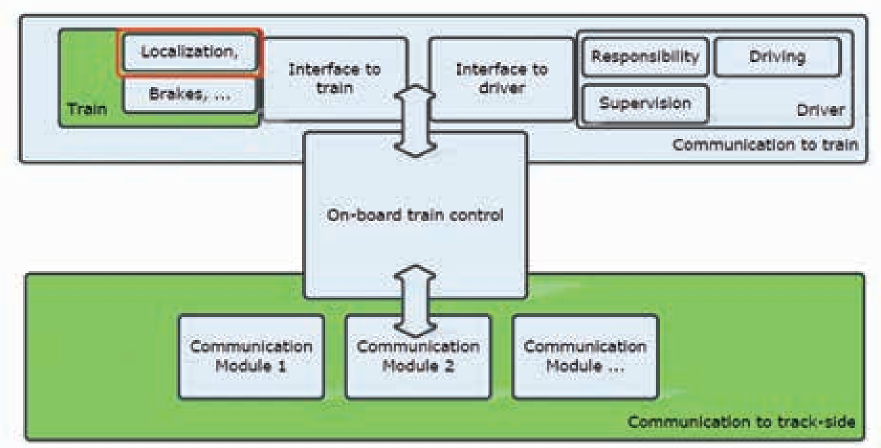


Figure 1. Overview of current setup of the localisation in the on-board train control

The last level of ETCS shall also improve line capacity by making it possible to manage circulations with moving blocks. ERTMS Level 3 will allow reducing the number of trackside equipment on the one hand, and the implementation of moving block in order to reduce the train spacing and then enhance capacity on the other hand.

Mainly driven by economic reasons and thus, acceleration of ERTMS penetration into the networks, the use of GNSS-based systems and, in particular, their introduction in signalling systems is seriously investigated today as a positioning opportunity included in ETCS and tested all around the world. In this context, GNSS is investigated to be the basis for a new embedded train locator.

The use of GNSS for low cost signalling solution and, in particular, in the highest level of ETCS (level 3) and the ERTMS Regional operation context, is an issue that has been investigated since the beginning of the 2000's [Raymond, 2004]. The concept of "virtual Balises" emerged to allow this new technology to penetrate the railway operational modes. The virtual Balises are dematerialized points, recorded in an embedded geographic database as illustrated in figure 2. These points are basically coordinates of the real (removed) physical Balises. The goal of the virtual Balise concept is to detect the position of the train when passing over the identified points by comparing the GNSS-based location of the train with the database. The train delivers then the same telegram that would have been sent with the use of the physical Balise. As such, the process could be as transparent as possible for the global system. The virtual Balise has been investigated in lots of projects since the beginning of the studies such as in RUNE [Albanese, 2004] or recently in 3InSat [Rispoli, 2013] or ERSAT [Facchinetti, 2015].

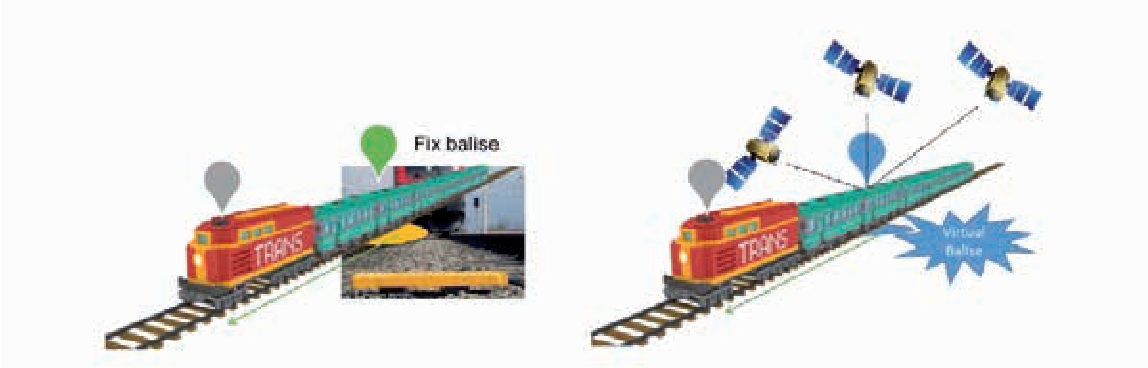


Figure 2. Virtual Balise use instead of physical trackside Balise

With several projects, the European Commission, through the successive Framework Programs, funded research in order to explore and promote the use of global navigation satellite systems for low cost signalling solutions and, in particular, in the highest level of ETCS (level 3) and the ERTMS Regional. First projects were APOLO [Filip, 2001], GADEROS [Urech 2002] and LOCOPROL [Mertens, 2003] but one can mention tens of others in the past decade until the recent Satloc [Gradinariu, 2017], GaLoROI [Manz, 2014] or 3inSat [Rispoli, 2014] projects. The concept of "virtual Balises" has been investigated in lots of projects since the beginning of the studies such as in RUNE [Albanese, 2004] or recently in 3InSat [Rispoli, 2013] or ERSAT [Facchinetti, 2015]. On-going projects are STARS [Gurnik, 2016] and Rhinos [Neri, 2016] focusing respectively on GNSS signal reception in railway environments and integrity concepts for railways. A complementary national project was DemoOrt in Germany [Meyer zu Hörste, 2009]

One European operated line embeds a light GNSS solution today. It is a single track line of 150 km length with a system developed by the University of Applied Sciences Upper Austria (FH OO) and operated by Stern & Hafferl, respective Salzburg AG. An on-board unit communicates the GNSS-based train position to the trackside train controller which is responsible for the movement authority of the train [Stadlmann; 2006]. The system is implemented in SIL2.

In April 2012, a big recognition step has occurred when the European Commission, the European Railway Agency and the European Rail Sector Associations signed together an ERTMS Memorandum of Understanding concerning the strengthening of cooperation for the management of ERTMS. It mentions in particular that GNSS can play a major role in the rail sector [ERTMS, 2012].

In the same period, new initiatives have been launched such as the NGTC project (Next Generation Train Control) that aims to study how the new developments for ERTMS/ETCS (for interoperable networks) and those for CBTC (Communication-Based Train Control) systems (for urban networks) can be mutualized. Satellite positioning is one of the NGTC focuses.

GNSS provides a means of absolute positioning with achievable accuracies at the order of metres using basic consumer-level devices to the order of centimetres based on more sophisticated solutions. It is today used in an extremely wide range of applications from smartphone navigation to network timing synchronisation in critical infrastructure to safetycritical aircraft operations. In the railways, GNSS is today widely used for applications where its use is not critical such as basic tracking and synchronisation of onboard systems. The industry is also starting to recognise the potential for using GNSS for more advanced applications, from those that are regulated and include an element of liability to train control and operations. For example, GNSS is currently used for Selective Door Opening systems which have an associated safety-criticality. Many of the more advanced applications have demanding requirements which are not straightforward to meet with GNSS due to the challenges presented by the railway environment in terms of GNSS signal availability and quality. Features such as tunnels, cuttings, trackside buildings and vegetation can block signals as well as reflect them, inducing multipath effects in GNSS receivers. There is a general consensus that GNSS must be integrated with other sensors and sources of information in order to meet the very demanding safety requirements of train control applications, which can be more demanding than in civil aviation applications. GNSS is also vulnerable to interference and spoofing threats and solutions must be identified to mitigate this.

In recent years there has been significant research and development undertaken into the Virtual Balise application. This uses an onboard GNSS-based positioning system to detect when a train passes well-defined ERTMS ‘Balise’ locations stored in a database onboard, enabling the odometer calibration to be performed. This allows physical Balises to be removed from the track, making savings in terms of CAPEX and OPEX (maintenance). The current focus for this application is on local and regional lines, largely since it represents an economically justifiable way of introducing ERTMS onto such lines which will improve safety and/or capacity. Other advanced developments include the mandated Positive Train Control concept in the USA, for which GNSS is an optional source of positioning, and the research into the use of GNSS for Degraded Mode Working in the UK.

3. The Role of On-Board Localisation in Future Train Control Systems

3.1 The SmartRaCon approach in Shift2Rail

The European research for Railways will be done for the next coming years mainly in the Shift2Rail Joint Undertaking. The European Union, two railways and six industry members founded the Joint Undertaking.

The Joint Undertaking Shift2Rail comprises five Innovation Programs (IP) and five Cross Cutting Activities (CCA) of railway research and development for the near future. More than 40 Technical Demonstrators will be developed in the field of rolling stock, traffic management and control systems, infrastructure, IT solutions and rail freight.

Four Partners with a strong expertise in research and development in the area of train control, communication and localisation formed the consortium “Smart Rail Control - SmartRaCon”:

- German Aerospace Centre (Deutsches Zentrum für Luft- und Raumfahrt – DLR), Germany
- Centro de Estudios e Investigaciones Técnicas (CEIT), San Sebastián (Gipuzkoa), Spain
- Fondation de Cooperation scientific RAILENIUM, Famars, France
- Nottingham Scientific Ltd. (NSL), Nottingham, England

The vision of SmartRaCon is to realise a fail-safe, multi-sensor onboard positioning system at minimal cost which not only provides train positioning to the ETCS kernel but also acts as an enabler for multiple areas within the area of signalling, potentially including Traffic Management, Train Integrity and Virtual Coupling. The objective of the consortium is to develop the constituents of the on-board positioning system based on multi-constellation GNSS, complemented by other positioning technologies. The activities will also include the definition of the expected performance thanks to field-testing, certification process definition and testing tools.

3.2 Concept and Approach

The overall concept is based on the need to ensure that the safety levels provided by existing signalling and control systems are not compromised in the movement towards an on-board positioning system. SmartRaCon will provide various contributions to ensure that this is achieved. This includes setting up and undertaking test campaigns, analysing the data from such campaigns, improving specifications, providing various inputs to the development of a safety case, as well as other more specific contributions which build on the positioning technology expertise within the consortium such as simulation based KPI evaluation, multiconstellation or sensor integration, etc.

Testing processes and the route to acceptance of GNSS and associated technology will be enhanced such that standardised methods are put in place in terms of the equipment used, measurements made, and analysis tools and results delivery. This includes a Route Clearance service, simulation tools for railway KPIs evaluation, Digital Route Maps (DRM) as an input to on-board positioning systems, in terms of their utilisation, distribution and management. Contributions will be made to the formation of a consolidated set of specifications, critically including methodology for testing the capabilities of common-off-the-shelf equipment against these specifications. There will be a focus on DRM technology and the Virtual Balise concept. The need for lab simulations will be identified here and it is proposed to develop a Local Environment Model as a specific task. Based on the outcomes of initial testing, solutions for performance optimisation will be proposed through in-depth knowledge in the hybridisation of GNSS with inertial sensors, odometry, dead reckoning, DRM and Wireless Communications Technology (WCT). Further specific tasks proposed are the development of a GNSS Railway Integrity Concept, input to the design of a Safety-Critical Railway GNSS Receiver, a Failure Modes and Effects Analysis (FMEA), formation of a GNSS Railway Threat Model, and demonstration of a Route Clearance service.

3.3 Expected Impact

In terms of impacts there is a target for the work to generate future business in various ways. Work will lead to an established core of safety expertise within the consortium concerning the use of on-board positioning technology for railway applications and this will be important as the technology grows and improves. There is an ambition to develop a Route Clearance service model for use in safely introducing the technology to specific new lines and for new applications. The partners will also gain a holding in the supply chain(s) involved in developing and certifying dedicated hardware, algorithms and the infrastructure required to deliver the services e.g. local integrity monitoring stations.

1. Improved services and customer quality: the use of enhanced communication technologies will derive in the following outcomes: trains will be managed more efficiently providing improved reliability, enhanced capacity of the lines and providing communication capabilities for real-time data transmission for improved passenger information, offering a better customer experience.
2. Reduced system costs: the design and implementation of the on-board gateway will be based on mass-market technologies and open interfaces, which will enable flexible architectures and application of engineering standards, allowing a correct system design adapted to the requirements of different market segments.
3. Enhanced interoperability: remaining technical “open points” related to electromagnetic compatibility will be closed.
4. Simplified business processes as modular architectures are introduced to divide the validation effort and improve standardization. The interface from the core system to the new technologies to be incorporated to the gateway will be defined as a seamless continuous upgrade possibility.

3.4 Implementation

Regarding the on-board positioning, the priority tasks are:

- to contribute to the test set-up, data collection and analysis according to specifications,
- to assist with further development of those specifications, • to provide input to the safety case,
- to build a DRM prototype.

SmartRaCon expect the contribution to data collection/test campaigns to be part of a wider effort including founder members to analyse and validate positioning system performance, throughout the various stages of the testing outlined in the Scoping Paper. The same may be said regarding progression of the specifications output from the NGTC project. The specific tasks proposed to contribute to the safety case will be led by SmartRaCon but will in certain aspects require input and cooperation with founder members. Other proposed tasks are seen as generally internal to the consortium requiring a low level of interaction with the founder members in order to set requirements and gain feedback on results. This is the case for example with the proposed Route Clearance service. In all cases a close level of cooperation with the founder members is important in order to take into account the specific characteristics of the proposed positioning system solution.

3.5 Integrated Assessment

A technology and impact evaluation is an essential element within the Shift2Rail Joint Undertaking in order to show the effect that this initiative will have on its key target KPIs: to double the capacity and availability and to reduce the costs of the railway system by 50%. The Shift2Rail Project “IMPACT-1” is producing a comprehensive bottom-up KPI model with the aim to show the interdependencies between the technological or procedural developments and the high level KPIs of the railway system. The overall objective is to prove the achievement of the objectives of Shift2Rail by determining to which extend the aims of reducing costs and improving availability and capacity will be reached. In a first step, these interdependencies will be analysed as cause-and-effect chains in order to obtain a qualitative model. Subsequently, the qualitative relations will be replaced by mathematical and logical descriptions. This is necessary in order to apply the model to data of the different market segments like high speed, regional, urban / suburban and freight rail. The analysis of the interdependencies as well as the application of the model is done in close collaboration between industry, infrastructure managers, railway operators and scientific institutions. Thus a KPI model will be generated which covers all aspects of the entire railway system. The presentation will cover the approach that has been chosen to develop the qualitative and quantitative model, share the experience made during this process and show the first results of the impact assessment of the Joint Undertaking Shift2Rail.

4. Impact on the Competitiveness of the European railway systems

The introduction of GNSS in the rail domain is now recognized as a powerful tool for ERTMS deployment, old system renewal. A study performed by Bocconi University for the ERSAT EAV projects shows that GNSS-based ERTMS proves to be especially convenient because of relevant savings in operating expenses: -67% each year compared to the traditional ERTMS [Galileo Services, 2016]. Moreover, the cost/benefit ratio will be maximized if satcoms are integrated in the global system.

This introduction is supported both by the GNSS domain through the GSA that identified rail as a real future market (figure 7), and by the railway sector in order to facilitate ERTMS diffusion, reduce costs and enhance safety on local and regional lines. As shown on figures 7 and 8, if asset management applications are currently driving shipments of GNSS devices, safetyrelevant applications (signalling and train control) based on GNSS will be increasingly developed in the next 15 years.

First application is through the Virtual Balise Concept in ERTMS L2 or 3 but next steps to be prepared are also Moving block operation and automatic train operation.

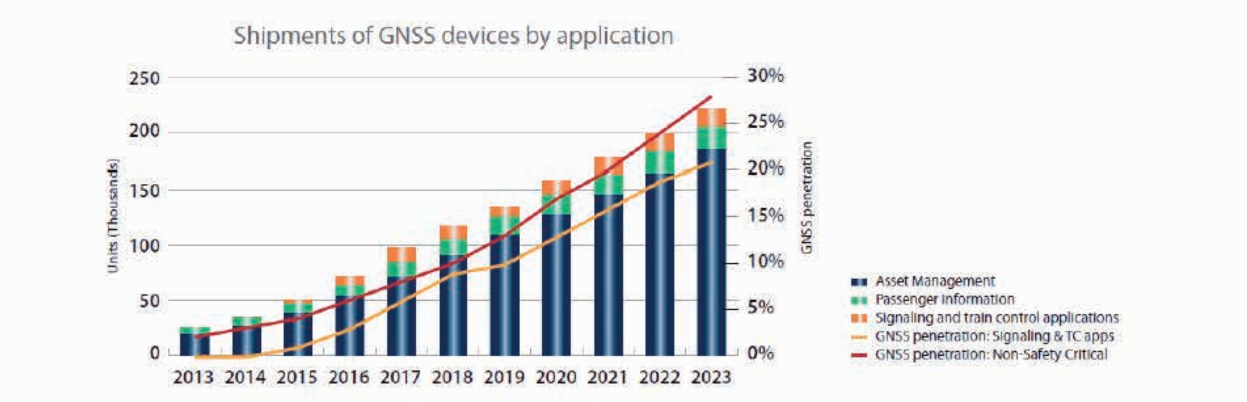


Figure 3. Future market evolution in rail applications

The NGTC project, that gathered together the main industries involved in ETCS and CBTC, consolidated the use of GNSS in the Virtual Balise concept and has developed and validated Generic Moving Block principles applicable to different railway types. Designed as a lighthouse project for Shift2Rail IP2, the outputs of NGTC regarding the analyses of the Satellite Positioning Signal Receiver Parameters that are relevant for the signalling applications and the process of their qualification and validation and, Initial Safety Analyses in regards to use of satellite positioning in ERTMS, will be major inputs for several Shift2Rail projects.

5. Conclusions and Perspective

Safe on-Board localisation is – besides reliable and safe communication – one of the key elements of future train control systems. Four partners from the science and innovation sector have founded the consortium SmartRaCon to perform research and innovation activities in the context of the European Joint Undertaking Shift2Rail. The aim is to develop and demonstrate technologies for safety-critical positioning applications in the railway sector where it is mandatory to guarantee a high safety and reliability the positioning result.

Several Shift2Rail projects will start in the next years to develop the technologies which will be demonstrated to show the results of the innovation. The aim of these projects is to demonstrate innovative onboard localisation technologies for future train control systems.

6. References

- [1] Albanese, A., Marradi, L., Campa, L., & Orsola, B. (2004). The RUNE project: navigation performance of GNSS-based railway user navigation equipment. In Proceedings of NAVITEC.
- [2] Arrizabalaga, S., Mendizabal, J., Pinte, S., Sanchez, J. M., Bauer, J., Themistokleous, M., & Lowe, D. (2014, April). Development of an Advanced Testing System and Smart Train Positioning System for ETCS applications. In Transport Research Arena (TRA) 5th Conference: Transport Solutions from Research to Deployment.
- [3] Beugin, J., Filip, A., Marais, J., & Berbineau, M. (2010). Galileo for railway operations: question about the positioning performances analogy with the RAMS requirements allocated to safety applications. In European Transport Research Review, 2(2), 93-102.
- [4] Facchinetti, C., Ansalone, L., & Tuozzi, A. (2015, October). Trends in GNSS Italian application scenarios in transportation. In Navigation World Congress (IAIN), 2015 International Association of Institutes of (pp. 1-8).
- [5] Slides presented by Galileo Services about rail applications for Galileo, 16 February 2016, Brussels, Belgium: "Space Strategy for Europe" - Views from the European GNSS Downstream Industry
- [6] Gradinariu T., Stadlmann B.; SATLOC – An innovative train control system for low trafficdensity lines; IRSE News; February 2017, Issue 230
- [7] Gurnik P., Stamm B. (2016) STARS presentation, InnoTrans, 22/09/2016, http://www.stars-rail.eu/wp-content/uploads/2016/04/STARS_InnoTrans_final.pdf
- [8] Legrand, C., Beugin, J., Marais, J., Conrard, B., El-Koursi, E. M., & Berbineau, M. (2016). From extended integrity monitoring to the safety evaluation of satellite-based localisation system. Reliability Engineering & System Safety.
- [9] Marais, J., Beugin, J., & Berbineau, M. (2017). A Survey of GNSS-Based Research and Developments for the European Railway Signaling. IEEE Transactions on Intelligent Transportation Systems.

- [10] Meyer zu Hörste et al. (2009). Entwicklung eines Demonstrators für Ortungsaufgaben mit Sicherheitsverantwortung im Schienengüterverkehr (DemoOrt) Final Report oft he Project
- DemoOrt. Reports from the DLR Institut of Transportation Systems, Volume 8. Dmeo
- [11] Neri, A., Filip, A., Rispoli, F., & Vegni, A. M. (2012). An analytical evaluation for hazardous failure rate in a satellite-based train positioning system with reference to the ERTMS train control systems. In 25th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2012) (pp. 2770-2784). The Institute of Navigation.
- [12] Neri A., Palma V., Rispoli F., Pullen S., Zhang S., Lo S., Enge P. (2016) "A Method for Multipath Detection and Mitigation in Railway Control Applications", ION GNSS+ 2016, Portland, OR, U.S.A
- [13] Raymond, G., Marais, J., Berbineau, M. (2004) Innovations Bring Satellite Control within Reach, Railway Gazette International, Déc. 2004, p835-837
- [14] Rispoli F., Castorina M., Neri A., Filip A., Di Mambro G., Senesi F. (2013). Recent Progress in Application of GNSS and Advanced Communications for Railway Signaling. In 23th Radioelektronika conference, April 16-17, Pardubice, Czech Republic.
- [15] UNISIG Subset026. System requirement specification. Issue 2.3.0
- [16] UNISIG Subset 091 (2014). Safety Requirements for the Technical Interoperability of ETCS in Levels 1 & 2. Union industry of signalling, issue 3.3.0, 2014-05-08.
- [17] UNISIG Subset-088 (2014). ETCS Application Levels 1 & 2 - Safety Analysis. Issue 3.5.4, 2014-05-08.

Safety and signalling systems

360.revista de alta velocidad
número 5 - junio - 2018. Pág 159 - 163

The application of the upcoming standard on ATO over ETCS

Costa, Raúl
Villalba, Manuel

CAF Signalling[1]

Abstract:

Interoperable ATO over ETCS is close to become a new standard, seen as key for the future automation of main line Railway operations in Europe.

Valuable experience and feedback is expected from an early implementation project on the Mexico City – Toluca suburban line.

Issues that could affect the deployment of systems according to the standard are identified, together with proposals to mitigate the impact.

Keywords: ATO, ETCS, Standard, Automation, Shift2Rail

---

[1] *Costa, Raúl. Sales & Business Development Director (Spain) CAF Signalling. Email: rcosta@cafsignalling.com*
Villalba, Manuel. Innovation Director CAF Signalling. Email: mvillalba@cafsignalling.com

1. The application of the upcoming standard on ATO over ETCS

1.1 ATO

Automatic Train Operation is widely seen as the next big boost for the productivity and performance of main line railways. Automation has been used by metros for almost half a century, and today the technology is well proven in this environment. Urban transit operators around the world have been reaping important benefits, yet the number of main line applications remains limited.

It is relatively straightforward to provide automation within a closed system, where the trains are all essentially the same, with similar performance characteristics and a handful of standard stopping patterns. But this is clearly not the case where fast and slow passenger and freight trains must share the same tracks, sometimes run by different operators. Some form of interoperable technology is needed, which as yet is not available off the shelf.

When discussing ATO, it is important to distinguish between the four grades of automation (Table I), which require diminishing levels of human involvement up to GoA 4 which is fully unattended.

TABLE I. GRADES OF AUTOMATION

Grade of Automation	Train Operation	Setting train in motion	Stopping Train	Closing Doors	Operation in the event of disruption
GoA1	ATP with Driver	Driver	Driver	Driver	Driver
GoA2	ATO with Driver	Automatic	Automatic	Driver	Driver
GoA3	Driverless	Automatic	Automatic	Train Attendant	Train Attendant
GoA4	Unattended	Automatic	Automatic	Automatic	Automatic

Being developed by the ERTMS Users Group, Unisig and ERA, the proposed architecture for an interoperable ATO over ETCS is initially envisaged for GoA 2, retaining a driver in the cab.

2. ATO over ETCS

In the main line arena, ETCS offers a standardised ATP function as part of the architecture of the European Rail Traffic Management System. This was originally created to achieve signalling interoperability across Europe, but it is being applied more and more outside Europe, thanks in part to the availability of most elements off the shelf from multiple suppliers.

It seems logical for any main line automation project to incorporate ETCS to provide the ATP functions. But ERTMS is also about interoperability, and it will be essential to ensure seamless operations across boundaries if the European rail network is to become more competitive against other transport modes.

To ensure the maximum benefit from automation, similar system approach and interoperability requirements should extend to the ATO. Any compliant train should be able to run safely with the target level of automation as long as it is running over a compliant infrastructure. That means complementing ETCS with interoperable ATO, while minimising the impact on the existing ETCS specifications.

This approach at the level of the wider railway system goes far beyond the traditional projectoriented approach, in order to ensure a smooth migration towards automation.

2.3 From concept to deployment

The AoE concept has been developed by a team drawn from the ERTMS User's Group and the UNISIG suppliers' association as part of a TEN-T project running from 2012 to 2014. This produced a system requirement specification and a number of interface specifications, which were presented to the European Agency for Railways at the end of 2014.

Additional technical work in 2015 included discussions with representatives of the wider European rail sector about integrating AoE into the agency's longer-term strategy for the development of new functions linked to ERTMS.

These discussions took off with the launch of the Shift2Rail technical work in September 2016. The project plan of the IP2 workstream of Shift2Rail includes the consolidation of the specifications, development of prototypes and demonstrations at different levels of integration.

In parallel, UNISIG and the ERTMS Users Group have started the work with the European Agency for Railways for adapting the ETCS specifications to allow integration with the ATO function, in such a way that the impact on the ETCS should be minimised, facilitating a smooth migration of the systems to include the Automation functions. The aim is to have the first GoA2 Shift2Rail demos running by 2018-19.

As part of the IP2 workstream, Shift2Rail will also address higher levels of automation up to GoA 3 and 4. Early implementation projects are also envisaged to provide feedback under real conditions and take the specifications to a higher level of maturity.

2.4 Early implementation: México-Toluca suburban line

While the final specifications are still under discussion, CAF Signalling has had the opportunity to pioneer a commercial application of AoE on the Mexico City – Toluca commuter line now under construction.

ATO system for this line includes both onboard and trackside equipment.

The project began in 2015, allowing the train control system to be designed from the outset using the AoE draft specifications. Grade of automation required is GoA2. Driving between stations will be automatic, but the drivers will be required to close the doors at each stop; they will also be able to operate the trains manually during periods of disruption.

The line will be fitted with ETCS Level 2 following SRS Version 2.3.0d. Although AoE specifications are mainly being made in the context of ETCS Baseline 3, this early implementation is demonstrating that the interoperable concept will also work with Baseline 2. The critical interface specifications follow the draft AoE standards, notably Subset 126 for the interface between ATO onboard and trackside and Subset 130 for that between the ETCS and ATO onboard equipment.

2.5 Early conclusions

Although the Mexico City – Toluca line is not yet operational, the project development work so far confirms that interoperable AoE offers many advantages. Most significantly, it has confirmed that a standard and interoperable approach to GoA2 is feasible, combining ETCS as the ATP function to ensure safety with an ATO overlay based on the draft specifications. This configuration ensures a high level of functional independency between ATO and ETCS. The automation functionality and performance is not limited by ETCS, except for the safety constraints imposed by the ATP. No modifications have been needed to the RBC, and the concept can be implemented using ETCS trackside equipment to either Baseline 2 or 3.

AoE is a powerful concept built on interoperability and standardisation. While it can be applied to a standalone project such as Mexico City – Toluca, the concept itself is oriented to a wider railway network. Having a configuration that allows an ATO-equipped train to run automatically over any equipped section of the network should simplify the deployment and migration process, reducing the level of investment needed and delivering huge benefits to railway operators and passengers.

2.6 AoE: Benefits and Challenges

ATO benefits are widely demonstrated in all the studies performed about the system. Automatic Operation will provide important improvements in operation, cost and user experience, such as:

- Capacity Increase, reducing headways through optimizing driving.
- Higher punctuality by using the line data, schedule data and TMS information from the trackside in order to drive efficiently to achieve the foreseen timing.
- Higher comfort
- Energy efficiency by using the line data and real-time information in order to drive at an optimized speed profile, providing relevant energy savings.
- Decrease of maintenance cost by reducing track and train wear due to the optimized driving.
- To achieve these improvements and take advantage of these benefits, railways sector must face some challenges:
- Migration from an ATP system to an ATP+ATO system will affects to both On-Board and trackside system. This fact demands some degree of coordination between Infrastructure Managers and Train Operators.
- Communications should evolve to allow a higher capacity.
- Even if ATO can be standardized, trains where the system will be implemented were completely different, so integration of On-Board equipment should be study carefully and independently for each case. Depending on the automation grade, ATO system will interact with the train (traction, brakes, etc.) more or less.
- Drivers' function will change. Depending on the automation grade, the tasks to be performed by the driver will change more or less. This change will have human and social impact in drivers. They should to learnt and adapt to their new functions.
- The only chance to take advantage of the important benefits offered by the automation is to achieve a high degree of commitment within the full railway sector. Nowadays the initiatives to develop AoE are supported by the different actors within the sector, but results must be important to keep with it.
- Some actions have been performed in order to ease the achievement of this challenges but other ones should be analysed and developed to meet the goals. Some of them are identified as follows:

Flexible implementation: Specifications and requirements should be defined in order to allow the implementation of ATO system affecting as less as possible the ETCS performance and the underlying signalling system in case of existing ones.

- Clear migration facilities and plans. This is needed to define realistic planning and deployment of such a complex system. Equipment to be provided by the suppliers must also facilitate step-by-step migration strategies.
- Conformance Testing: Standardization of testing seems to be one of the most important topic to be defined in order to meet the real interoperability of ATO system.
- Railway Sector Commitment is needed mostly to facilitate the co-ordination of planning and deployment affecting the trains, the trackside signalling, and the trackside-train communications.

3. References

Villalba, M. (2016) Pioneering ATO over ETCS Level 2, Railway Gazette International.

High Speed Rail Infrastructure

FUNDACIÓN DE LOS
FERROCARRILES
ESPAÑOLES

360.revista de alta velocidad

High Speed Rail Infrastructure

International Congress on High-speed Rail:
Technologies and Long Term Impacts
Ciudad Real (Spain), October, 2017

Escuela de Caminos
Ciudad Real UCLM

Escuela de Ingenieros
de Caminos, Canales y Puertos
de Ciudad Real
UNIVERSIDAD DE CASTILLA-LA MANCHA

High Speed Rail Infrastructure

360.revista de alta velocidad
número 5 - junio - 2018. Pág 167 - 189

KEYNOTE

Alternate Double–Single Track Proposals for Low and Intermediate Demand High Speed and Conventional Lines

Castillo, Enrique

Grande, Zacarías

University of Cantabria[1]

Abstract

The paper introduces the alternate double single track (ADST) as an alternative to double track lines and discusses its main advantages and shortcomings. An optimization program for the design of such lines are presented, including a detailed analysis of the selected objective function and corresponding constraints. In particular, some interesting formulas for deriving the upper bound of the capacity of these lines are derived to illustrate how the efficiency of the line can range from the single to the double track extremes. Some possible international applications of this alternative are identified in which the ADST could be advantageous. Finally, some real examples of applications, which include the cases of Spain, Chile and Ireland, are used to illustrate the proposed methods.

Keywords: Optimization, railway line design, rational investment analysis, Rail capacity, relative time

---

[1] *Castillo, Enrique. University of Cantabria. Email: castie@unican.es*
Grande, Zacarías. University of Cantabria. Email: Zgandrade@gmail.com

1. Introduction and motivation

Though when you think of high-speed lines you immediately tend to associate them with double track lines, not all high-speed lines must necessarily be built in double-track. In fact, the single track multiplies its capacity almost proportionally to the speed of the trains that circulate through it. Therefore, an increase of speed immediately produces an increase in the efficiency of the single track line. This means that single track lines today have important advantages with respect to past single track lines because train speeds have been increased substantially.

When high speed lines are used in a country for the first time, they are usually oriented to link two large populations (case 1). However, other lines may involve small and at most one large population (case 2). In this paper we point out that the above two cases 1 and 2 are completely different and require different solutions, because: (a) the number of users is necessarily much lower in the second than in the first case leading to very different train frequencies, which are determined mainly by the small size city, and (b) we must question ourselves whether or not the expensive double track is necessary in Case 2, or a new alternative should be contemplated. In addition, since decisions with optimal criteria are required, computer programs are unavoidable to design, develop preliminary projects and evaluate the proposed solutions.

It appears that the first known application of the mathematical programming methodology applied to railway optimization problems was due to (Amit & Goldfarb, 1971) Optimization and simulation were already the most commonly used methods, even before the eighties, when computers had no power enough to deal with these complex problems (see (Assad, 1980) and (Haghani, 1987)). However, nowadays problems related to railway network management and similar cannot be conceived of without computers (see (Petersen, et al., 1986), (Hellström, 1998), (Yang & Hayashi, 2002) or (Ouyang, et al., 2009)). An exhaustive review of existing optimization methods was done by (Cordeau, et al., 1998).

One of the main reasons for these railway optimization problems to be complicated is that a huge amount of continuous and binary variables and constraints are involved, leading to mixed integer (MIP) linear and non-linear related programming problems of high complexity (see, for example, (Kraay & Harker, 1995) (Carey & Lockwood, 1995) , (Higgins, et al., 1996), (DÁriano & Pranzo, 2004), (DÁriano, et al., 2007)), which require a lot of memory and CPU resources (see (Burdett & Kozan, 2010)). As indicated in (Castillo, et al., 2011), (Castillo, et al., 2015), a reduction in binary variables has an important effect on the CPU time required.

A possible alternative to single track and double track lines is the alternate double-single track (ADST) line, in which single track segments are combined with double track segments. As will be verified, the capacity of a single track line is clearly overcome by an ADST line with a very low construction and maintenance costs increase. In this paper, first, a general formula for dimensioning the lengths of single track and double track section is derived. Then, these formulas are applied to several cases that will serve us to illustrate the possibilities that the ADST lines provide from a practical point of view.

The design of a railway line must always include the predicted demand. If the demand is very high a double track is the right solution. Contrary, if the demand is very low, the right solution is the single track. However, in the most common cases of intermediate demands the optimal solution involves segments in single and segments in double track. Since in this intermediate case we must decide which segments must be in single and which in double tracks, the demand plays a relevant role in the optimal solution because the timetable must be optimized in order to reduce travel time. This implies that the costs of single and double tracks must be included in the objective function of our optimization problem. In fact, we are in front of a bi-level problem in which the first level decides about single or double tracks and the second level optimizes the timetables to reduce travel times and reduce the occupancy of the line.

The second level problem of selecting the optimal timetables, that is, to determine time departures and arrivals of trains from and at stations so that not only safety constraints but user requirements are satisfied is a very difficult one (see, for example, (Carey, 1994), (Higgins, et al., 1996), (Caprara, et al., 2002) or (Cacchiani & Toth, 2012)). There exist a wide collection of related works, such as (Jia & Zhang, 1993), (Sahin, 1999), (Carey & Crawford, 2007) and (DÁriano, et al., 2007), and a complete analysis done by (Pachl, 2014) of the timetable design principles that must rule engineering work in this area.

An alternative to the bi-level approach is to consider all into a single problem 3. For example, (Castillo, et al., 2009) combines into a single optimization problem the line and timetable designs, producing an important change in the alternate double-single track direction, as an intermediate and efficient solution between the traditional double and single track solutions.

No matter which alternative is chosen, the size of de associated problems can reach a point in which serious complexity problems appear. To solve this problem, (Lin & Ku, 2013) proposed genetic algorithms and (Castillo, et al., 2016) proposed a time partitioning technique that reduces drastically the computation time without serious losses in optimality.

2. The alternate double-single track

An alternate double-single track (ADST) line essentially consists in utilizing single track throughout expensive segments (tunnels and viaducts) and double track in cheap segments (plain areas) and only where it is necessary.

It should be clearly stated that (see (Castillo, et al., 2015)): (a) the ADST line is not a single track line, (b) the ADST line is not a double track line, (c) the ADST performance is much closer to double than to single track performance, (d) the ADST cost is much closer to single track than to double track cost, (e) it reaches practically the same performance as the double track solution for the expected demand and even slightly superior to it, (f) it reduces the construction and maintenance costs (until a 40%) and finally, (g) lines, which are nor economically viable as double track lines can become viable as ADST lines.

The design and management of an alternate double single track line is complex, because it requires:

1. Deciding which segments should be constructed in single track and which in double track.
2. Satisfy the safety and timetable constraints of the different services with the aim of obtaining small travel times when we have a single track in some segments.
3. Minimize costs and travel times and optimize the infrastructure usage.
4. Obtain all rail timetables of the whole network

Due to the complexity of the problem, the use of an optimization program is necessary in order to satisfy all the imposed safety and service conditions.

2.1 Some considerations about travel times

Since travel times of the different trains circulating along the network or line could be very different and it is not the same a delay of five minutes in a one hour trip than in a three hours trip, we use in our model relative travel times.

The relative travel time is the quotient:

$$Relative\ travel\ time = \frac{Travel\ time}{Travel\ time\ at\ maximum\ speed} \qquad (1)$$

Then, a relative time *1* means that we travel at maximum speed; contrary a relative time value of *1.10* or *1.20* means that we have been used for the trip a *10%* and a *20%* more time, respectively.

This has the advantage of allowing the combination of different trains correcting the travel duration effect.

3. Elements of the optimization problem. Proposed model

This section is devoted to describe the used optimization problem (see (Castillo, et al., 2015)). In any optimization problem we must deal with four basic elements: data, variables, constraints and objective function. We analyze them in detail below.

3.1 Data

Our data will be: (a) the set of candidate segments to determine which ones should be built as single or double tracks, (b) the segment construction costs of each segment as single and double track, (c) the demand, that is, the number and priority of daily services, (d) the desired departure times and corresponding flexibilities of all services, (e) the maximum segment speeds, (f) the safe headway times between consecutive trains, and the service dwell times.

3.2 Variables

The list of variables used in our model is given in Table 1.

Table 1 List of variables used in our model.

Variable	Meaning
b_j	Binary variable which value is 0 for single tracks and 1 for double
$e_{i,j}$	Entry time of train *i* in segment *j*
$s_{i,j}$	Exit time of train *i* from segment *j*
$t_{i,j}$	Travel time of train *i* along segment *j*
ε	Maximum relative travel time of all trains (\varepsilon\ge 1)
$q_{i,j,t}$	Binary variable which value is 1 if train *i* uses track t in segment j
r_i	Departure time of train *i*
$x_{i_1,i_1,j}$	Binary variable which value is 1 if the train i_1 uses segment *j* before the train i_2
$y_{i_1,i_1,j}$	Binary variable which value is 1 if the train i_1 uses segment *j* after the train i_2.
η_{il} , η_{iu}	slacknesses in the real departure time of train *i* (before and after desired times).
t_i^0	Minimum travel time of train *i*
γ_i	Priority of train *i*
r_{0i}	Desired departure time of train *i*
r_i^l , r_i^u	Maximum allowable slacknesses in departure time of train i (before and after desired times)
t_i^0	Minimum travel time of train *i*
γ_i	Priority of train *i* $(0< \gamma i \le 1)$
r_{0i}	Desired departure time of train *i*
r_i^l , r_i^u	Maximum allowable slacknesses in departure time of train i (before and after desired times)

3.3 Constraints

In this section we provide most of the constraints used in our model.

1. Departure times equal to entry times in the first segment.

$$s_{i,1} = r_i \qquad \forall i \in I \tag{2}$$

2. Departure times are upper and lowerly bounded.

$$r_{0i} - r^i_{\ell} \le r_i \le r_{0i} + r^i_u \qquad \forall i \in I \tag{3}$$

3. Exit time of a segment equal to entry time plus travel time through the segment.

$$e_{i,k} = s_{i,k} + t_{i,k} \qquad \forall i \in I \qquad \forall k \in K_i \tag{4}$$

4. Bounded segments' travel times.

$$t^i_{k,\ell} \le t_{i,k} \le t^i_{k,u} \qquad \forall i \in I \quad \forall k \in K_i \tag{5}$$

5. Lower bounded dwell times.

$$s_{i,k} \ge e_{i,k-1} + d^{\ell}_{i,k} \qquad \forall i \in I \quad \forall k \in K_i \tag{6}$$

6. Null stop at no TOPP[1] nodes.

$$s_{i,k} = e_{i,k-1} \qquad \forall i \in I \quad \forall k \notin PAET \tag{7}$$

7. Enforce headway times among trains circulating in opposite direction.

$$s_{i1,j} - e_{i2,j} \ge h_j^{\,0} y_{i1,i2,j} - M\,(1 - y_{i1,i2,j}) \tag{8}$$

$$s_{i2,j} - ei1,j \ge h_j^{\,0} xi1,i2,j - M(1 -_{xi1,i2,j})\ \forall_{i1,i2} \in I\ \forall j \tag{9}$$

8. Enforce headway times among trains circulating in same direction.

$$s_{i2,j} - s_{i1,j} \ge h_{i1,i2,j} x_{i1,i2,j} - M(1 - x_{i1,i2,j});\ \ \forall i_1, i_2, j \tag{10}$$

$$s_{i1,j} - s_{i2,j} \ge h_{i1,i2,j} y_{i1,i2,j} - M\,(1 - y_{i1,i2,j});\quad \forall i_1, i_2, j \tag{11}$$

$$e_{i2,j} - e_{i1,j} \ge h_j^{\,0} y_{i1,i2,j} - M(1 - x_{i1,i2,j});\quad \forall i_1, i_2, j \tag{12}$$

$$e_{i1,j} - e_{i2,j} \ge h_j^{\,0} y_{i1,i2,j} - M\,(1 - y_{i1,i2,j});\qquad \forall i_1 < i_2 \in I\ \rho_{i1,i2,j} = 1 \tag{13}$$

9. If two trains use the same track in the same segment \$j\$, at least one of them should have priority

$$b_j + q_{i1j2} + q_{i1j2} - 2 \le x_{i1,i2,j} + y_{i1,i2,j}\ \forall_{i1,i2} \in I\ \forall j \tag{14}$$

$$q_{i1,j,1} + q_{i2,j,1} - 1 \le x_{i1,i2,j} + y_{i1,i2,j,} \forall i1 < i_2 \in I, \forall j, t, \tag{15}$$

10. At most a train has priority

$$x_{\{i1,i2,j\}} + y_{\{i1,i2,j\}} \le 1; \forall i_1 < i_2 \in I \qquad \forall j \tag{16}$$

---

[1] Train Overtop and Parking Post

4. Objective function

Though we can use many other criteria, the main one used here is to minimize the maximum relative travel time of all circulating trains.

Since there are infinitely many solutions with this minimum relative travel time, we can use other criteria to select one among them, such as minimizing the sum of all relative travel times of circulating trains.

Since we still can have infinitely many optimal solutions we can add a third criterion, such as minimum fuel consumption, which leads to strict dwell time in stations and speeds as small as possible.

Consequently, we clearly face a multiobjective hierarchical optimization problem.

The selected objective function for our model includes five hierarchized goals: (a) construction cost, (b) maximum relative travel time ε , (c) relative travel times sum, (d) fuel save and (e) sum of departure time slacknesses.

Hence, our objective function can be chosen as:

$$\underset{b,s,e,r,t,x,y,\eta}{Min} \quad Z=c(\boldsymbol{b})+\alpha_1\varepsilon+\alpha_2 f(\boldsymbol{e},\boldsymbol{s})-\alpha_3 g(\boldsymbol{e},\boldsymbol{s})+\alpha_4 p(\boldsymbol{e},\boldsymbol{s})$$

Where

$$c(\boldsymbol{b}) = \sum_j (1-b_j)\,STC_j + b_j DTC_j \quad : \quad Construction\ Cost \qquad (18)$$

$$\varepsilon = \max_{i \in I}\left[y_i \frac{s_{i,last}-e_{i,1}}{t_i^0}\right] \quad : \quad Max\ relative\ time \qquad (19)$$

$$f(\boldsymbol{e},\boldsymbol{s}) = \sum_{i \in I} y_i \frac{e_{i,last}-s_{i,1}}{t_i^0} \quad : \quad Relative\ times\ sum \qquad (20)$$

$$g(\boldsymbol{e},\boldsymbol{s}) = \sum_{i \in I} (e_{i,k}-s_{i,k-1}) \quad : \quad Energy\ consumption \qquad (21)$$

$$p(\boldsymbol{e},\boldsymbol{s}) = \sum_{i \in I} |r_{0i}-r_i| \quad : \quad Sum\ of\ departure\ time\ slacknesses \qquad (22)$$

5. ADST Capacity

In this section we derive some interesting formulas for dimensioning the lengths of double and single track segments. They are based on the hypothetical case of assuming a maximum capacity when trains circulate in both directions at the same speed s and a headway of duration t is used to guarantee safety. Though this assumption is an ideal situation, it gives an upper bound of the capacity of the line and provides us with a good understanding of ADST lines.

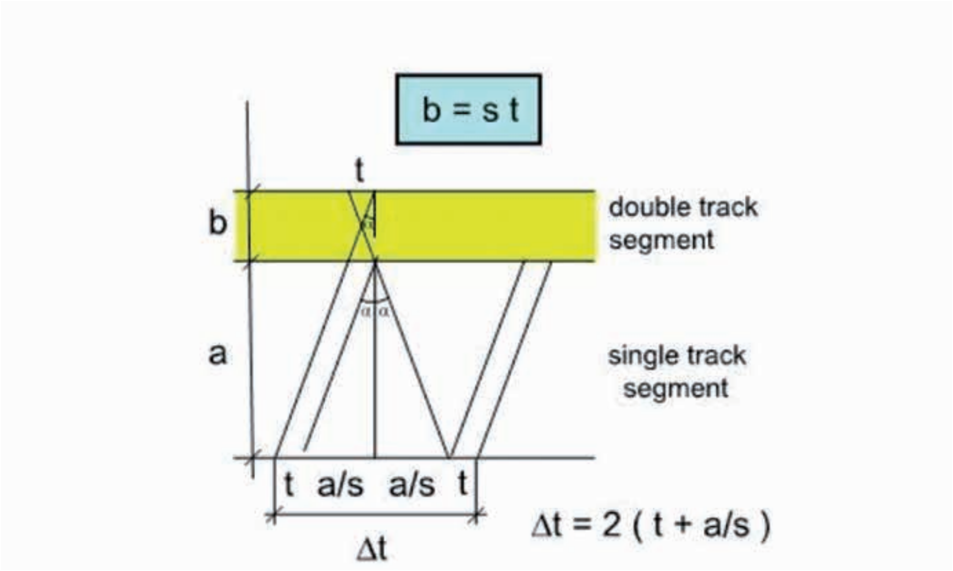


Figure 1. Circulating diagram of one single track segment (the bottom one) and one double track segment (the top one).

Figure 1 represents a diagram showing one single track segment (the one on the bottom) and one double track segment (the one on the top) and three trains circulating at speed $\mathrm{s}=\mathrm{ctg}\alpha$, together with the headway t and lengths a and b of the segments. It represents the optimal situation in terms of efficiency of the segments, that is, the travel times of the segments is a minimum and corresponds to a minimal occupation of the line.

According to figure Figure 1, the train time lag Δt, that is, the time between two consecutive trains circulating in the same direction, satisfies the equation

$$\Delta t=2\ (t+a/s) \tag{23}$$

Similarly, from the double track segment in the top of figureFigure 1 we obtain

$$b=st \tag{24}$$

where b is the double track segment length.

This formula is very useful to design the optimal double track length in terms of s and t, which are usually selected without any problem.

From expressions (23) and (24) we have

$$\Delta t/t=2\ [1+a/st]=2\ [1+a/b], \tag{25}$$

which implies

$$\mathrm{b}/a+b = 2t/\Delta t. \tag{26}$$

This equation provides the proportion of double track length with respect to the total length in terms of t and Δt.

It is relevant to note that the two equalities in (25) involve dimensionless ratios $\Delta t/t,\ a/st$ and a/b, as suggested by the well-known $\Pi-Buckingham\ theorem$. This means that these two equations depend only on two variables (ratios) and not on four variables, $\Delta t,\ t$ a and s the first and $\Delta t,\ t$ a and b, the second, as it could appear in an initial or superficial analysis of them.

Figure 2 shows different train flow diagrams illustrating how the theoretical capacity of a line improves from 34 services per day to 264 services per day when we increase the number of double track segments from zero (single track line) to a maximum value (double track line).

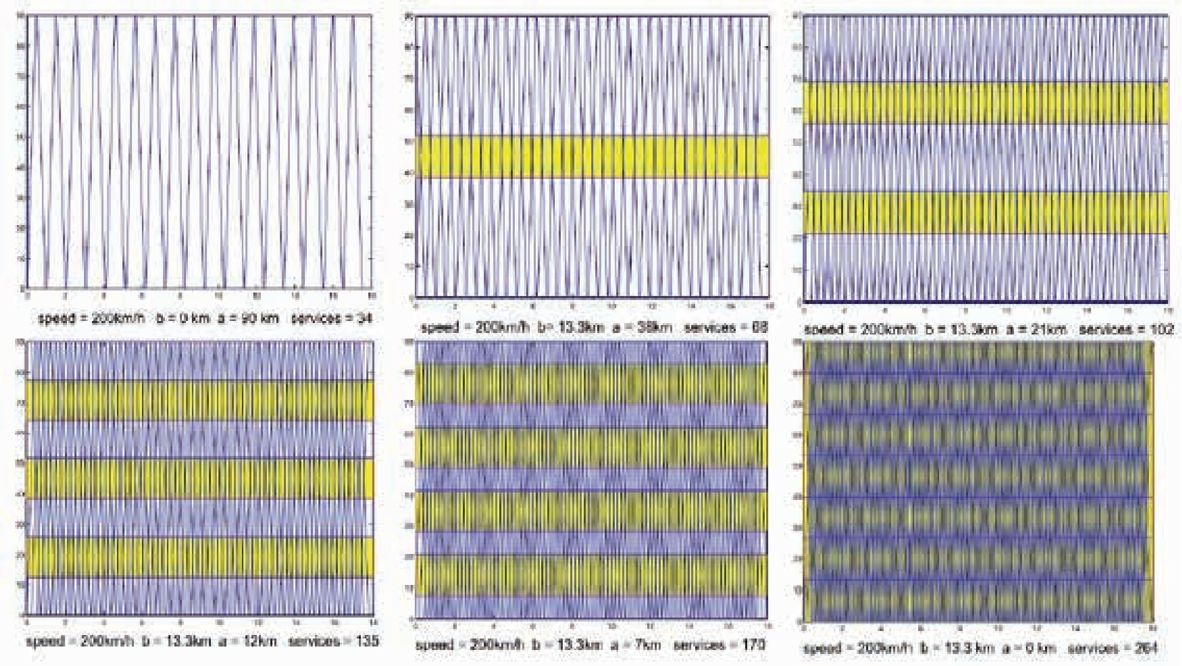


Figure 2. Train flow diagrams illustrating how the theoretical capacity of a line improves from 34 services per day to 264 services per day when we increase the number of double track segments from zero (single track line) to a maximum value (double track line).

6. Software

To solve the optimization problem (17) subject to constraints (like (2) to (16) we have written a computer program in Matlab, which provides the optimal solution. This means that the optimal sequence of single and double track segments is obtained together with the optimal timetable of the line.

In addition, the GAMS program automatically generates a Mathematica code, which after execution produces plots as those in figures \ref{f522} and \ref{f522a}, and a Latex code such that when compiled produces tables with the arrival and departure times for all stations and services (trains).

Finally, a user interface, developed in JAVA, permits a friendly data input.

7. Internationalization

There are many locations where ADST could be implemented. Some examples are given in figures Figure 3 and Figure 4, which include Thailand, Ireland, Ukraine, Morocco, Italy, Spain, Turkey, France and the United Kingdom.

KEYNOTE

Figure 3 Some possible ADST in Thailand, Ireland and Ukraine

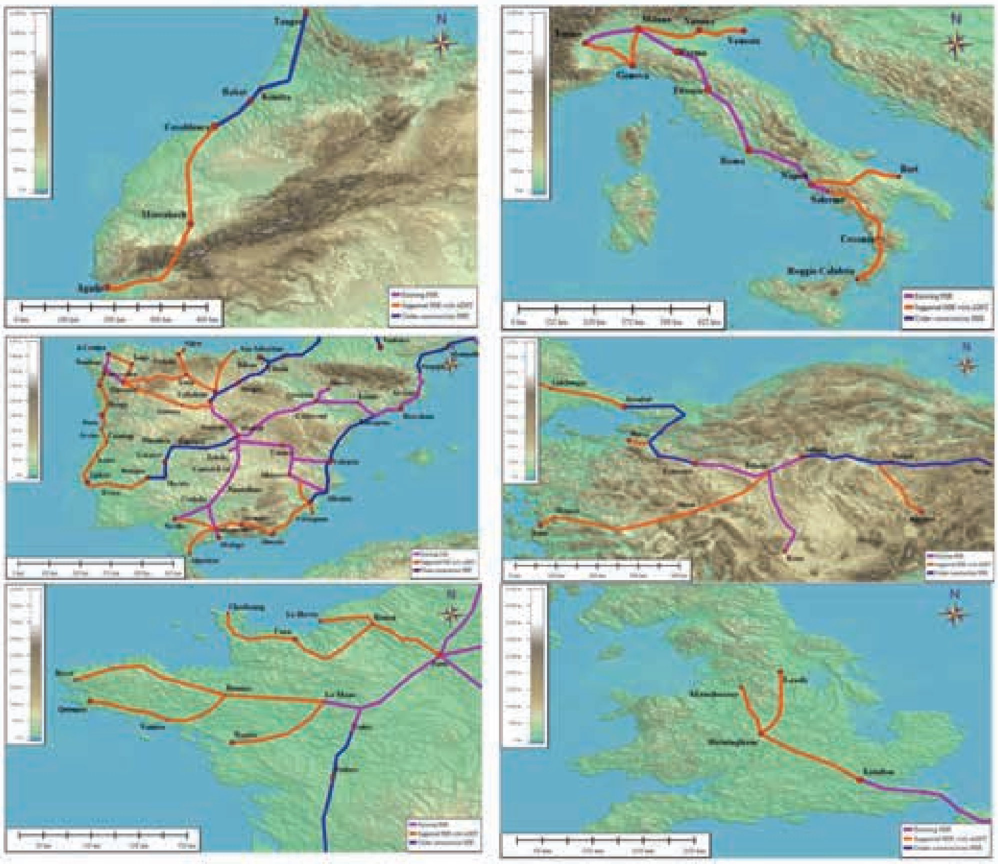


Figure 4 Some possible ADST in Morocco, Italy, Spain, Turkey, France and United Kingdom.

8. Some examples

8.1 Santiago - Valparaíso and Viña del Mar Case

We start with the simplest case, that is, when we built a new railway line. This means that we are free to choose single and double track segments.

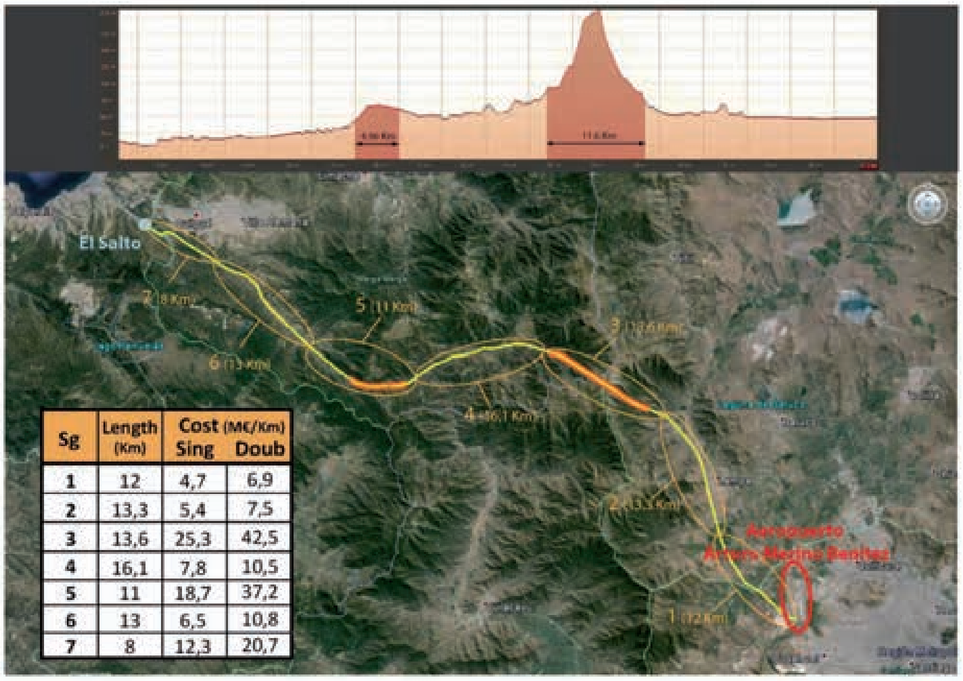
This is the case of Santiago - Valparaíso and Viña del Mar, in which among others they have the following problems:

1. 15,000 people travel for work on weekdays by bus or car from Valparaíso and Viña del Mar to Santiago and vice versa and they expend between 4 and 5 hours in transportation.
2. Valparaíso port is losing competitiveness with respect other ports in the country because it does not have transport by rail.
3. The beaches and the casino of Viña del Mar have difficult access by road from Santiago, which produces significant losses in tourist activities.
4. Citizens from Valparaíso and Viña del Mar have difficult access to the Santiago international airport.

With the idea of solving these and generating new activities, we proposed to the community to construct a new high speed railway line, based on the following ideas.

- A mixed line for passenger and freight trains was considered.
- The maximum service speeds have been 200 km/h for passenger trains and 120 km/h for freight trains.
- Two paths: Santiago-El Salto and vice versa have been studied.
- This implies a total of 100 daily services: 76 passenger trains (38 each way) and 24 freight trains (12 in each direction), 1 train every 20 minutes at peak (6:00-8:20 and 17:20-20:20), and every 30 minutes during valley hours (8:20-17:20 and 20:20-22:20).
- Five night hours have been reserved for line maintenance.

Figure 5 shows the proposed layout of the Santiago-Valparaíso line, in which only two short tunnels are required. It also shows the seven elements used for our analysis together with their lengths and construction costs as single and double track.



Sg	Length (Km)	Cost (M€/Km)	
		Sing	Doub
1	12	4,7	6,9
2	13,3	5,4	7,5
3	13,6	25,3	42,5
4	16,1	7,8	10,5
5	11	18,7	37,2
6	13	6,5	10,8
7	8	12,3	20,7

Figure 5. Layout of the Santiago-Valparaíso line showing the seven elements considered.

KEYNOTE

Introducing these data in our computer program together with the desired departure times and their corresponding flexibilities we obtained the solution indicated in figure \ref{f502}, where the double track segments are shown in yellow color.

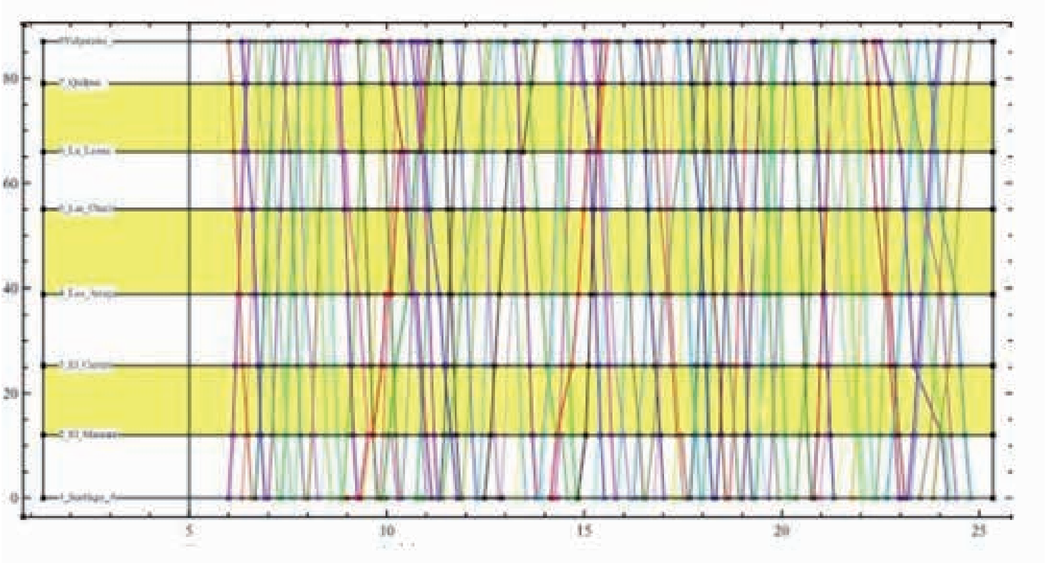


Figure 6. Proposed timetable for the Santiago - Valparaíso/Viña del Mar line

From the study above, we can conclude the following:

1. The proposed line Santiago - Valparaíso/Viña del Mar can be integrated with the metropolitan transport network (metro, buses, etc.).
2. The alternate double-single track solution (1,084 M€) allows us to achieve a saving of *36%* compared to the double track (1,700 M€) and important maintenance savings.
3. If the line were integrated with the subway of Valparaíso (same gauge), freight trains could reach Valparaíso port

This ADST project permits:

- Daily displacement of 15,000 workers from Valparaíso and Viña del Mar to Santiago, and vice versa.
- Promote tourism and leisure activities to the citizens of Santiago (beaches, casino, etc.).
- Provide Valparaíso and Viña del Mar with an international airport accessible in half an hour.
- Allow the survival of the port of Valparaíso giving it an exit by rail to Santiago.
- Generate new activities and population centers in Santiago, Valparaíso and Viña del Mar.

8.2 Palencia-Santander case

This example deals with the improvement of the Palencia-Santander line, with a length of 217.2 km. connecting Palencia (80.000 dwellers) with Santander (176.000 dwellers).

Figure 7. Layout of the Santander-Palencia line.

In this case, we have an existing conventional line, whose layout is shown in figure Figure 7 and we want to improve it to reach Madrid from Santander in 3 hours, which would allow railways to compete with air transport. This means, a target travel time of 1 hour and 45 minutes from Santander to Palencia (the actual travel time is more than 2 hours and 25 minutes).

Several years ago a high speed double track line was designed with a length of 186.6 km and a travel time of 1 hour 40 min. with a cost of 3,200 M€. Unfortunately, due to the high cost, it was neither financed nor built.

The profile of this Palencia-Santander line is shown in figure Figure 8 where we can be seen that de first 120 km are rather flat, the intermediate segments correspond to a very steep line with alternating tunnels (in red color) and viaducts (in blue color) and the final 23.7 kms are rather flat again.

It is clear that the construction of the intermediate segments is very expensive and is not justified by the actual or expected future demand, so that we expect that our optimization program will decide single track segments in this zone.

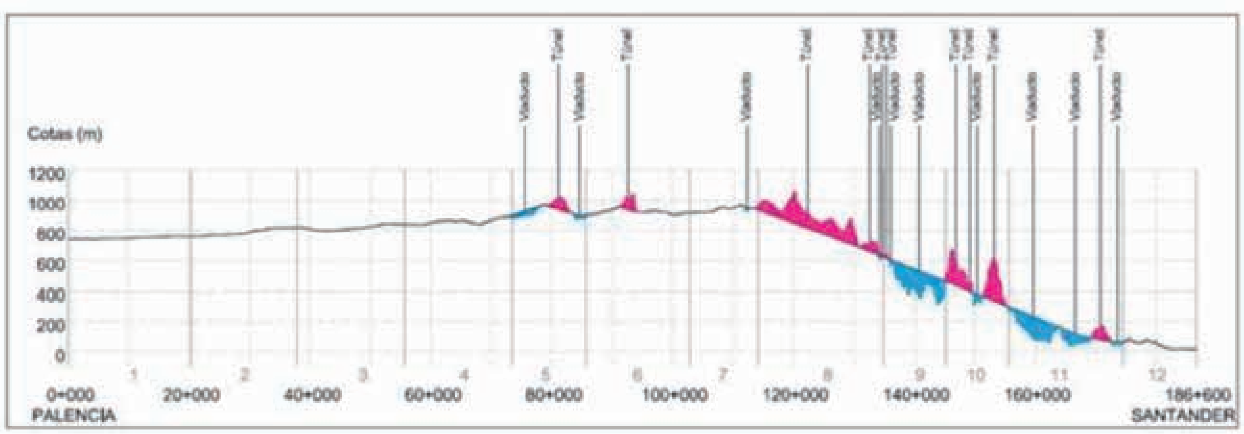


Figure 8. Profile of the Santander-Palencia line.

KEYNOTE

Figure 9 shows the Palencia-Santander trace with the 12 segments that have been considered in our analysis, from which the optimization program will select those to be built in double track.

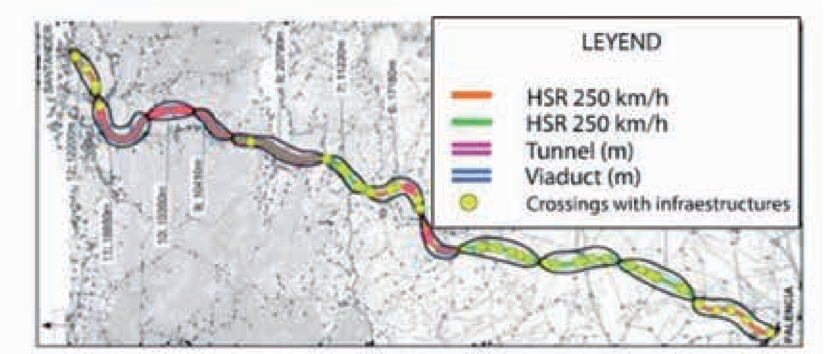
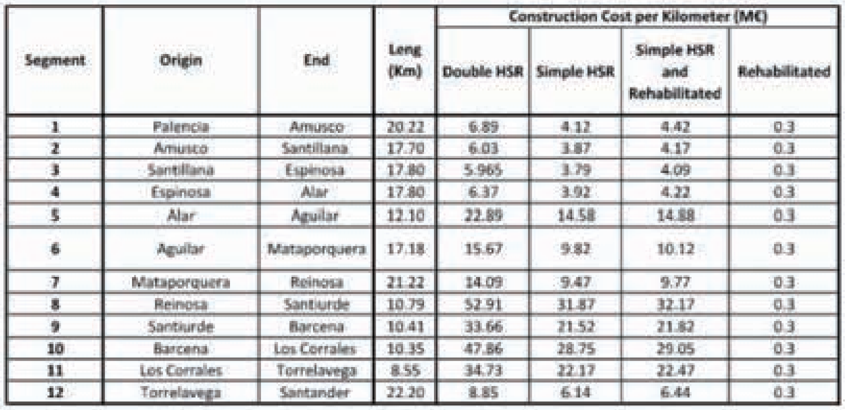


Figure 9. Description of the Palencia-Santander line, showing the 12 segments that have been considered in our analysis.

Table 2 shows the selected segments and cost per kilometer for the different alternatives (single, double track, etc.)

Table 2. Selected segments and cost per kilometer from the different alternatives.



Segment	Origin	End	Leng (Km)	Construction Cost per Kilometer (M€)			
				Double HSR	Simple HSR	Simple HSR and Rehabilitated	Rehabilitated
1	Palencia	Amusco	20.22	6.89	4.12	4.42	0.3
2	Amusco	Santillana	17.70	6.03	3.87	4.17	0.3
3	Santillana	Espinosa	17.80	5.965	3.79	4.09	0.3
4	Espinosa	Alar	17.80	6.37	3.92	4.22	0.3
5	Alar	Aguilar	12.10	22.89	14.58	14.88	0.3
6	Aguilar	Mataporquera	17.18	15.67	9.82	10.12	0.3
7	Mataporquera	Reinosa	21.22	14.09	9.47	9.77	0.3
8	Reinosa	Santiurde	10.79	52.91	31.87	32.17	0.3
9	Santiurde	Barcena	10.41	33.66	21.52	21.82	0.3
10	Barcena	Los Corrales	10.35	47.86	28.75	29.05	0.3
11	Los Corrales	Torrelavega	8.55	34.73	22.17	22.47	0.3
12	Torrelavega	Santander	22.20	8.85	6.14	6.44	0.3

The line is constructed to satisfy a certain demand. The actual demand is shown in figure Figure 10, which show the actual demand optimized circulation diagram. It consists of 70 daily trains.

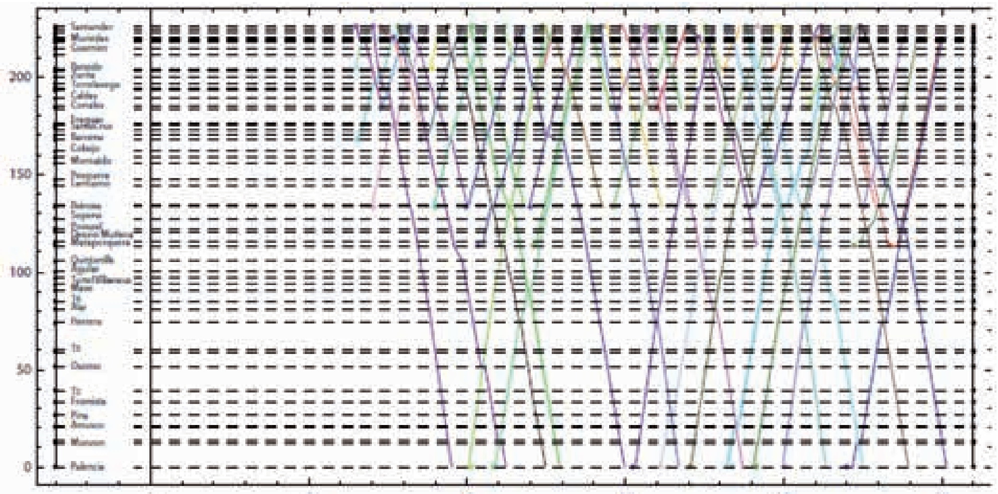
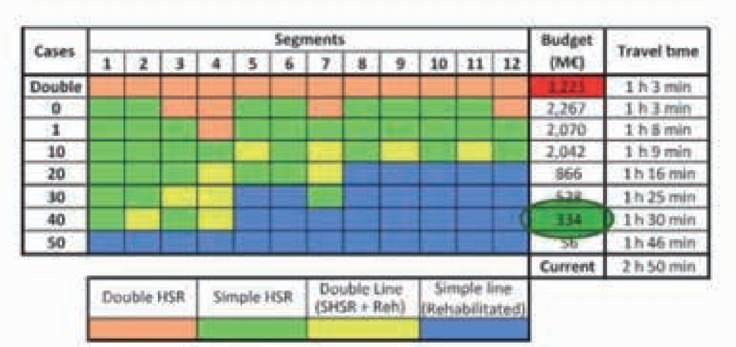


Figure 10. Current optimized circulation diagram: Santander-Palencia



Cases	Segments												Budget (M€)	Travel time
	1	2	3	4	5	6	7	8	9	10	11	12		
Double	🟧	🟧	🟧	🟧	🟧	🟧	🟧	🟧	🟧	🟧	🟧	🟧	3,221	1 h 3 min
0	🟩	🟩	🟧	🟧	🟩	🟩	🟧	🟩	🟩	🟩	🟩	🟧	2,267	1 h 3 min
1	🟩	🟩	🟩	🟧	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	2,070	1 h 8 min
10	🟩	🟩	🟩	🟩	🟨	🟩	🟨	🟩	🟨	🟩	🟨	🟩	2,042	1 h 9 min
20	🟩	🟩	🟩	🟨	🟩	🟩	🟨	🟦	🟦	🟦	🟦	🟦	866	1 h 16 min
30	🟩	🟩	🟨	🟨	🟦	🟦	🟩	🟦	🟦	🟦	🟦	🟦	528	1 h 25 min
40	🟩	🟨	🟩	🟨	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	334	1 h 30 min
50	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	56	1 h 46 min
													Current	2 h 50 min

Double HSR	Simple HSR	Double Line (SHSR + Reh)	Simple line (Rehabilitated)
🟧	🟩	🟨	🟦

Figure 11. Chart showing the different solutions an associated budget and travel times for the Santander-Palencia line

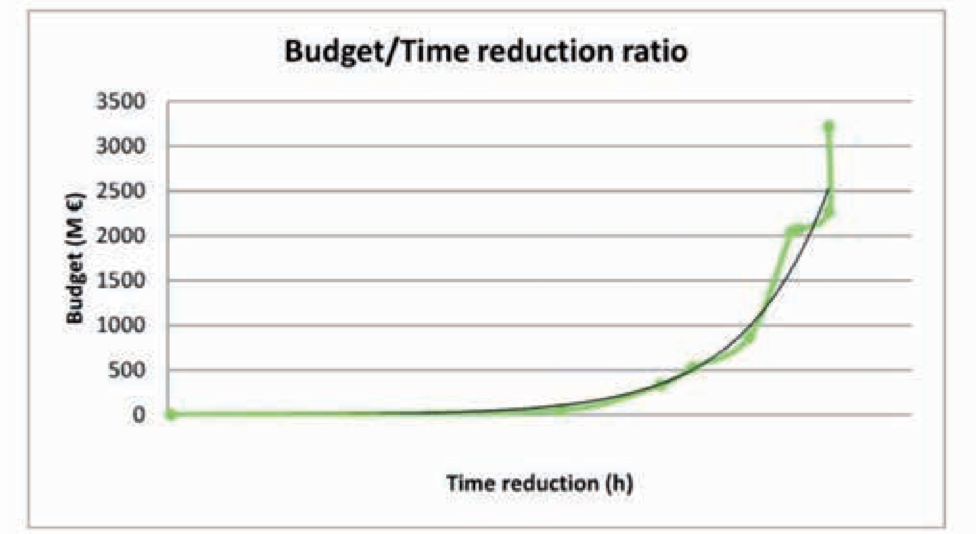
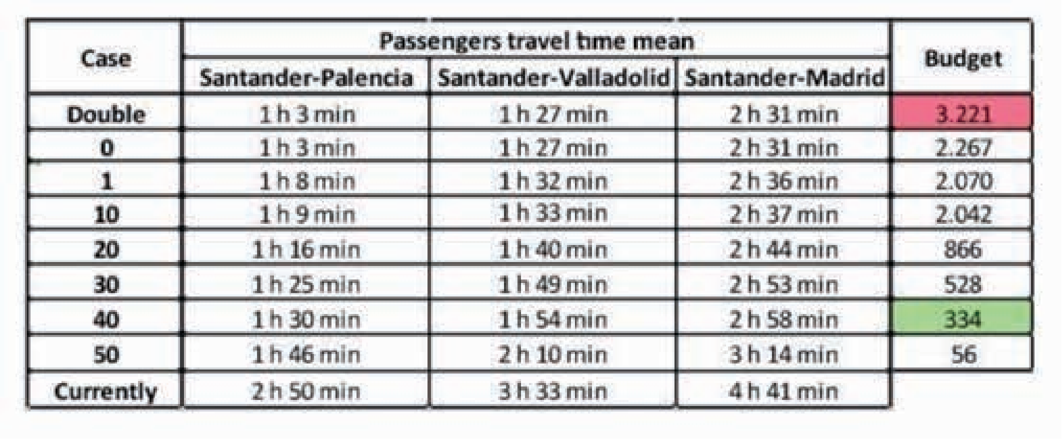


Figure 12. Cost versus reduction in travel time (in hours).

Table 3 Travel times and costs of the different analyzed solutions.



Case	Passengers travel time mean			Budget
	Santander-Palencia	Santander-Valladolid	Santander-Madrid	
Double	1 h 3 min	1 h 27 min	2 h 31 min	3.221
0	1 h 3 min	1 h 27 min	2 h 31 min	2.267
1	1 h 8 min	1 h 32 min	2 h 36 min	2.070
10	1 h 9 min	1 h 33 min	2 h 37 min	2.042
20	1 h 16 min	1 h 40 min	2 h 44 min	866
30	1 h 25 min	1 h 49 min	2 h 53 min	528
40	1 h 30 min	1 h 54 min	2 h 58 min	334
50	1 h 46 min	2 h 10 min	3 h 14 min	56
Currently	2 h 50 min	3 h 33 min	4 h 41 min	

KEYNOTE

Once the Cantabria Government had knowledge of the results of the previous study changed his mind and decides to abandon the HSR double track solution of 3221 M€ and to commission a study of the solution whose budget was 334 M€.

The optimal timetable results as in figure Figure 13 where the green segments are double track segments.

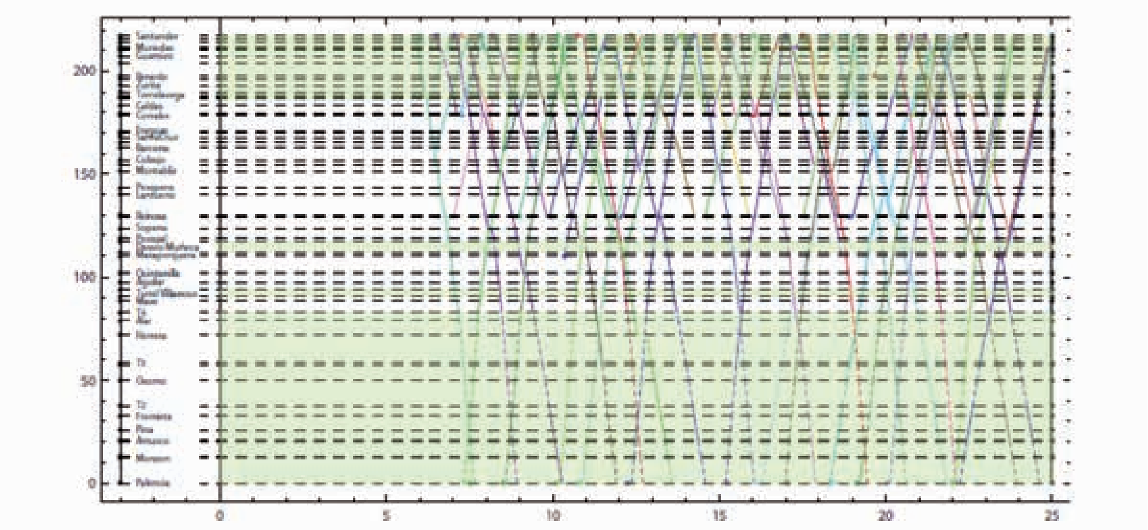


Figure 13. Example of circulation diagram with 8 additional long trip trains in the Santander-Palencia line.

Consequently, the final proposed actions to improve the Cantabrian railway service are depicted in Figure 14.

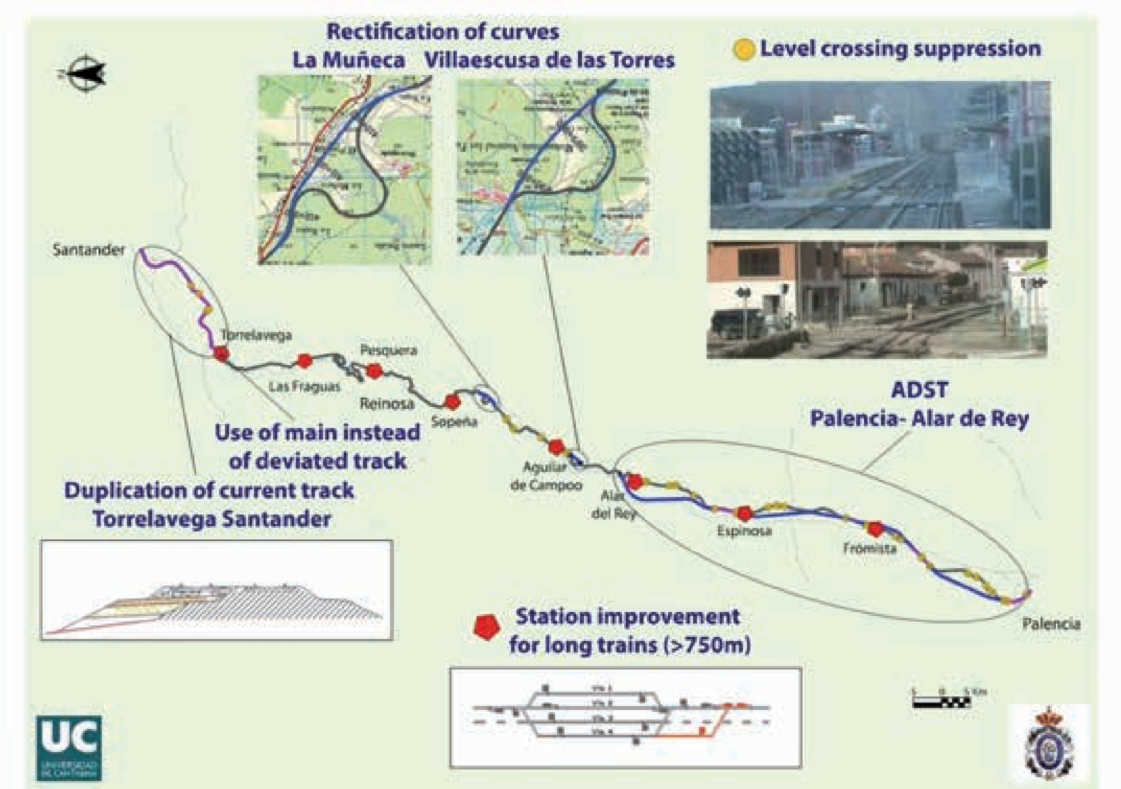


Figure 14. Set of actions.

8.3 Dublin Belfast case

In this section we analyze the existing conventional Dublin-Belfast line and analyze several alternatives for a new high speed line, showing the associated costs and travel times.

8.3.1 Current Line

Before designing a railway line it is very important to study the affected population and the possible demand. The list of the connected cities and their actual populations are shown in figure Figure 15 which shows that the Dublin-Belfast link could connect 2 million people.

City	Inhabitant
Dublin	1.273.069
Dorgheda	38.578
Dundalk	37.816
Newry	29.946
Portadown	22.000
Belfast	720.000

Type	Dublin - Belfast		Services	Trip fares (€)	
	CTC Travel Time	Ocuppancy	per day	Single	Return
Car	2 h 2 min	40%	-	11	23
Bus (705X)	2 h 33 min	19%	44	17	24
Train	2 h 47 min	40%	16	38	55

*CTC Center to Center

Figure 15. Current line description.

The trace of the current line between Dublin and Belfast together with their main stations are shown in figure Figure 15. Its length is 181 Km (113 miles) and the current travel time is 2 h and 9 min. It must be noted that, in addition to offering 16 daily services between the two cities, the network shares 312 daily services including the Dublin and Belfast commuter and freight transport services.

In addition, Figure 15 shows the travel times between both cities, occupancies and travel costs associated with the different alternatives such as, car, bus and other trains, together with their single and return trip fares. It can be seen that the railway transport is the most expensive conveyance and its daily service is significantly lower than the bus connection.

In the following sections we propose some alternatives to improve the efficiencies of the DublinBelfast services.

Figure 16. High speed line proposal showing the selected eight segments used for the analysis and the considered speeds in each segment.

8.3.2 High speed line proposal

As our first proposal we consider a new double track high speed line from Dublin to Belfast.

We propose the line shown in figure Figure 16 where the traces of the new and the existing lines are shown and the selected eight segments used for the analysis and the considered speeds in each segment. The new proposed line has a length of 160.4 Km (100.25 miles), that is, a reduction of 20 km (12.5 miles) with respect to the existing conventional line. In order to reduce travel times only one intermediate stop at Dundalk has been assumed (see Figure 16).

Table 4 Data used for our high speed proposals including the segment definition, their lengths and costs in single and double track.

Segment			Length		Cost (M€/km M£/mi)			
id	Origin	Destination	Km	Miles	Simple		Double	
1	Dublin	Donabate	16,5	10,3	7,8	9,1	14,0	16,3
2	Donabate	Pilltown	27,7	17,3	4,9	5,7	7,8	9,1
3	Pilltown	Dromiskin	28,1	17,6	5,3	6,2	8,4	9,8
4	Dromiskin	Dundalk	9,1	5,7	6,9	8,0	12,4	14,5
5	Dundalk	Newry	22,1	13,8	5,8	6,8	8,9	10,4
6	Newry	Banbridge	18,1	11,3	5,5	6,4	8,7	10,1
7	Banbridge	Mazetown	22,7	14,2	6,1	7,1	9,5	11,1
8	Mazetown	Belfast	16,1	10,1	7,4	8,6	13,9	16,2

Table 4 shows the data used for our high speed proposals including the segment definition, their lengths and costs in single and double track.

Our first high speed line proposal is a traditional double-track line, i.e., all segments are built as double-track segments. Consequently, our computer program does not need to make any decision on which segments must go in single or double-track, but it needs to optimize travel times. he resulting optimal travel time between Dublin and Belfast is 50 min., which means a reduction of 1 hour and 19 minutes with respect to the current travel time.

To summarize, the main characteristics of the proposed high speed line in the case of a doubletrack solution are: a total length of 160.4 km/100.25 Miles, a travel time of 50 min. and a construction cost of M€ 1589.54 (M£ 1158.98).

This solution is the most expensive one, but offers maximum capacity and flexibility. However, the main questions that should be asked are: whether or not to resort to a double-track solution is necessary and whether or not there are more efficient alternatives.

8.3.3 ADST proposals

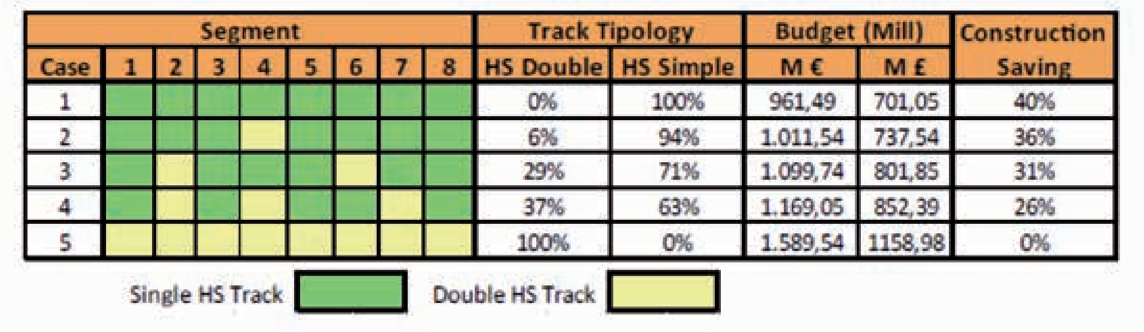
With the aim of reducing construction and maintenance costs, in this section some alternative ADST line solutions are discussed. The decision of which segments should be in double- and single-track under the assumed demand assumptions and the corresponding timetable optimization will be made with the help of the optimization program.

Only the routes of Dublin-Belfast and Dublin-Dundalk-Belfast in both directions and a demand of 16 and 32 daily services have been considered in this preliminary analysis.

The commercial speed used in the analysis was assumed 90% of the maximum speed for each segment. In addition, a headway of 4 minutes and a delay of 4 minutes per hour of travel has been assumed. This commercial speed and delay values were considered to guarantee a certain robustness of the timetables under a regular operation.

After using the proposed method, we have considered 5 different solutions that range from single track to double track with 3 intermediate cases that include one and three double track segments. They are illustrated in table Table 5, where it can be shown that the budget ranges from € 962 M/ £ 701 M, for single track to € 1590 M/ £ 1160 M, for double track.

Table 5 Cost comparison of the 5 alternatives considered.

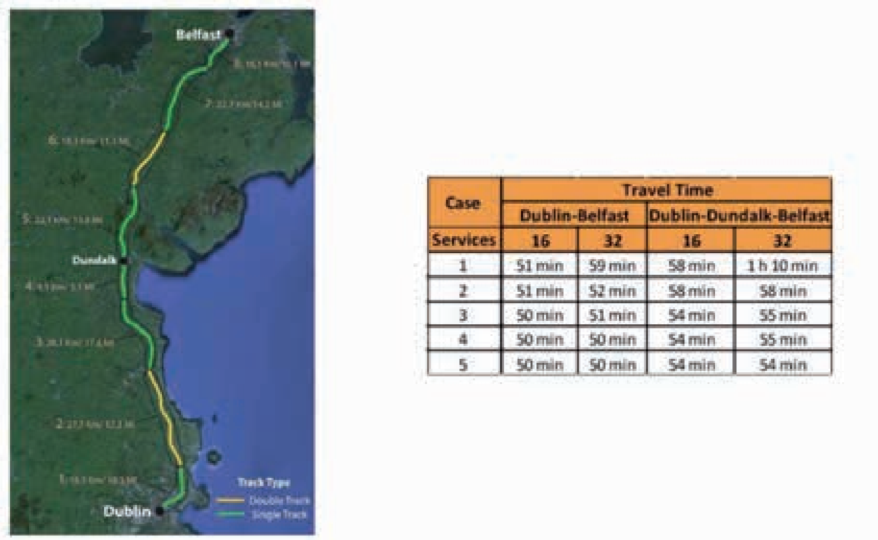


Segment									Track Tipology		Budget (Mill)		Construction Saving
Case	1	2	3	4	5	6	7	8	HS Double	HS Simple	M €	M £	
1	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	0%	100%	961,49	701,05	40%
2	🟩	🟩	🟩	🟨	🟩	🟩	🟩	🟩	6%	94%	1.011,54	737,54	36%
3	🟩	🟨	🟩	🟩	🟩	🟨	🟩	🟩	29%	71%	1.099,74	801,85	31%
4	🟩	🟨	🟩	🟨	🟩	🟩	🟨	🟩	37%	63%	1.169,05	852,39	26%
5	🟨	🟨	🟨	🟨	🟨	🟨	🟨	🟨	100%	0%	1.589,54	1158,98	0%

Single HS Track 🟩 Double HS Track 🟨

The travel time comparison of the 5 alternatives considered in Figure 17 show that any DublinBelfast service does not take longer than one hour and in the route Dublin-Dundalk-Belfast travel time exceeds one hour only in the single track case.

KEYNOTE



Case	Travel Time			
	Dublin-Belfast		Dublin-Dundalk-Belfast	
Services	16	32	16	32
1	51 min	59 min	58 min	1 h 10 min
2	51 min	52 min	58 min	58 min
3	50 min	51 min	54 min	55 min
4	50 min	50 min	54 min	55 min
5	50 min	50 min	54 min	54 min

Figure 17. Optimal ADST railway line solution and Travel time comparison of the 5 alternatives considered.}

The optimal option is given by Case 3, a high speed line proposal composed of 2 double-track segments (in yellow color) and 6 single-track segments as shown in Figure 17.

Its main characteristics are: total length 160,4 Km , construction budget € 1.099,74 M (31 % saving w. r. t. double track), travel time 50-51 min (Dublin-Belfast) and \\ 54-55 min (DublinDundalk-Belfast), 16 and 32 daily services with a capacity of 6.200 and 12.400 daily passengers.

The resulting timetable for 32 services is shown in figure Figure 18.

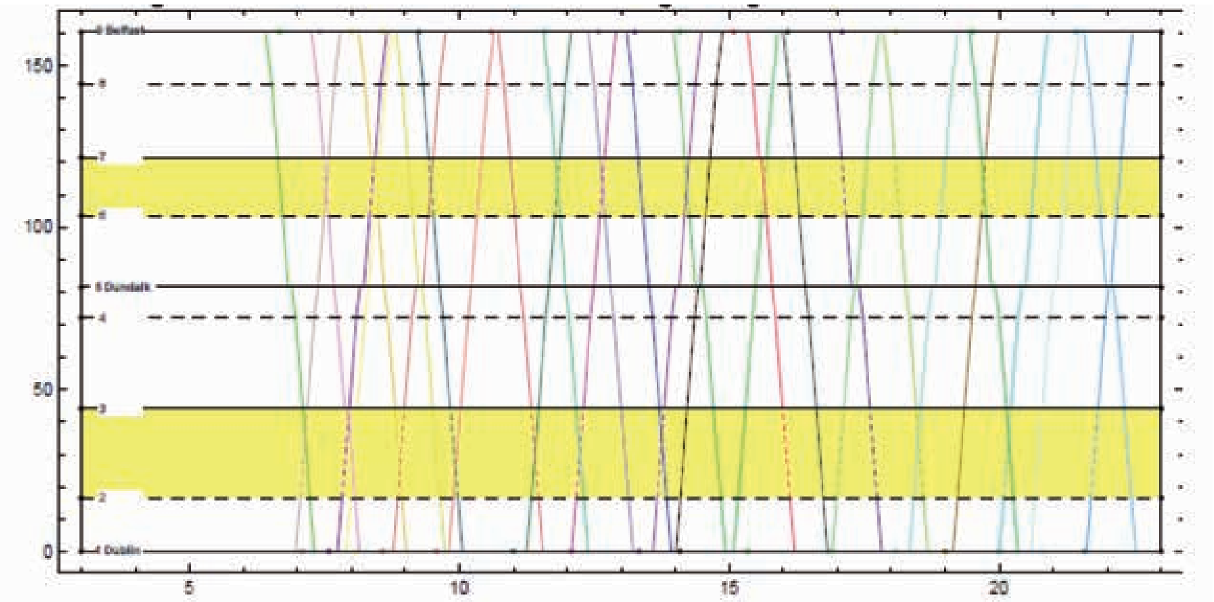
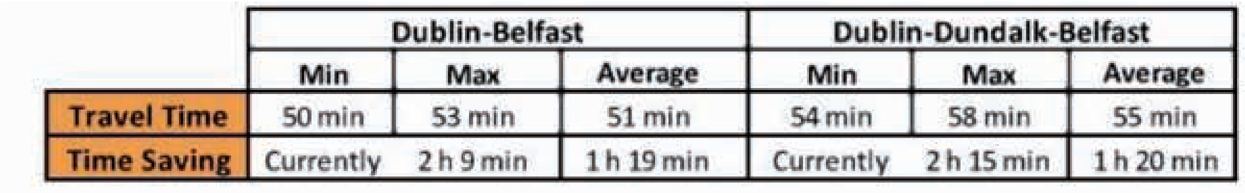


Figure 18. 32 services Timetable for the Dublin-Belfast line.

Table 6 Travel times and travel time reductions w.r.t. actual travel times.



	Dublin-Belfast			Dublin-Dundalk-Belfast		
	Min	Max	Average	Min	Max	Average
Travel Time	50 min	53 min	51 min	54 min	58 min	55 min
Time Saving	Currently	2 h 9 min	1 h 19 min	Currently	2 h 15 min	1 h 20 min

Table 6 shows the travel times and travel time reductions w.r.t. actual travel times.

8.3.4 Improvement of the existing line

In this section a final alternative that combines the existing conventional line with segments of new construction is proposed.

We propose to build 4 new segments, as shown in figure Figure 19 that is, between DonabateJulianstown-Dromiskin (segments T1-T2 and T2-T3) and Newry-Banbridge-Lisburn (segments T6T7 and T7-T8). These new segments would operate to complement the current line, so that the services (long distance link, commuter and freight traffic) could circulate using the current and new tracks, without any restriction.

The new segments should be designed to accept trains with a range of speed between 130 km/h (80 mph) and the future high speed 250 km/h (155 mph).

We study an improved line with 16+18 services between Dublin and Belfast where the new segments are assumed to be in operation with only 28 fast connection services of a total 34 links between Dublin and Belfast circulating along the new segments. With this, the total amount of services along the network is 346 trains.

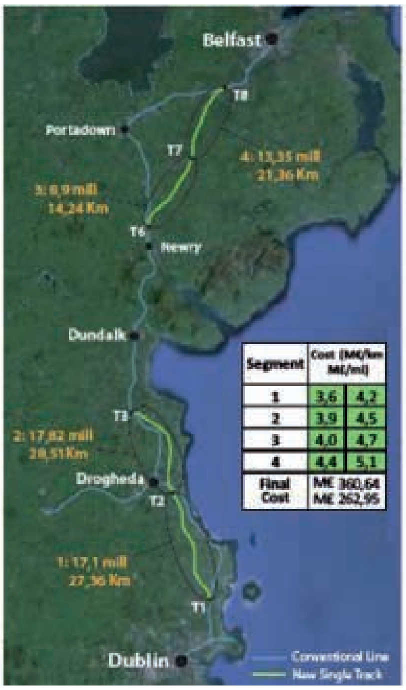


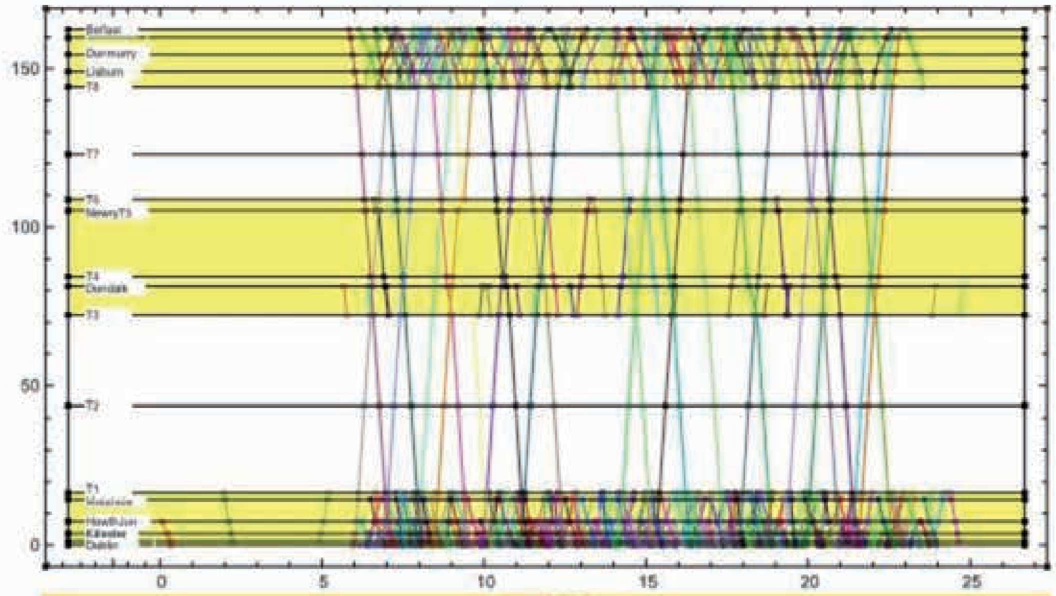
Figure 19. 4 segments to be constructed to improve the existing line.

KEYNOTE

The characteristics of the resulting optimized line are: it is a line with total length 161 km/100.6 Miles with 4 new single track segments of length (91.5 km/57.2 mile), which imply a total cost of M€ 360.64/M£ 262.95.

Finally, the result of this analysis provides the following conclusions:

1. Fast connections, less than 1 hour 25 minutes, between Dublin-Belfast are possible. They would increase Dublin-Belfast traffic rate, and generate new users.
2. New construction segments are suggested only out of the two congested metropolitan areas.
3. Four single track segments with a total length of 91.5 km (57.2 Miles) are proposed, because double-track segments are not necessary. Therefore, the resulting construction cost of the proposed line is only M€ 360.64 (M£ 264.95).
4. The new segments permit fast connection, reduce congestion in the whole network and favor cross-border freight transport.



Travel Time						
Route	**11**	**12**	**101**	**102**	**111**	**112**
Min	2 h 2 min	2 h 7 min	1 h 25 min	1 h 25 min	1 h 21 min	1 h 23 min
Max	2 h 12 min	2 h 13 min	1 h 32 min	1 h 32 min	1 h 27 min	1 h 28 min
Average	2 h 7 min	2 h 11 min	1 h 29 min	1 h 31 min	1 h 24 min	1 h 26 min

Figure 20. Improved Network (16+18 Dublin-Belfast services).

9. Conclusions

As final conclusion, Alternate Double-Single track projects, as different examples reflect, allow to:

- Minimize the construction cost with reduced travel times.
- Design railway lines under current and future demands.
- Reduce maintenance costs.
- Optimize timetables.
- Model the timetable in response to incidents on the network.

10. Bibliography

- Amit, I. & Goldfarb, D., 1971. The timetable problem for railways. *Developments in Operations Research,* Volume 2, pp. 379-387.
- Assad, A., 1980. Models for rail transportation. *Transportation Research Part A,* Volume 14, pp. 205-220.
- Burdett, R. L. & Kozan, E., 2010. A disjunctive graph model and framework for constructing new train schedules. *European Journal of Operational Research,* Volume 2010, pp. 8598.
- Cacchiani, V. & Toth, P., 2012. Nominal and Robust Train Timetabling Problems. *European Journal of Operational Research,* Volume 219, pp. 727-737.
- Caprara, A., Fischetti, M. & Toth, P., 2002. Modeling and solving the train timetabling problem. *Operations Research,* Volume 50, pp. 851-861.
- Carey, M., 1994. A model and strategy for train pathing with choice of lines, platforms and routes. *Transportation Research Part B,* Volume 28, pp. 333-353.
- Carey, M. & Crawford, I., 2007. Scheduling trains on a network of busy complex stations. *Transportation Research Part B,* Volume 41, pp. 159-178.
- Carey, M. & Lockwood, D., 1995. A model, algorithms and strategy for train pathing. *Journal of the Operational Research Society,* Volume 46, pp. 988-1005.
- Castillo, E. et al., 2015. An alternate double-single track proposal for high-speed peripheral railway lines. *Computer Aided Civil and Infrastructure Engineering,* Volume 30, pp. 181201.
- Castillo, E., Gallego, I., Ureña, J. M. & Coronado, J. M., 2009. Timetabling optimization of a single railway track line with sensitivity analysis. *TOP,* Volume 17, pp. 256-287.
- Castillo, E., Gallego, I., Ureña, J. M. & Coronado, J. M., 2011. Timetabling optimization of a mixed double- and single-tracked railway network. *Applied Mathematical Modelling,* Volume 35, pp. 859-878.
- Castillo, E., Grande, Z., Moraga, P. & Sánchez Vizcano, J., 2016. A time partitioning technique for railway line design and timetable optimization. *Computer Aided Civil And Infrastructure Engineering,* Volume 31, pp. 599-616.
- Cordeau, J. F., Toth, P. & Vigo, D., 1998. A survey of optimization models for train routing and scheduling. *Transportation Science,* Volume 32, pp. 380-404.
- DÁriano, A., Pacciarelli, D. & Pranzo, M., 2007. A branch and bound algorithm for scheduling trains in a railway network. *European Journal of Operational Research,* Volume 183, pp. 643-657.

- DÁriano, A. & Pranzo, M., 2004. *A real time train dispatching system based on blocking time theory.* Delft, DUP Science, pp. 129-152.
- DÁriano, A., Pranzo, M. & Hansen, I. A., 2007. Conflict Resolution and Train Speed Coordination for Solving Real-Time Timetable Perturbations. *IEEE Transactions on Intelligent Transportation Systems,* Volume 8.
- Haghani, A. E., 1987. Rail freight transportation: a review of recent optimization models for train routing and empty car distribution. *Journal of Advanced Transportation,* Volume 21, pp. 14-172.
- Hellström, P., 1998. *Analysis and evaluation of systems and algorithms for computer-aided train dispatching,* Sweden: s.n.
- Higgins, A., Kozan, E. & Ferreira, L., 1996. Optimal scheduling of trains on a single line track. *Transportation Research Part B,* Volume 30, pp. 147-161.
- Jia, L. M. & Zhang, X. D., 1993. Distributed intelligent railway traffic control based on fuzzy decision making. *Fuzzy Sets and Systems,* Volume 62, pp. 255-265.
- Kraay, D. R. & Harker, P. T., 1995. Real-time scheduling of freight railroads. *Transportation Research Part B,* Volume 29, pp. 213-229.
- Lin, D. Y. & Ku, Y. H., 2013. Using Genetic Algorithms to Optimize Stopping Patterns for Passenger Rail Transportation. *Computer-Aided Civil and Infrastructure Engineering,* pp. n/a--n/a.
- Ouyang, Y. et al., 2009. Optimal Locations of Railroad Wayside Defect Detection Installations. *Computer-Aided Civil and Infrastructure Engineering,* Volume 24, pp. 309-319.
- Pachl, J., 2014. Railway Timetabling and Operations. Analysis - Modelling - Optimisation - Simulation - Performance Evaluation. In: Hamburg(): Eurailpress, pp. 23-24.
- Petersen, E. R., Taylor, A. J. & Martland, C. D., 1986. An introduction to computer aided train dispatching. *Journal of Advanced Transportation,* Volume 20, pp. 63-72.
- Sahin, I., 1999. Railway traffic control and train scheduling based on inter-train conflict management. *Transportation Research Part B,* Volume 33, pp. 511-534.
- Yang, Z. & Hayashi, Y., 2002. GIS-Based Analysis of Railways Origin/Destination Path-Selecting Behavior. *Computer-Aided Civil and Infrastructure Engineering,* Volume 17, pp. 221-226.

High Speed Rail Infrastructure

360.revista de alta velocidad
número 5 - junio - 2018. Pág 191 - 217

Testing railway tracks at 1:1 scale at CEDEX Track Box

Estaire, José
Cuéllar, Vicente
Pardo de Santayana, Fernando
Santana, María

Laboratorio de Geotecnia, CEDEX[1]

Abstract

The paper describes the experimental facility built at CEDEX to bridge the gap between simple laboratory tests and in-situ measurements to study the mechanical behaviour of railways tracks. The texting facility, called CEDEX Track Box (CTB), is a 21 m long, 5 m wide and 4 m deep device whose main objective is to test, at 1:1 scale, complete railway track sections of conventional and high speed lines for passenger, freight and mixed traffics, at speeds up to 450 km/h. The railway track response, in terms of displacements, velocities and accelerations, is collected from a great number of linear variable differential transformers (LVDTs), geophones, accelerometers and pressure cells installed both inside the embankment and the bed layers (ballast, sub-ballast and form layer) of the track. On the other hand, the railway superstructure response is recorded with a number of mechanical displacement transducers, laser sensors, geophones and accelerometers installed on the different track superstructure components: rail, sleeper and railpad.

After checking the results obtained in CTB with in-situ measurements carried out at different points in the Spanish high speed line network, CTB has been used to research on different matters: measurement of track vertical stiffness under different track conditions by imposing static loads step by step through a set of servo-hydraulic actuators and measuring the rail deflection as a function of the load applied; calibration of 3-D numerical models under static conditions to show the crucial role played by the non-linear behaviour of the different materials constituting the track bed layers and the embankment or the trench supporting ground on the behaviour of ballasted and slab tracks; use of a commercially available machine to determine the effect of repeating tamping operations in ballast degradation; assessment of track lateral stability in CTB with the aid of a special tool that pushes away the sleeper while its horizontal movement is recorded; study of the short and the long term behaviour of railway track sections through the obtainment of ballast, subballast and form layer permanent settlement curves in fatigue tests in which, at least, one million axles were applied, under different test conditions; determination of rail deflections under the pass-by of trains at very high speeds, up to 450 km/h; optimization of bituminous subballast thickness for high speed lines under mixed traffic conditions; abatement of noise by the use of sleepers with USP.

Furthermore, CTB can be also used to: study the influence of track or ground irregularities in the behaviour of slab tracks and the corrective actions to be taken; optimize the maintenance works in slab-ballast transition zones; analyse the behaviour of curved sections; and to assess the ballast degradation in switches and crossings.Recent developments already made, such as the deployment of a 3Hz natural frequency spring system at the CTB base, will allow in the future analysing experimentally the behaviour of ballast and slab tracks founded on soft soils. A summary of some of the results obtained during the first working years are shown in this paper.

Keywords: railway infrastructure, testing facility, track stiffness, lateral resistance, ballast degradation, critical speed.

---

[1] *Estaire, José. Laboratorio de Geotecnia, CEDEX, Madrid (Spain). Email: jose.estaire@cedex.es. (corresponding author)*
Cuéllar, Vicente. Laboratorio de Geotecnia, CEDEX, Madrid (Spain).
Pardo de Santayana, Fernando. Laboratorio de Geotecnia, CEDEX, Madrid (Spain).
Santana, María. Laboratorio de Geotecnia, CEDEX, Madrid (Spain)

1. Introduction

The experimental research of any track component or configuration usually begins with the performance of simple laboratory tests, being the last step to perform tests with real traffic in a real railway track. However this last activity is a quite complicated task due to the difficulties to have access to the track and the complications derived of the in-situ instrumentation installation.

To bridge the gap between simple laboratory tests and in-situ measurements in railways tracks, CEDEX built the CEDEX Track Box 12 years ago, as part of Supertrack project (2001-05) in the frame of European Union Fifth Framework Program (Manzanas et al, 2007).

Since that, this testing facility has been extensively used in the following European Union projects: Innotrack (2005-09), Rivas (2009-13) and Capacity for Rails (2013-17) (Cuéllar, 2016).

2. General description of CEDEX Track Box

CEDEX Track Box (CTB) is a 21 m long, 5 m wide and 4 m deep facility whose main objective is to test, at 1:1 scale, complete railway track sections of conventional and high speed lines for passenger and freight trains, at speeds up to 450 km/h. Figure1 shows a general view of the testing facility.

Figure 1. General view of the testing facility

The testing facility was designed, built and developed as part of SUPERTRACK (“Sustained Performance of Railway Tracks”, 2001-05) and INNOTRACK (“Innovative Track Systems”, 2005-2009) projects funded by the European Union Fifth and Sixth Framework Programs, respectively.

Its principal advantage is the possibility of performing fatigue tests in a fast way as in one working week, the effect of the passing-by of trains during a year in a real section can be modelled.

The reproduction of the effect of an approaching, passing-by and departing train in a test cross-section, as it occurs in a real track section, is performed by application of loads, adequately unphased as a function of the velocity of the train which is being simulated, produced by three pairs of servo-hydraulic actuators (that can apply a maximum load of 250 kN at a frequency of 50 Hz), placed on each rail and 1.5 m longitudinally separated, as seen in Figure 2.

Figure 2. Detail of the hydraulic actuators

Furthermore, the reproduction of wheel and track imperfection effects that produces low amplitude high frequency dynamic loads can also be carried out by the use of two piezoelectric actuators that can apply loads up to 20 kN at 300 Hz.

The railway track response, in terms of displacements, velocities, accelerations and pressures, is collected from a great number of linear variable differential transformers (LVDTs), geophones, accelerometers and pressure cells installed inside both the embankment and the bed layers (ballast, sub-ballast and form layer) of the track.

On the other hand, the railway superstructure response is recorded with mechanical displacement transducers, laser sensors, geophones and accelerometers installed on the different track components (rail, sleeper and railpad), as seen in Figure 3. The acquisition data unit can receive information from 150 sensors at the same time.

Figure 3. View of some of the instrumentation installed in the superstructure

Lastly, CTB is equipped with a little commercially available tamping machine, shown in Figure 4, to tamp the ballast if it is required in some of the tests.

Figure 4. General view of the tamping machine used in CTB

The 1:1 scale models that can be built in CTB can reproduce the following track conditions:

- Ballasted or slab tracks.
- Sections in straight line or in curve.
- Switches and crossings.
- Transitions zones.
- Sections with different kinds of ballast, sub-ballast, form layer or embankment.
- Sections with standard, polyvalent and three-rail sleepers.
- Sections equipped with new materials: sleepers with USP, under ballast mats, artificial ballast, bituminous sub-ballast, geotextiles and soils treated with lime or cement.

With these 1:1 scale models, a set of different tests can be performed having the following characteristics:

- Tests with passenger, freight trains and mixed traffic.
- Tests with static loads to determine track stiffness.
- Tests with quasi -static loads to simulate the pass-by of trains at different speeds up to 450 km/h.
- Tests with dynamic loads to simulate the effects induced by track and wheel irregularities.
- Tests to determine the fatigue behaviour of any track component (mainly, the fastening system, the ballast, sub-ballast layer) by the simulation of pass-by of millions of axle trains.
- Tests on vibration propagation and abatement solutions.
- Tests to determine the lateral and longitudinal track resistance.

The results obtained in the tests can be used mainly to analyse the short and long term behaviour of railway track sections submitted to any kind of train traffic and to calibrate 3D numerical models to be used in other type of studies or to widen the aim of the tests.

CTB has been used in the past to research on different matters, which are described in the following sections.

- Measurement of track vertical stiffness under different track conditions.
- Determination of track lateral stability.
- Study of short and long term behaviour of railway track sections by obtaining ballast, subballast and form layer permanent settlement curves in fatigue tests and under tamping operations.
- Measuring of track mechanic behaviour under the pass-by of trains at very high speeds, up to 450 km/h.
- Behaviour of sleepers with USP.
- Optimization of bituminous sub-ballast thickness.
- Behaviour of High Speed Lines subjected to mixed traffic.
- Calibration of 3D numerical models.

3. Static tests

3.1 Measurement of track vertical stiffness

The measurement of track vertical stiffness in any track condition is made imposing static loads by a pair of servohydraulic actuators that are installed in each of the cross-sections (A, B or C in Figure 5) on each of the rails (see Figure 2).

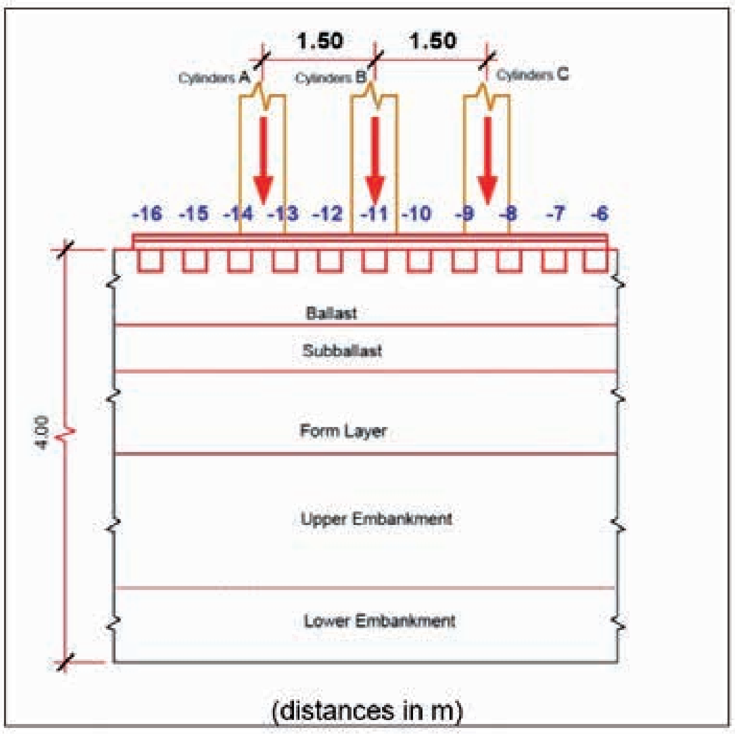


Figure 5. Cross section of the 1:1 scale model used in the tests with the position of the servo-hydraulic actuators

Each pair of actuators is activated simultaneously with the same time-load history. The static tests consist usually on two loading-unloading cycles. In the case shown in Figure 6 a maximum load level of 92.5 kN was reached, step by step, and a minimum load of 2.5 kN was applied to avoid a loss of contact between the actuators and the rail. Once a load level of 82.5 kN was attained in each loading cycle, two small excursions of ±10 kN were programmed to determine the dynamic track stiffness. The duration of each static test was 12 minutes. A total of four static test were performed, one with each actuator independently and another with the three pairs of actuators acting together.

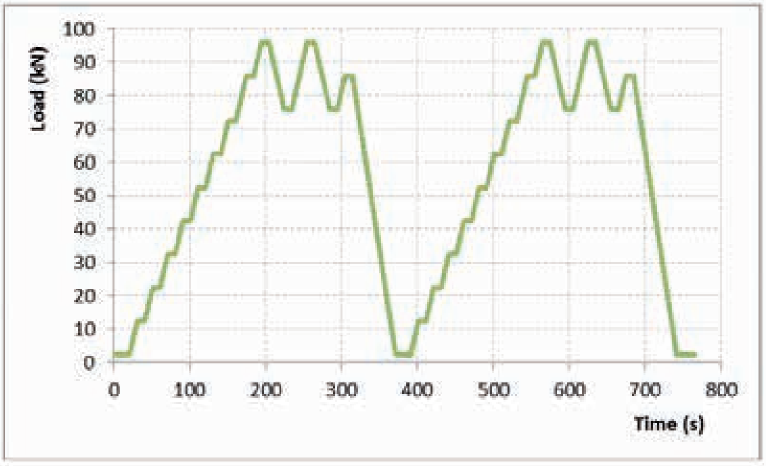


Figure 6. Time-load history used in the static tests

The rail vertical displacements were recorded with eight laser systems distributed in the three test sections. Figure 7 shows the distribution of the laser systems installed (blue labels) and the location of the three pairs of actuators (yellow squares).

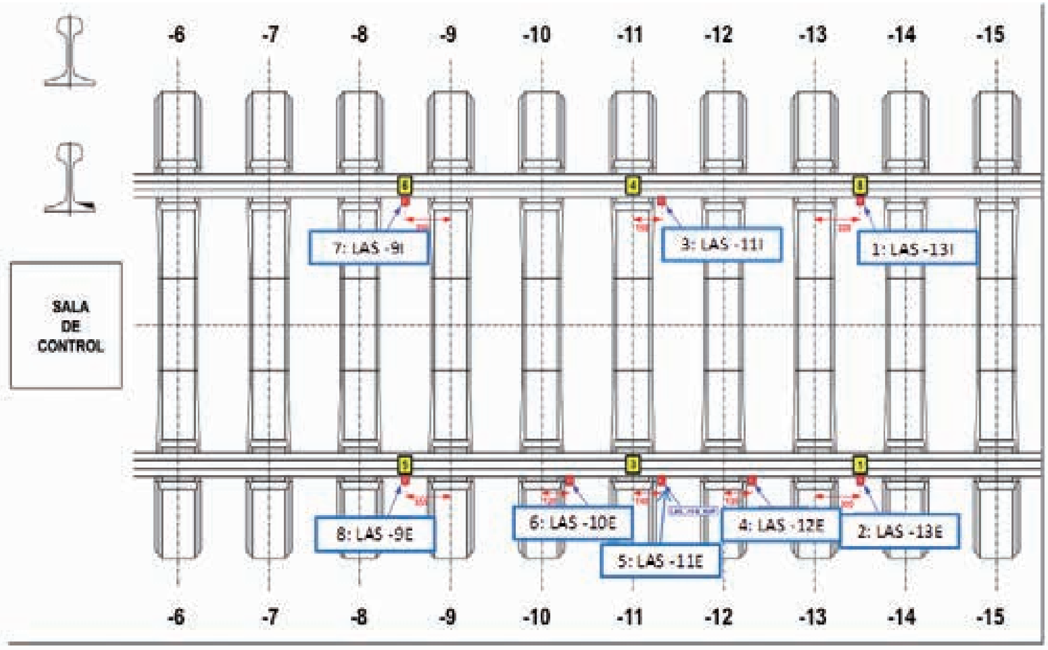


Figure 7. Location of eight laser systems (blue labels) and location of the three pairs of actuators (yellow squares)

Figure 8 shows the evolution with time of the rail displacement, measured with all the laser systems installed for the four static tests performed. These data can be used to analyse the rail deflection, as a function of the load applied, as shown in Figure 9. It can be seen that, in the case of a hard embankment combined with hard bed layers, the deflection curves are clearly non-linear so railway stiffness deduced from them should always be referred to the load in which it is being calculated. It is important to notice that this non-linear effect can even increase in the case of hard track bed layers on a soft soil foundation.

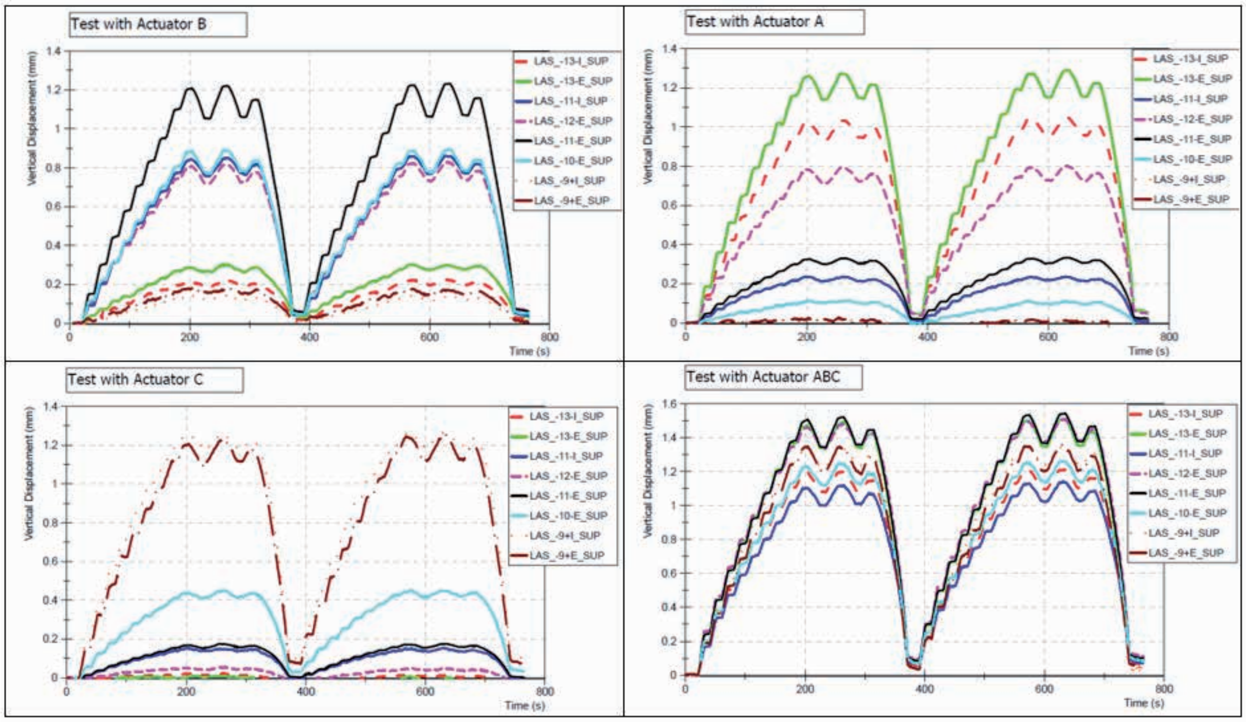


Figure 8. Rail displacement evolution with time for all the static tests

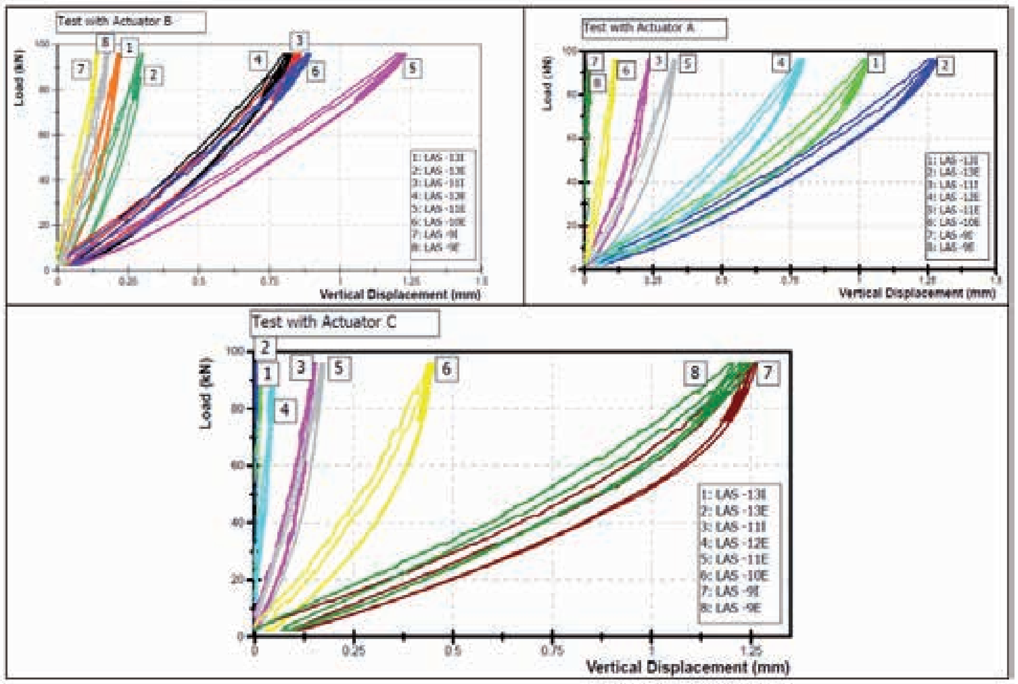


Figure 9. Deflection curves obtained in the static tests for hard embankment and hard bed layers

The next step is to analyze the rail deflections, for a certain load level, along the rail as a function of the distance from the load application point, as seen in Figure 10 for the static tests in which the three actuators were acting independently.

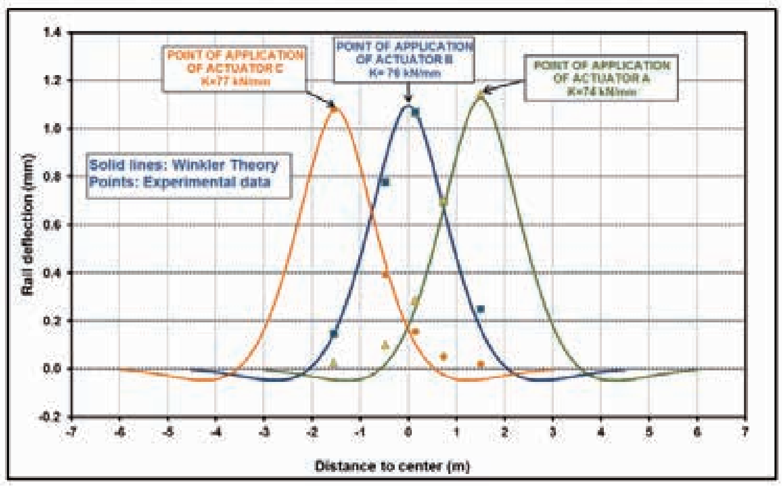


Figure 10. Rail deflection at different points during a set of static tests, including rail stiffness values obtained for a load of 83.34 kN per wheel

The good fitting between the measures and the theory, seen in Figure 10 and in a great number of other static tests, proves that the rail deflections due to a vertical single load can be predicted with high accuracy assuming the rail track has a Winkler-type behaviour quantified by Equation 1.

$$\delta(x,t) = \frac{Q}{K}\, e^{-\frac{|x-vt|}{L}} \left[cos\left(\frac{|x-vt|}{L}\right) + sin\left(\frac{|x-vt|}{L}\right) \right] \qquad (1)$$

As mentioned previously, the rail deflection-load curves were non-linear so track stiffness must be referred to the load imposed. Figure 11 shows the variation of the track stiffness for the load application point of the central actuator (B) in the two cycles. It can be seen that, for the second cycle, track stiffness increases logarithmically from 0 to 75 kN/mm when the load applied is in the range between 2 and 95 kN, while for the first cycle the valid values form a straight line beginning at a load of 12 kN. Below that load, the values cannot be considered valid due to the concavity of the results. These anomalous results can have been produced by a rearrangement of ballast particles during the first load cycle which indicated ballast was not enough stabilized at the beginning of the test.

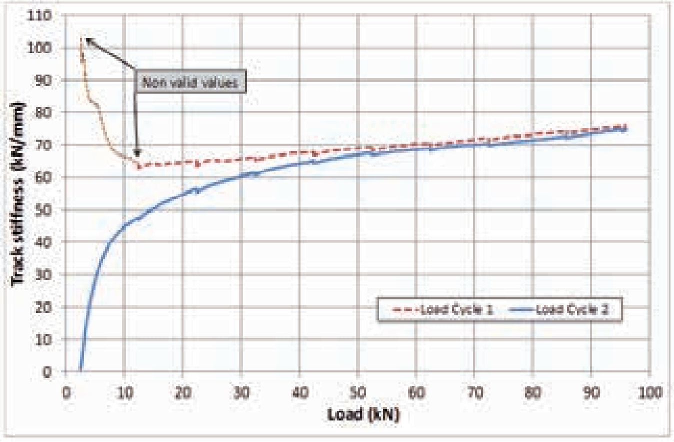


Figure 11. Track stiffness as a function of the load applied

3.2 Contribution of the track layers to the total rail deflection

In many of the static tests performed in CTB, besides the rail deflection recorded with laser systems, the track movements of other track elements can be measured:

- rail-pad vertical displacements, measured as rail-sleeper relative displacements, recorded with resistivemeters.
- ballast layer settlements, measured by LVDT inductive sensors, mounted attaching their heads to the sleepers and fixing their base to the top of the sub-ballast layer.

Figure 12 shows some photographs of the instrumentation used in the static tests.

Laser system Resistivemeter LVDT

Figure 12. Instrumentation used in the static tests

With the data recorded in the different elements previously listed, the contribution of each railway track layer to the total rail deflection was possible to be determined, as Table 1 reflects.

Table 1. Summary of results of the static tests performed

Train	Type	Passenger				Freight	
	Speed (km/h)	300				120	
	Load (kN/axle)	165				220 - 245	
Physical model (1)		1	2	3	4	2	3
Track stiffness K (kN/mm)		100	120	125	118	130	140
Contribution (%)	Pad	25	35	32	34	40	39
	Ballast	45	43	45	41	38	43
	Subballast	15	2	3	4	2	3
	Form layer	15	7	20	21	5	15
	Embankment		13			15	

(1) Subballast layers in the physical models:
Model 1: 30 cm granular; Model 2: 8 cm bituminous
Model 3: 12 cm granular; Model 4: 16 cm bituminous

Physical models 1 to 4, referred in Table 1, only differ in the sub-ballast layer, as collected in the table, while the other elements are common: 100 kN/mm stiff pad, 35 cm thick ballast layer, 60 cm thick form layer and 2.5 m high embankment with a shear wave velocity higher than 200 m/s. Figure 13 (left) is the section type of Model 1, while Figure 13 (right) shows a schematic drawing of physical Models 2, 3 and 4.

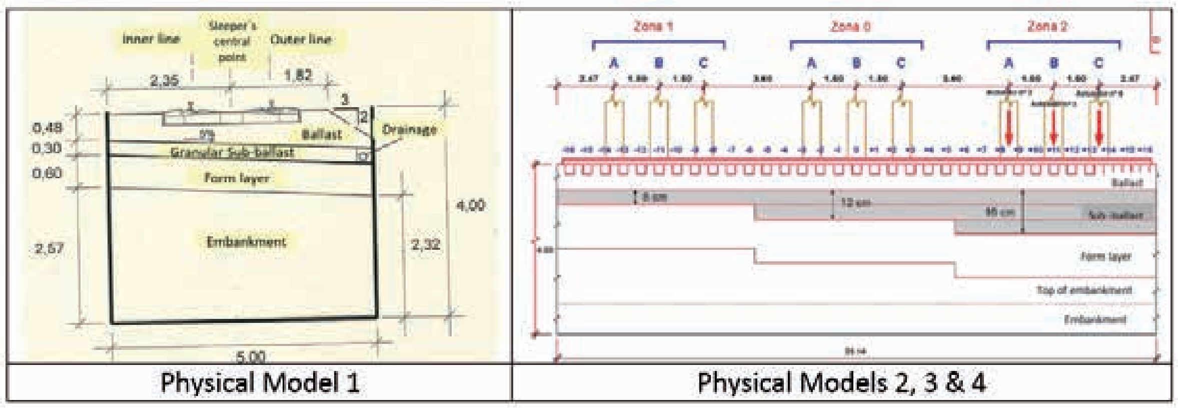


Figure 13. Section type of physical Model 1 (left) and schematic drawing of physical Models 2, 3 and 4 (right)

It must be highlighted, from Table 1, the track consolidation observed after the pass-by of 4 million axles of passenger trains at 300 km/h, thanks to the increase in track stiffness, in the Physical models 2 (from 120 to 130 kN/mm) and 3 (from 125 to 140 kN/mm). On other hand, it can be stated that the ballast layer settlement represents about 40-45% of the total rail deflection, the pad contributes between 25 and 40% and the rest of the rail deflection is due to subballast, form layer and embankment (Table 1).

3.3 Determination of track lateral stability

The study of the track lateral stability can also be carried out in CEDEX Track Box with the aid of a special tool, shown in Figure 14, which pushes away the sleeper while its horizontal movement is recorded.

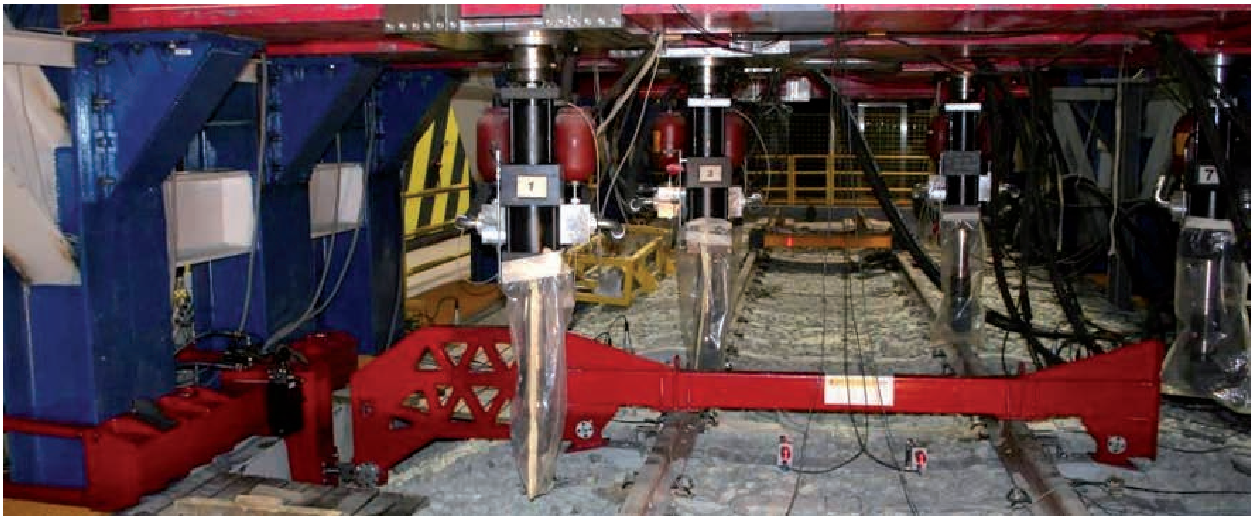


Figure 14. Tool designed to perform track lateral stability tests

The measurement equipment installed consisted on a load cell, two laser systems (to record the sleeper horizontal movement) and two potentiometers (to control the relative displacement between the sleeper and the rail).

Two different track lateral stability tests have been performed: in one of them the sleeper rested in a clean ballast layer while in the other one, the ballast layer was completely fouled with dry desert sand (Estaire et al. 2017). Test results showed a peak horizontal load of 12.5 and 16 kN in the tests performed with clean and fouled ballast, respectively, as shown in Figure 15. In both tests, peak load was reached when horizontal sleeper displacement was quite small (out 2.5 and 1.2 mm.

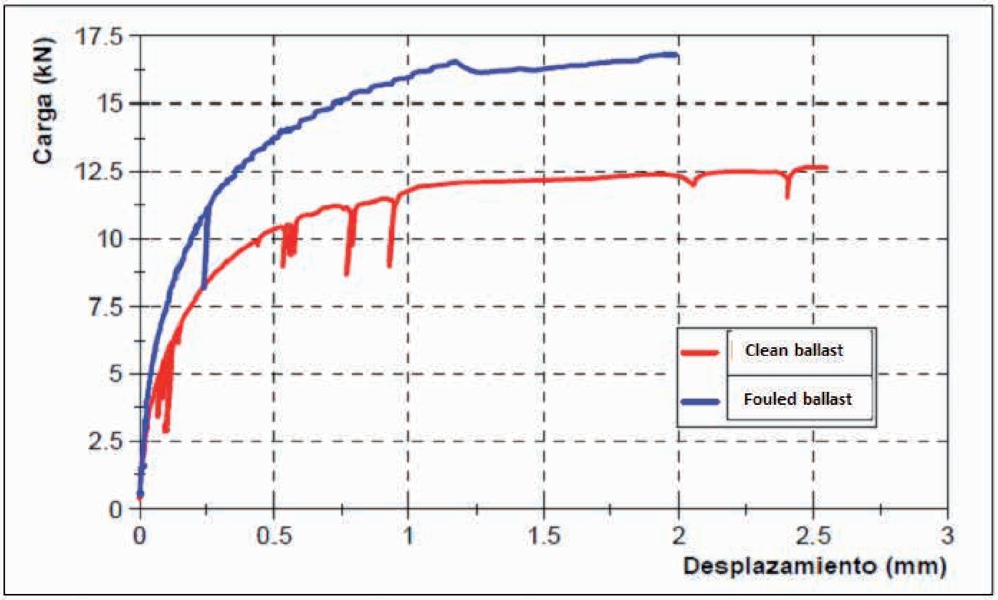


Figure 15. Result of the test performed with the sleeper on clean ballast

The results obtained in these tests show similar shapes and results than Single Tie Push Tests (STPT) performed in real railways in similar track conditions (Samavedam et al. 1999).

The track lateral stability tests were numerically modelled taking into account the following three mechanisms: the friction in the sleeper base with the ballast, the friction in the sleeper lateral faces with the ballast and the passive and active resistances of ballast in the sleeper shoulders (Kish, 2011).

The results of such modelling make it possible to draw the following conclusions:

- The main contributing components of the track lateral stability, when there is no vertical load applied, are the sleeper bottom friction (contribution between 25 and 45%, according to the cases analysed), the shoulder restraint in the frontal sleeper end (contribution between 25 and 60%), while the friction in the lateral faces accounts for the resting 15-30% (Estaire et al. 2017).
- When the sleeper is vertically loaded, the most important factor is the sleeper bottom friction, with a contribution higher than 80% and even increasing with the vertical load up to 95%, while the rest is shared between the passive resistance of ballast in the sleeper shoulder and the friction in the sleeper lateral faces, in a 1.5:1 proportion.
- Ballast friction angles used in the numerical modelling were in the range between 60 and 75° that can be considered quite high, although they are in accordance with the direct shear test results obtained in the very large shear box (1x1 m) belonging to CEDEX (Estaire and Santana, 2017).
- This similarity between the friction angles obtained in direct shear tests performed in

laboratory with the angles deduced from track lateral resistance tests makes it possible to give a great consistency and verisimilitude to both group of tests performed.

4. Quasi-static tests

4.1 Determination of ballast settlement curves

Since its beginning of work, 26 fatigue tests have been performed in CEDEX Track Box (CTB) in which, at least, one million axles were applied, under different test conditions:

- Two types of trains: passenger trains (with speeds between 300 and 320 km/h and axle loads mainly between 165 and 190 kN) and freight trains (running at a speed of 120 km/h and axle loads in the range between 220 and 245 kN).
- Two different types of sub-ballast layer: granular, with a thickness of 20 and 30 cm, and bituminous, with thickness of 8, 12 and 16 cm.
- Two different types of track systems: a) GIF AI-99 sleepers with a weight of 3.44 kN and rail pads with a stiffness of 100 kN/mm and b) B90.2 sleepers with a weight of 6.10 kN, equipped with an G04 (SLN 1010) type USP with 0.1 N/mm3 of static bedding modulus and rail pads with a stiffness of 450 kN/mm.
- Two different situations in the ballast layer: clean and fouled with desert sand in different proportions between 0 and 100%.
- The thickness of the ballast layer was 35 cm in all the tests, being formed by andesite particles.

In these tests, permanent settlement curves for the ballast, sub-ballast and form layers were obtained.

The set of the settlement curves obtained for the ballast layer in the tests performed, such as the ones shown in Figure 16, were analysed to discriminate the main factors that have influence in the track settlement and to obtain a mathematical expression to fit the results (Estaire et al. 2017).

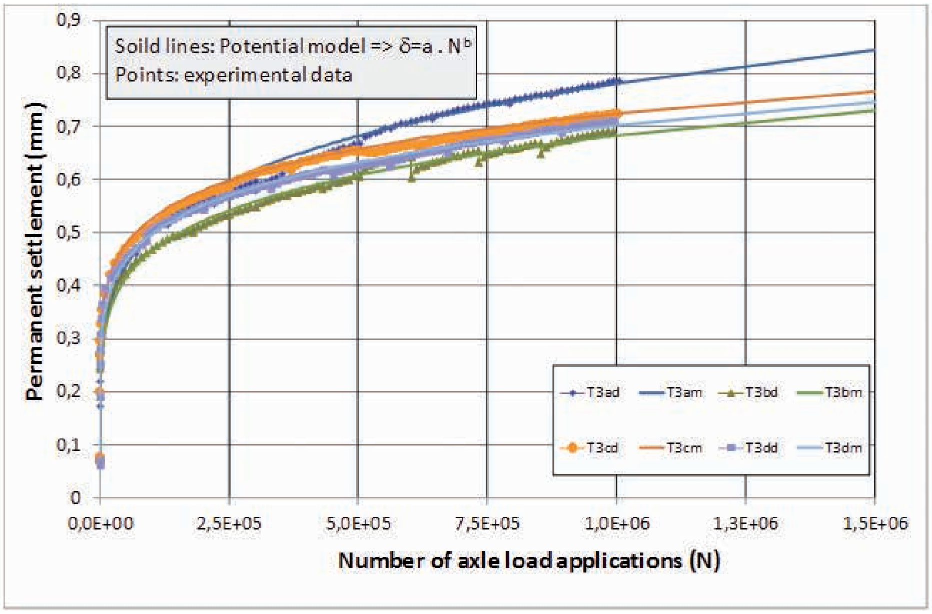


Figure 16. Some ballast settlement curves and their modelling with a potential model

The principal aspects that can be highlighted from the analysis of the experimental curves and their numerical modelling are:

- The values of the permanent settlement obtained in the tests are, in average, around 1 mm in the ballast layers, 0.03 mm in the bituminous sub-ballast layers and 0.02 mm in the form layers for 1 million of load axles, regardless the speed of the trains and the axle loads applied.
- The ballast settlement curves were modelled using a potential expression (δ = a . Nb), in which “a” represents theoretically the permanent settlement in the first cycle and “b” the rate of settlement growth with the number of axles applied. From a conceptual point of view, parameter “a” can be related with the axle load and parameter “b” with train speed, although in the final expression (Eq.2) train speed does not appear as external parameter.
- This model is different of the settlement models existing in literature, as shown with the review performed.
- A remarkable good adjustment of the test curves was obtained that confirms the validity of the potential model.
- The summary of the analysis performed for train speeds between 120 and 320 km/h leads to the following general expression of the ballast settlement law, as a function of the number of axle load applications (N), which is considered valid for axle loads (Q) between 110 and 250 kN:

$$\delta\ (\mathrm{mm}) = 0.0004\ .\ Q\ (\mathrm{kN})\ .\ N^{0.155} \qquad (2)$$

4.2 Kynematic behaviour of the different track elements

During the fatigue tests described before, the kinematic behaviour of the track elements could be determined.

On one hand, it has been shown that the track stiffness obtained in the quasi-static tests practically coincide with the ones shown in Table 1, for static tests, being the differences found in all the cases around 2-4%.

On other hand, the installation of geophones and accelerometers made it possible to measure velocities and accelerations in different track elements, as shown in Table 2, for trains travelling at about 300 km/h.

Table 2.Velocities and acceleration peaks obtained in the fatigue tests

	Velocity peaks (mm/s)		Acceleration peaks (g)	
Train	**Passenger**	**Freight**	**Passenger**	**Freight**
Rail	40 – 45	15 – 20	1.0 – 1.5	0.50 – 0.80
Sleeper	20 - 30	10 – 15	0.5 – 1.0	0.15 - 0.30
Ballast	15 – 20	7 - 10	< 0.5	< 0.15
Form layer	10 – 15	7 - 10	---	---
Embankment	2 - 6	< 4	---	---

These values must be considered as reference values when track is in good mechanical conditions, so they can be used to indicate the need to perform maintenance or repair works.

4.3 Determination of track behaviour for different speeds

4.3.1 Introduction

Some tests were performed in CTB modelling the pass-by of trains at different speeds (50/ 100/ 150/ 200/ 250/ 300/ 350/ 400 km/h) to analyse the effect of speed in the global response of the track. To do that, a Siemens S-100 train (a train with 13 bogies, with a separation between bogies of a wagon of 18.7 m and almost 200 m long) was modelled, supposing a constant wheel load of 71 kN instead of the real loads, that are in the range between 63 and 76 kN, as shown in Figure 17.

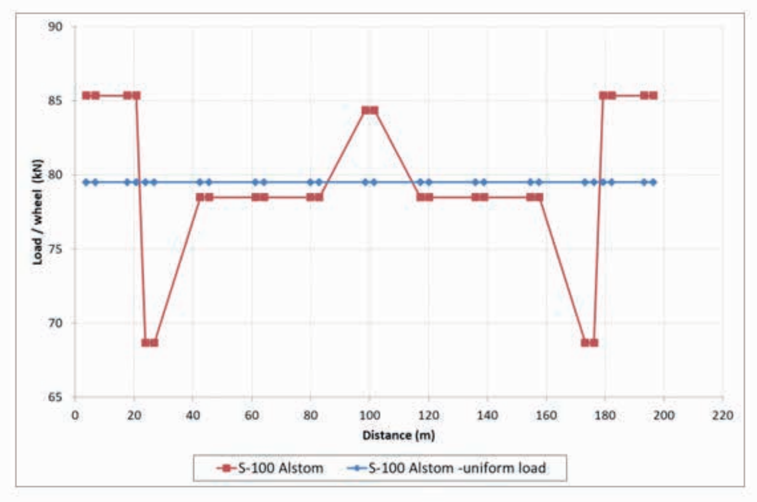


Figure 17. Loads per wheel in the modelled train used in CTB tests

Figures 18 and 19 show, as examples, the load-time history of the loads imposed by the modelled train and their frequency content, respectively, when the train speed is 150 km/h. It can be seen that, for that speed (v), the fundamental frequency is about 2.25 Hz which corresponds to the frequency (f) derived from the distance between bogies (d) in the same wagon (f=v/d).

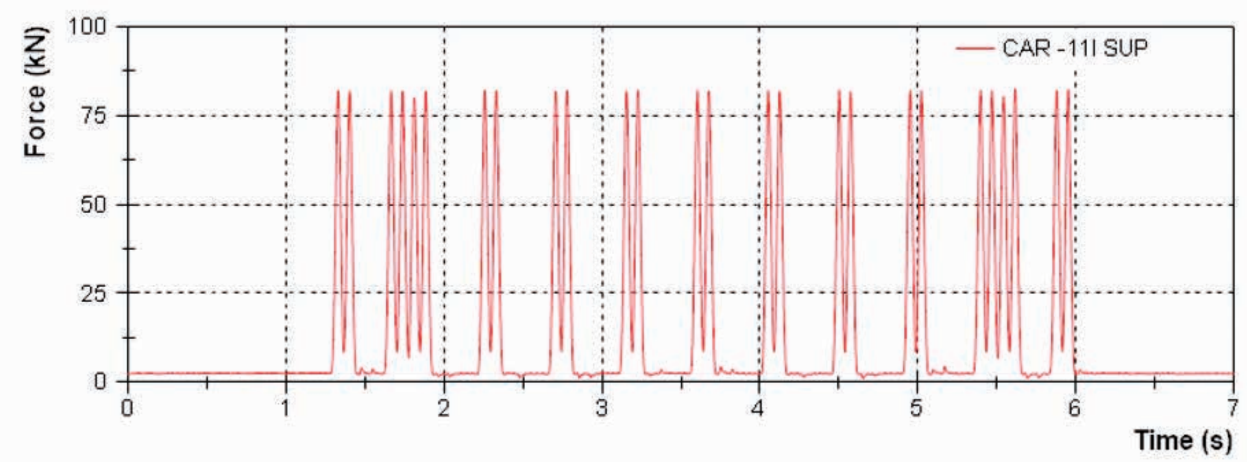


Figure 18. Time-load history in the modelled train used in CTB tests, running at 150 km/h

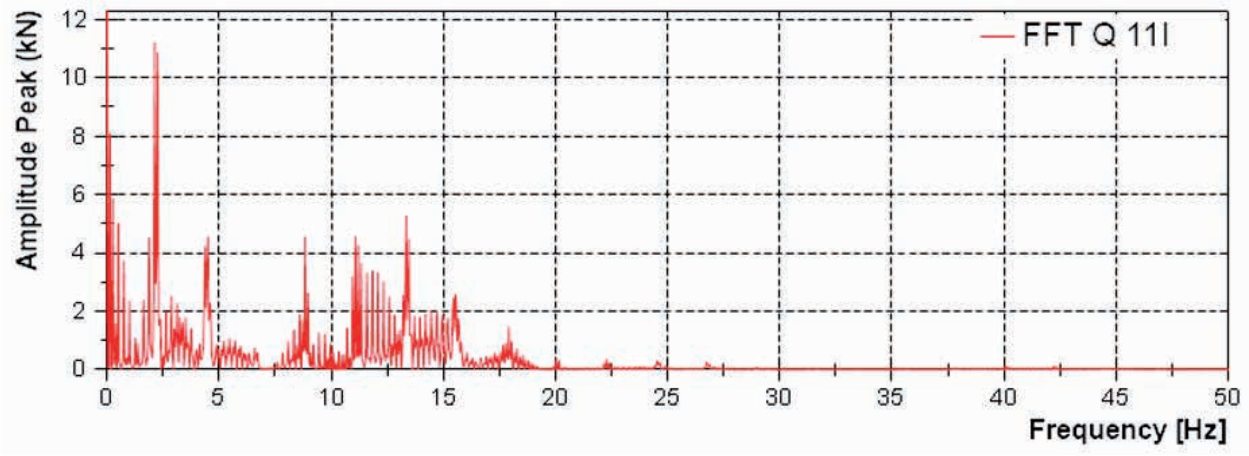


Figure 19. Frequency content of the load history in the modelled train used in CTB tests, running at 150 km/h

4.3.2 Rail deflection

Some of the test results performed with different speeds, in terms of rail deflections, are shown in Figure 20. The numerical results are collected in Table 3 that includes the result of a static test performed previously at the beginning of the tests.

Table 3. Test results for a S-100 train passing by at different speeds

Train speed (km/h)	Rail deflection (mm)	Relative rail deflection (1)
0	0.931	1.00
50	0.943	1.01
100	0.928	1.00
150	1.013	1.09
200	1.007	1.08
250	1.006	1.08
300	1.023	1.10
350	1.113	1.20
400	1.152	1.24

(1): Relative rail deflect (speed): Rail deflect (speed) / Rail deflect (v=0 km/h)

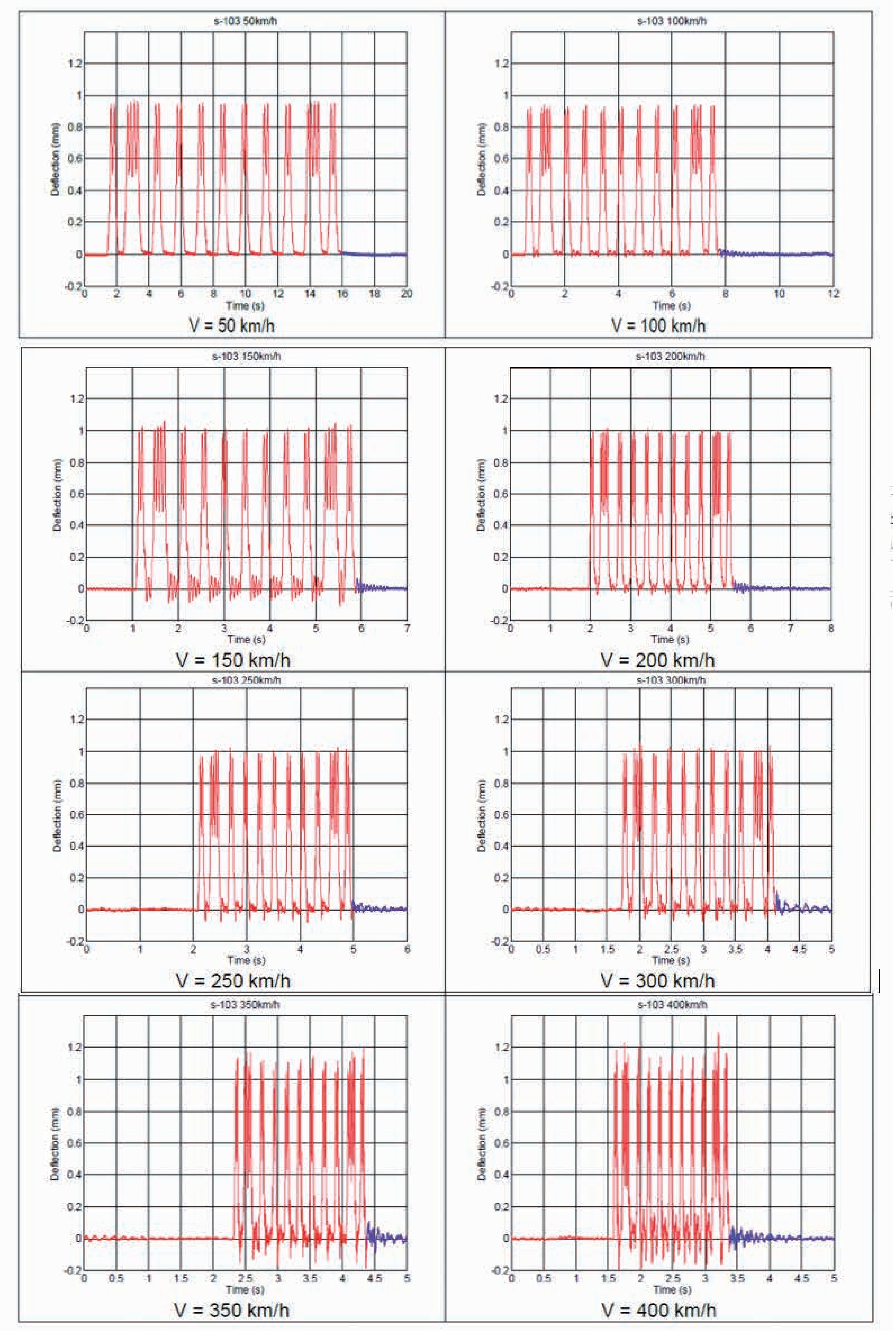


Figure 20. Rail deflection curves obtained for different train speeds in CTB tests

The analysis of the test results shown above makes it possible to highlight the following aspects:

- The tests can be divided in three groups according with the rail deflection obtained, as it can be seen in Figure 21:
  - For tests with speeds below 100 km/h, the rail deflection is below 1 mm.
  - For tests with speeds between 150 and 300 km/h, rail deflection is a bit above 1 mm.
  - For tests with speeds above 300 km/h, rail deflection increases steadily up to 1.15 mm.

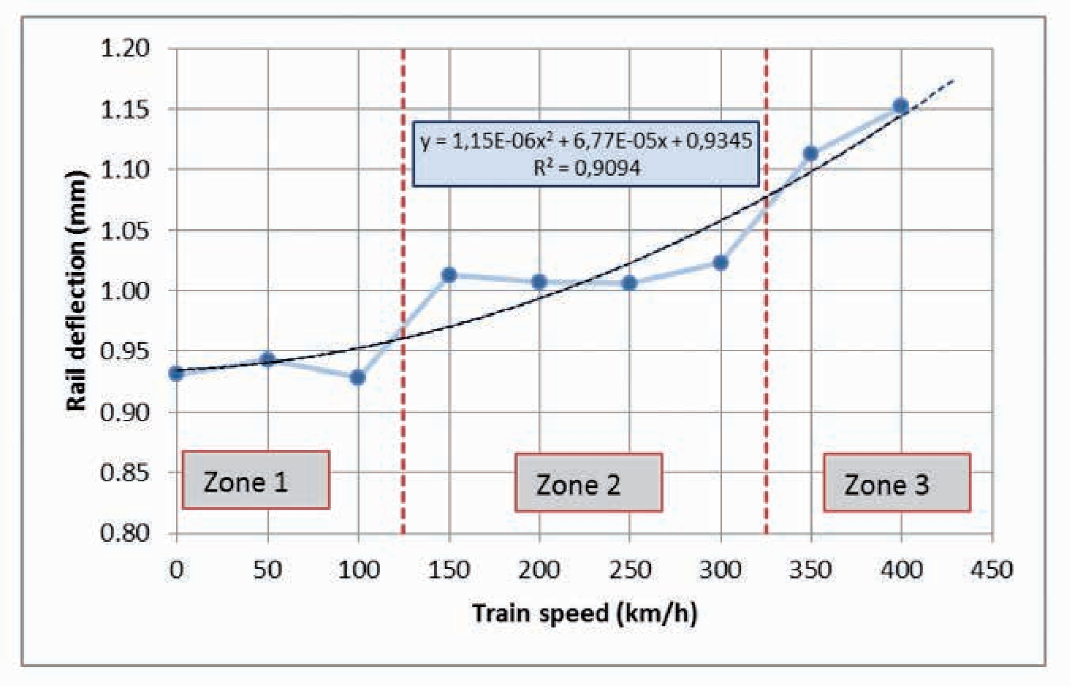


Figure 21. Rail deflection obtained in CTB tests as a function of the train speed

- The increase in rail deflection was from 0.93 to 1.15 mm when train speed varied from 0 to 400 km/h. That supposes an increase of 25%.
- The rail deflections obtained for speeds between 150 and 300 km/h can be considered quite steady, while for 350 and 400 km/h the peak rail deflections are more irregular and oscillate much more than the ones obtained for smaller speeds.
- It can be also observed that the free rail oscillation during the bogie pass-by significantly increases with the speed.

The results can be used to deduce the critical velocity of the 1:1 scale model built in CTB, with the aid of the graph in Figure 22 that relates the increase of rail deflections with the increase of train speed. The curves in Figure 22, for different damping ratios, appear as the solution of the differential equation that gives the vertical deflection of an infinite beam on an elastic foundation when a load is moving on it (Fryba, 1999).

The best fitting of the results is obtained, in this case, for a critical speed of 640 km/h which can be considered accurate enough for a very good railway track, with a hard embankment and hard track bed layers.

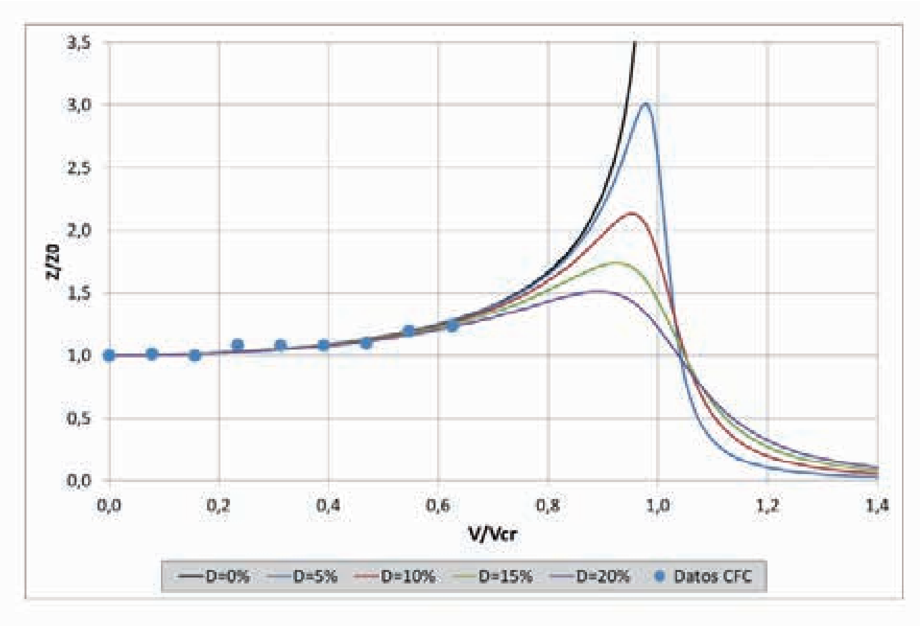


Figure 22. Analysis of the displacements obtained for different train speeds to deduce the critical speed of the 1:1 scale model built in CTB with a hard embankment and hard track bed layers

4.3.3 Rail and sleeper velocities

The tests performed at different speeds also made it possible to record velocities in the rail and in some sleepers by the use of geophones. The values obtained are collected in Table 4 and represented in Figure 23.

Table 4. Velocities obtained in rail and sleepers for different train speeds.

Train speed (km/h)	Outer Rail Amplitude (cm/s)	Inner Rail Amplitude (cm/s)	Outer Sleeper Amplitude (cm/s)	Inner Sleeper Amplitude (cm/s)
50	0.83	0.66	0.36	0.57
100	2.00	1.16	1.11	0.83
150	3.50	2.20	2.00	1.10
200	5.90	5.40	4.40	3.50
250	7.50	7.10	4.50	3.90
300	9.70	11.10	5.60	5.20
350	12.60	12.01	7.50	6.40
400	16.00	14.50	10.40	8.00

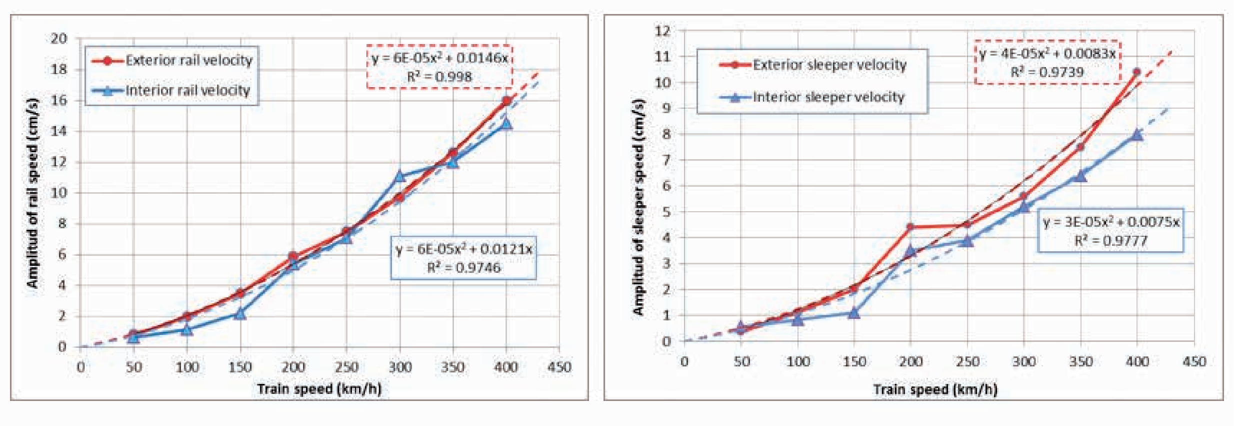


Figure 23. Amplitude of rail (left) and sleeper (right) speed

As it can be seen in Figure 23, the amplitude of velocities obtained in the rail and in the central sleeper increase with the train speed, showing a similar pattern than the increase of rail deflections. Note that the best fitting curves are, in all cases, second degree polynomials.

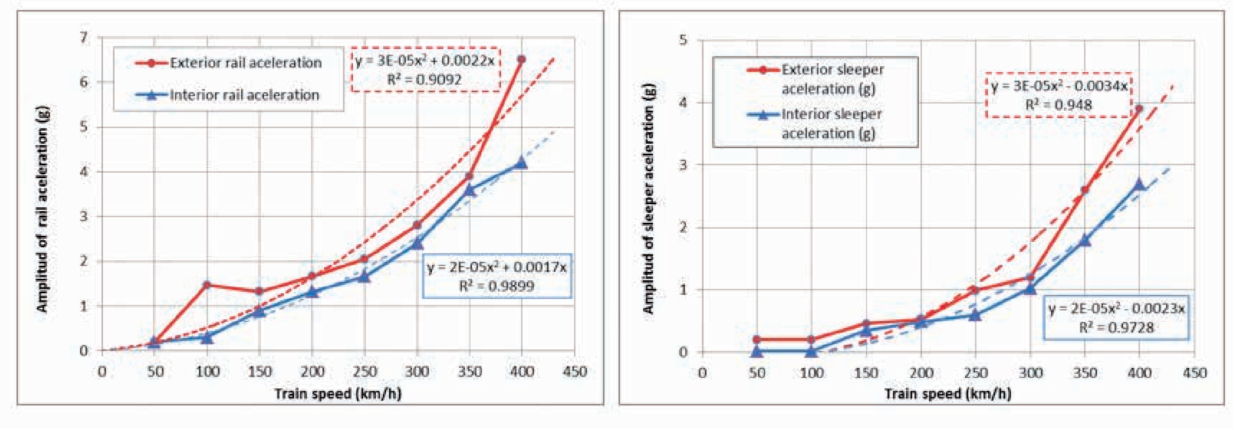
4.3.4 Rail and sleeper accelerations

In the tests performed at different speeds, the accelerations in the rail and in the central sleeper were also recorded with some accelerometers. The values obtained are tabulated in Table 5 and represented in Figure 24.

Table 5. Accelerations obtained in the rail and the central sleeper for different train speeds

Train speed (km/h)	Outer Rail Amplitude (g)	Inner Rail Amplitude (g)	Outer Sleeper Amplitude (g)	Inner Sleeper Amplitude (g)
50	0.20	0.20	0.20	0.02
100	1.47	0.30	0.20	0.02
150	1.32	0.90	0.46	0.35
200	1.66	1.32	0.53	0.49
250	2.05	1.66	0.99	0.60
300	2.80	2.40	1.20	1.03
350	3.90	3.60	2.60	1.80
400	6.50	4.20	3.90	2.70

Figure 24. Amplitude of rail (left) and sleeper (right) acceleration



As shown in Figure 24, the rail and sleeper accelerations have the same trend as deflections and velocities: an increase of the values recorded with an increment of train speed, following a second degree polynomial.

4.4 Tests with new materials: sleepers equipped with USP

The aim of this work was to analyze the behavior of a prototype of sleepers equipped with Under Sleeper Pad (USP) by the performance of both static and quasi – static tests. These tests were carried out as part of RIVAS project (Railway Induced Vibration Abatement Solutions) in the frame of European Union Seventh Framework Program.

4.4.1 Sleeper characteristics

The sleepers used in the tests were 13 mono-bloc concrete pre-stressed units of the B90.2 type. Their average weight was 6.10 kN and they had a length of 2.80 m, while the rest of the dimensions can be seen in Figure 25. The sleepers were provided with 0.1 N/mm3 static bedding modulus under-sleeper pads of the G04 (SLN 1010) type, glued along all the sleeper bottom, as can be seen in Figure 26.

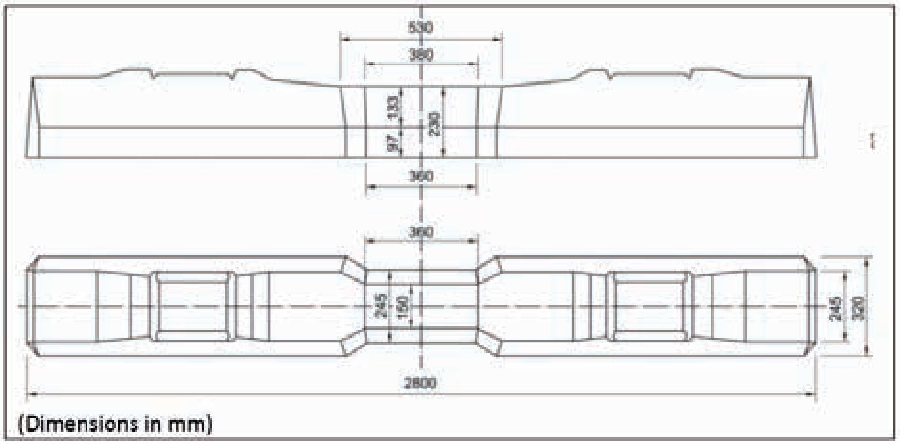


Figure 25. Schematic drawings of the sleeper used in the tests

Figure 26. General aspect of USP installed in the sleepers

4.4.2 Static tests

Four series of three step by step static tests, up to a maximum axle load of 225 kN each, were carried out loading the track subsequently in cross sections A, B and C, as shown in Figure 27 (left). For each cross section, two hydraulic actuators were used simultaneously to load both rails. The time-load history used in the static tests is indicated in Figure 27 (right).

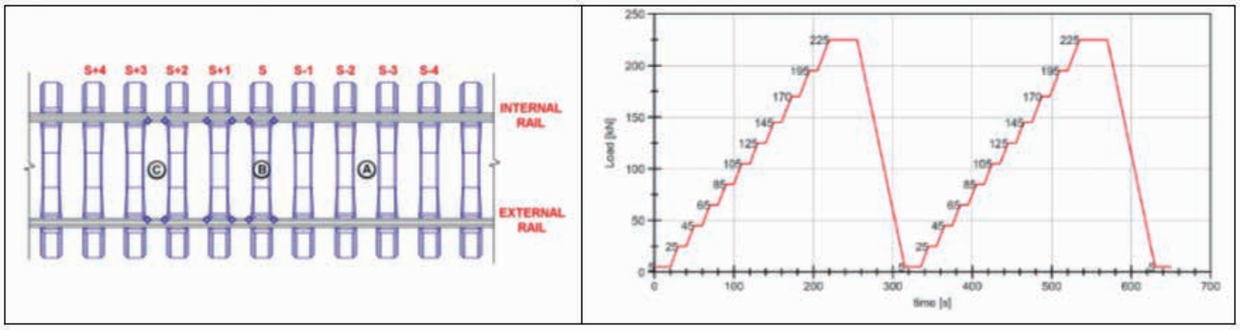


Figure 27. CTB static test cross sections (left) and load time history used in the tests (right)

The static tests were performed to determine the mechanical behavior of the track system with the ballast in the following states:

- BT state: after extended and compacted.
- BTS state: after tamped and stabilized with 70,000 freight vehicle tones.
- B1M state: after consolidated under the pass by of 1M freight vehicle axle loads.
- B2M state: after consolidated under the pass by of 2M freight vehicle axle loads.

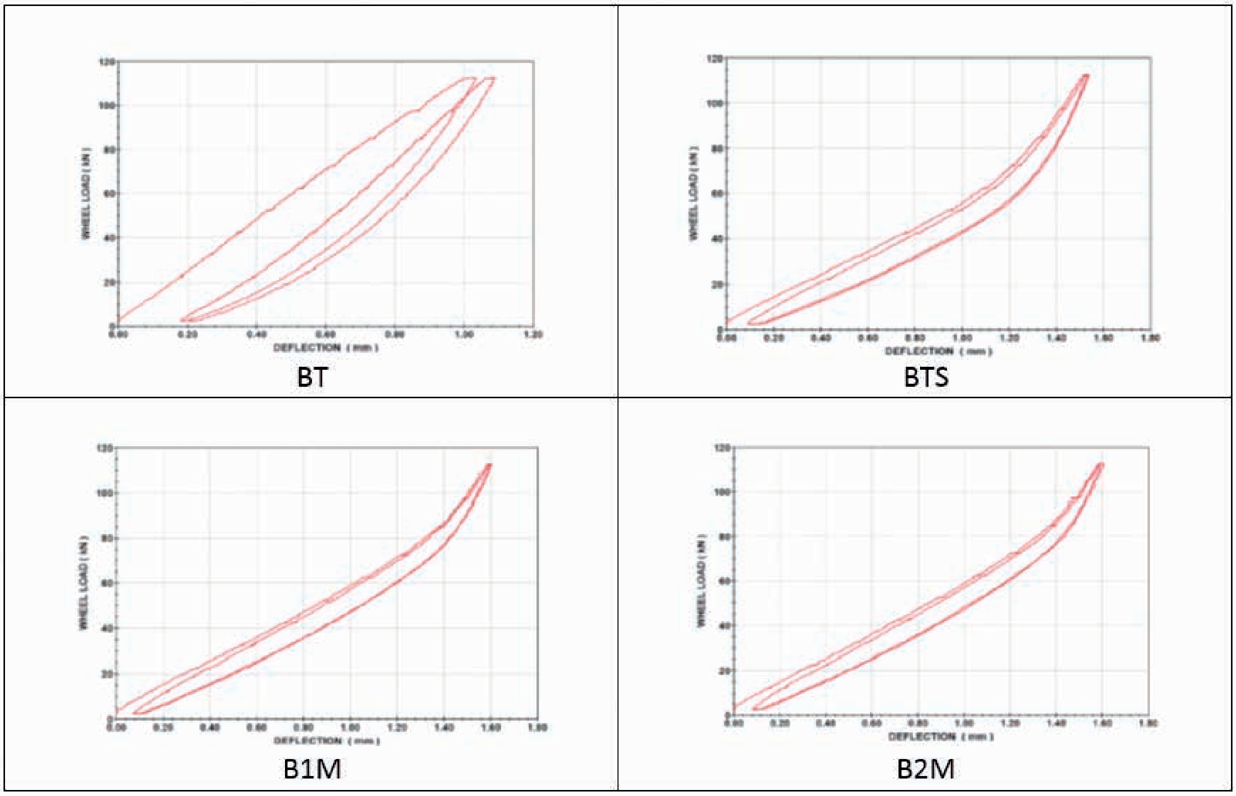
The following sets of sensors were used for each static test set:

- 8 laser beam systems: 4 for cross section B and 2 in each cross sections A and C.
- 16 inductive rail-pad displacement sensors: 8 for each rail.
- 3 inductive sleeper displacement sensors: 1 for each cross section.

The rail load-deflection curves obtained in the static tests carried out in the central cross section (cross section B) are given in Figure 28. On other hand, Figure 29 shows the supporting load/ballast + under sleeper pad compression curves obtained in the static tests carried out. From both set of curves it can be noticed the non-linear behaviour of the track components.

The variations of the track stiffness, along the rails, obtained in the four series of static tests are illustrated in Figure 30, for the different ballast state. In each one of them three values are given for each rail representing the global stiffness values obtained in cross sections A, B and C. Also one single value, corresponding to the average of the three global values, is given for each rail in those figures.

It can be seen that there are significant differences between the different points that reflects the heterogeneities produced during the construction operations in spite of the laboratory conditions and the careful control under those activities were performed in CTB.



BT

BTS

B1M

B2M

Figure 28. Rail load-deflection curves obtained in the static tests

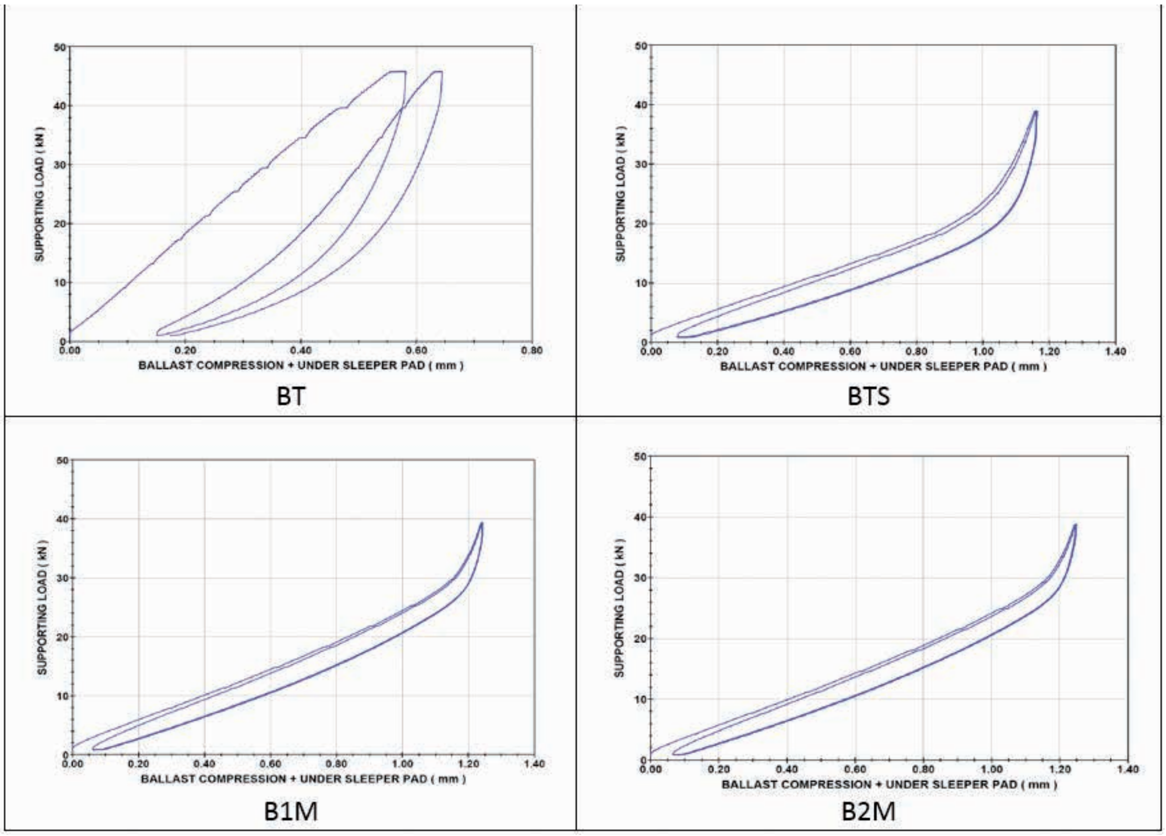


Figure 29. Supporting load/ballast + under sleeper pad compression curves obtained in the static tests

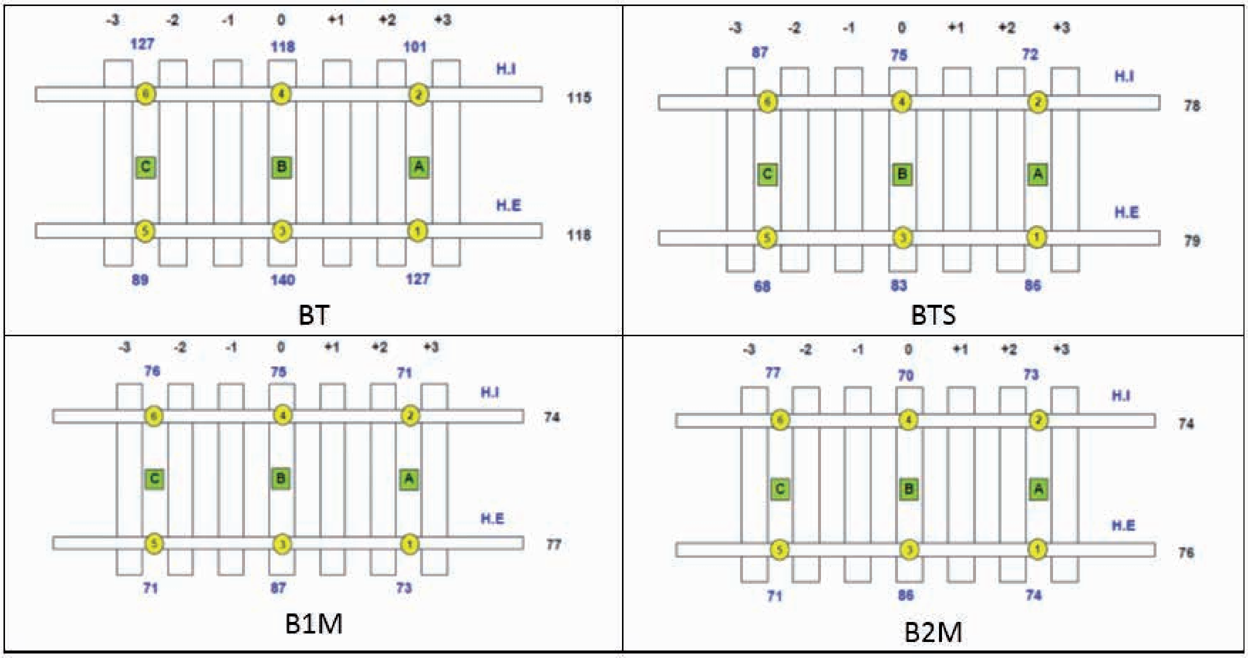


Figure 30. Track stiffness deduced from the static tests for different ballast states

The non-linear behaviour exhibited by both the rail load/deflection curves and the ballast load/compression curves have been quantified through the lower, global and higher load range stiffness values presented in Table 6.

Table 6: Track stiffness values obtained in the second loading cycle at the central cross section

Ballast State	Lower range value 2.5 - 42.5 kN	Global range value 2.5 - 112.5 kN	Higher range value 82.5 - 112.5 kN
BT	104	128	136
BTS	56	78	145
B1M	58	81	123
B2M	60	78	133

On other hand, the vertical displacement measured in the different potentiometers installed in both rails to measure the pad vertical compression are tabulated in Table 7 and represented in Figure 31. Additionally, the rail pad spring constant (kp) is aslo collected in Table 7.

Ballast state	Rail	Rail pad compression (x 10^{-3}mm)									Rail pad spring kp (kN/mm)
		Sleeper									Sleeper
		-4	-3	-2	-1	0	1	2	3	4	*0*
BT	Inner	1	2	8	18	56	26	8	1	1	795
	Outer	2	1	8	33	42	26	8	1	1	1121
BTS	Inner	3	4	5	5	50	17	10	28	2	767
	Outer	3	4	8	32	63	1	11	2	2	628
B1M	Inner	2	16	8	12	64	6	15	10	0	599
	Outer	2	16	9	19	74	17	4	2	0	543
B2M	Inner	4	12	18	17	61	3	22	14	0	617
	Outer	4	12	13	16	74	19	5	2	0	541

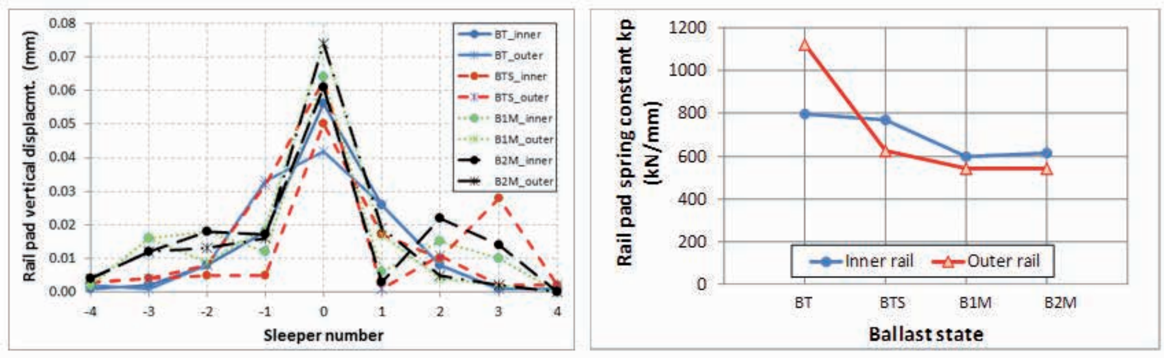


Figure 31. Rail pad vertical settlements measured in the static test performed at the central section for different ballast states (left) and rail pad spring constant (kp) deduced for the central section

The most significant feature of the results is that, for each section, the rail pad settlements, and consequently its spring constant, change during the test although the most apparently affected track component should be the ballast, since it was tamped, stabilized and loaded with 2 M load cycles.

4.4.3 Long lasting quasi - static tests

For the long lasting tests, the three couples of hydraulic cylinders situated at cross sections A, B and C (see Figure 27) were used simultaneously but with a time lag of 45 ms between each couple to simulate the pass by of freight vehicles having 164 axles that apply a load of 225 kN, at 120 km/h, until fulfilling the 2M axle loads.

The evolution of the vertical deflection amplitudes of the rail, the rail-pad and the sleeper, measured at the central cross section along the first million of axle loads is illustrated in Figure 32. It can be seen that those amplitudes have a light decrease once the first 100,000 cycles have been applied, while during the rest of the test they are quite constant.

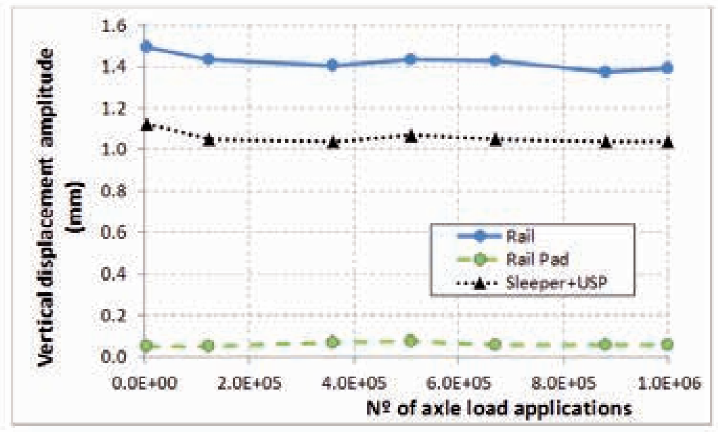


Figure 32. Track, rail pad and sleeper deflection amplitudes at central cross section

Figure 33 shows the track stiffness determined during the test, besides the values obtained in the static tests corresponding to the initial (BTS state) and final (B1M state) stages. It can be seen that the track stiffness is quite constant during all the test although there is a little increase in the outer rail values. On other hand, the static values are quite similar to those obtained in the quasi-static tests.

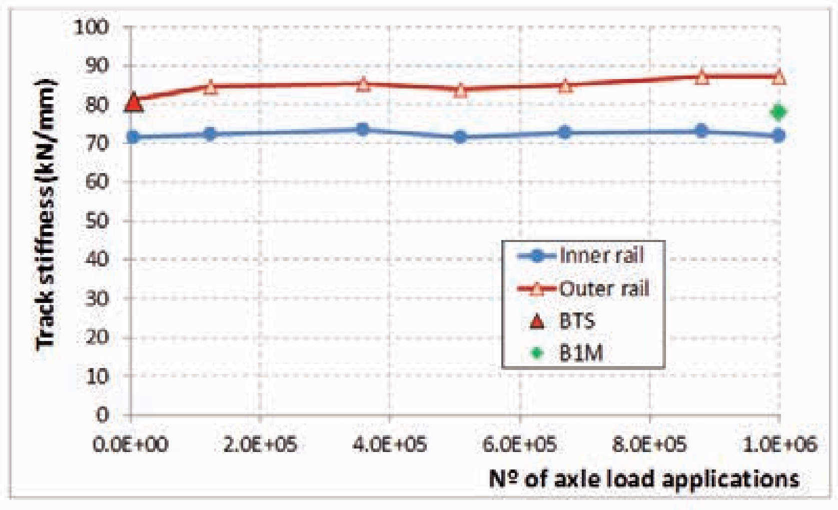


Figure 33. Quasi-static track stiffness values at central cross section obtained along the first million of axle loads

Finally, the evolution of the irreversible settlement experienced by the ballast is shown in Figure 34. This curve can be interpreted as the ballast fatigue curve. It can be seen that there is a great initial settlement around 0.6 mm which roughly represents the half of the total and that, from 250,000 cycles there is a linear increase with a rate of settlement of about 0.22 mm/ 1M axle loads.

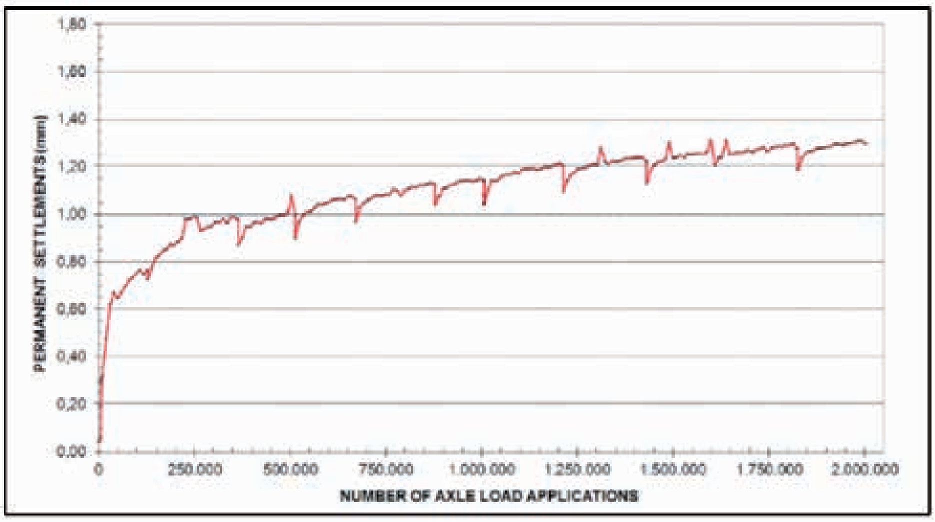


Figure 34. Ballast permanent settlement curve

5. Summary

CEDEX Track Box (CTB) is a 21 m long, 5 m wide and 4 m deep facility whose main objective is to test, at 1:1 scale, complete railway track sections of conventional and high speed lines for passenger, freight and mixed traffics, at speeds up to 450 km/h.

CTB has been used since its construction, 12 years ago to research on different matters of rail way tracks founded on hard soils such as: measurement of track vertical stiffness under different track conditions, determination of track lateral stability, determination and interpretation of ballast permanent settlement curves in fatigue tests, calculation of rail deflections under the pass-by of trains at very high speeds, up to 450 km/h, behaviour of sleepers with USP, optimization of bituminous sub ballast thickness, , behaviour of High Speed Lines subjected to mixed traffic and calibration of 3D numerical models.

Furthermore, it can be also used to study the influence of ground irregularities in the behaviour of slab tracks, to optimize the maintenance works in slab-ballast transition zones or to analyse the ballast degradation in switches or crossings.

The recent deployment of a 3 Hz natural frequency spring system at the base of CTB will allow in the future to assess experimentally the behaviour of ballasted and slab tracks founded on soft soils.

6. References

- CUÉLLAR, V. (2016). Comportamiento dinámico de líneas férreas con balasto bajo cargas verticales. Ingeniería Civil 183, pp. 79-108.
- ESTAIRE, J., CUÉLLAR, V. y CÁMARA, J. L. (2017). Settlement laws of bed layers of a ballast track as determined in 1:1 scale models performed in CEDEX Track Box. Int. Conf. Transportation Geotechnics and Geoecology (to be published)
- ESTAIRE, J., y SANTANA, M. (2017). Large direct shear tests performed with fresh ballast (to be published in Symposium on Railroad Ballast Testing and Properties, ASTM, 2018).
- ESTAIRE, J., CUÉLLAR, V. y SANTANA, M. (2017). Track lateral stability tests performed in CEDEX Track Box and their modelling. Revista Geotecnia, nº 140 – July 2017 (pp. 03-30)
- FRYBA, L. (1999). Vibration of solids and Structures under Moving Loads. Thomas Telford.
- KISH, A. (2011). On the fundamentals of track lateral resistance. AREMA.
- MANZANAS, J., MORENO, J., CUÉLLAR, V., ANDREU, M.A. y NAVARRO, F. (2007). CEDEX´s experimental facility for testing railway tracks. Proc. 14th European Conf. Soil. Mech. Vol 4, 2037-2055.
- RIVAS. Deliverable D3.7 Part A. (2013) Results of laboratory tests for ballasted track mitigation measures. CEDEX Track Box.
- SAMAVEDAM, G., SLUZ, A. y KISH, A. (1999). The effect of realignment on track lateral stability. AREMA

7. Acknowledgements

The authors of the paper want to acknowledge all the people who have worked in CTB during all its working life: Francisco Navarro, Miguel A. Andreu, Juan L. Cámara, Ricardo Fernández, Francisco González, Inés Crespo, Raul Reinoso, Ángel Carballo, Gema Villar. Pedro León, Antonio García, José M. Cerezo y José L. F. Rocha.

High Speed Rail Infrastructure

360.revista de alta velocidad
número 5 - junio - 2018. Pág 219 - 227

New design concepts for High-speed lines and the limits of the ballasted track

Escobar, Adrián
Zamorano, Clara Isabel
Jiménez, Pablo Lorenzo
Escobar, Jorge

Technical University of Madrid[1]

Abstract

The correct decision regarding to what type of track should be used in High-speed lines is one of the topics that has generated more discussion and controversy over the last few years. In addition, this decision is very influenced by the tradition, philosophy and design principles of the different countries. This fact is fundamentally due to the multitude of constraints and influencing factors that define as most appropriate, the use of a ballasted track against a ballastless track, or vice versa. However, the present research intends to look beyond and review the main High-speed projects of greater relevance and transcendence that are being or will be realized in the next years, around the world, to observe if some of the two typologies is presenting a greater implementation, and the main factors that are motivating this decision. In addition, the new conditions that are appearing in some countries will be analysed, related to the demand for greater performance of design speed and design life, as well as the influence that these new demands will have on the decisions related to the type of superstructure to be installed. Finally, a comparison of all these different design philosophies will be carried out, in relation to the Spanish High-speed model, in order to generate a brief discussion of the main differences and similarities found. The final objective of the present study is to analyse current trends in superstructure design and to help the debate on the type of track to be used in the High-speed lines that will be developed in the near future.

Keywords: superstructure comparison, High-speed trends, High-speed rail, ballasted track, ballastless track .

---

[1] *Escobar, Adrián. Technical University of Madrid. Email: adrian.escobar.pastor@alumnos.upm.es (Corresponding author)*
Zamorano, Clara Isabel. Technical University of Madrid. Email: clara.zamorano@upm.es
Jiménez, Pablo Lorenzo. European University of Madrid. Email: pablolorenzo.jimenez@universidadeuropea.es
Escobar, Jorge. Technical University of Madrid. Email: j.escobar@alumnos.upm.es

1. Introduction

High-speed railway systems have been widely accepted in modern society, positioning themselves as one of the most important means of transport to guarantee the territorial structure. These systems present different conceptions or design philosophies regarding the superstructure subsystem, which has led to the generation of different infrastructure models characterized by a predominance of ballasted track or by a predominance of ballastless track.

Although historically the first lines had only the typology of ballasted track, nowadays, it is possible to find regions in which the implementation of the slab track is majority. The choice of typology will depend on technical factors such as relief, environmental conditions, design speed, availability of materials, etc.; and economic factors mainly characterized by the financing capacity and operating and maintenance costs.

However, in the last years, different High-speed projects have been launched in different parts of the world, where new technological challenges are being pursued, among them, the achievement of a maximum speed of around 400 km/h or, the fact of leading a High-speed line through an environment with extreme meteorological conditions. The main objective of this article is to make a deep reflection on the typology of railway superstructure that will be used to face these new technological challenges and to discuss the main limitations that can present the typology of ballasted track regarding to its use in the future High-speed lines.

With this aim, a brief historical review of the main track typologies used in the High-speed lines will be carried out, then the main limitations of the ballasted track will be described, in the third place, the main High-speed projects that are currently being developed around the world will be shown and, finally, it will be analyzed if there is any trend for the implementation of a typology of superstructure in particular.

2. State of the art

2.1 Historical review

With a length of 515 kilometers, the first High-speed line came into operation in Japan in 1964 from Tokyo Central to Shin Osaka, as shown in Li et al. (2016). This line has currently a maximum operating speed that reaches 285 km/h and has a track superstructure belonging to the ballasted track typology, as per UIC (2015). However, it was not until 1975, with the completion of the Sanyo line, when the first High-speed line with a ballastless superstructure was implemented in Japan, as explained by Yokoyama (2010), which currently has a maximum operating speed that reaches the 300 km/h, as per UIC (2016). This line was the first Highspeed line in the world that introduced the use of the slab track in a percentage around 50% of the total length of the infrastructure. The Japanese High-speed network had a total length of 1,244 km of slab track in the year 2001, as shown in Ando et al. (2001).

However, before 1964 there were already several attempts by countries such as Germany, France or the United States to launch High-speed Rail projects, which eventually became separated historical records, as described by Hughes (2015).

In the case of Germany, in 1903, 200 km/h was exceeded with an electrically powered train between Marienfeld and Zossen, over a ballasted track, as shown in Ebeling (2005). Nevertheless, it will not be until 1991 when the regular High-speed service will finally be implemented with the ICE train, using a superstructure of ballasted track, on the lines from Mannheim to Stuttgart and from Hannover to Würzburg, with a maximum speed of 280 km/h, as explained by Beck (2006). Although, the first ballastless track system is developed in 1972 at the Rheda Station, it will not be until 1998 with the construction of the line between Hannover and Berlin, when it starts to use the slab track on the German High-speed network, as shown in Sugrue (2013).

France set its first world record in 1955 reaching a speed of 331 km/h, as described by Givoni (2007). However, it will be necessary to wait until 1981 to witness the inauguration of the TGV's first High-speed line between Paris and Lyon, reaching a speed of 270 km/h on a ballasted track, as shown in Kim (2000).

On the other hand, in 1992, the first High-speed line between Madrid and Seville was inaugurated in Spain, with a maximum speed of 300 km/h over a ballasted track, as described by Gutierrez Puebla (2004). This line reduced the travel time by train between Madrid and Ciudad Real from 160 minutes to 50, generating an increase in the mobility between the two cities from 310,161 passengers in 1990 to 740,972 in 2000, as shown in Ureña et al. (2005).

Both the Spanish High-speed model and the French model have developed their entire network under the typology of ballasted track for a maximum operating speed of 300 km/h and 320 km/h, respectively, as shown in UIC (2017).

2.2 Multicriteria assessment to choose the superstructure typology

As can be seen in the previous paragraphs, there are very different philosophies depending on each country, with respect to the type of superstructure to be used in High-speed lines. Therefore, this topic has been widely discussed by authors such as Esveld (1999) and Esveld et al. (2003), who explained the advantages of the use of the slab track against the ballasted track. Koriath et al. (2003) established the main criteria for the objective selection of the typology of the railway superstructure on the German rail network. Then, in 2008 the studies carried out by CEDEX (2008) showed the medium and long-term behavior of both superstructure typologies.

Ren et al. (2009) performed an economic analysis of the life cycle of a ballasted track and compared it with a slab track. Kollo et al. (2015) developed also a detailed technical-economic comparison between both types of superstructure. Finally, Giunta et al. (2017) made a comparison between the use of slab track or ballasted track in High-speed lines from the point of view of the overall study of the life cycle costs of both solutions.

2.3 Limits of implementation of a ballasted track in a High-speed line

On the other hand, some authors have highlighted the main problems that arise with the use of ballasted track with High-speed traffic. López Pita (2001) analyzed the acceleration experienced by the ballast deterioration when subjected to High-speed traffic. Riessberger (2006) developed a study to defend the use in High-speed lines of the ballasted track but applying certain improvements or modifications to the traditional design.

Al-Shaer et al. (2008) developed a study on the dynamic behavior of the ballasted track and its settlements, in High-speed lines, by using a physical scale model. Nguyen et al. (2011) made a model to study the mechanisms of deterioration of High-speed lines with ballasted track.

Finally, Giannakos et al. (2012) outlines the main requirements to be met by ballast for its use in High-speed lines.

However, slab track systems also have important disadvantages such as the impossibility of great geometrical corrections after track construction, the very tight limitation regarding to earthwork settlements and the higher installation costs, as shown in Blanco-Lorenzo et al. (2011).

3. Experimental analysis

3.1 Main ballasted track issues

Although ballast has been widely used for railway infrastructures throughout the world, its use in High-speed lines will present important limitations that cannot be ignored in the case of lines with speeds above the 300 km/h, as shown in González-Cancelas et al. (2012). Some of the main effects that ballast can suffer if placed in lines that have maximum speeds above this limit are: a higher speed of deterioration that will significantly reduce the useful life of the ballast and that will modify its elastic and resistant properties; a proliferation of differential settlements above the normally accepted limits; and, in the third place, second-order effects such as ballast flight, which may cause significant damage to the rolling stock and imply a risk to the safety of the circulation.

With respect to the two main effects, the accelerated ageing of the ballast layer and the appearance of higher differential settlements in the track, it is worth mentioning that both phenomena are connected and both follow a direct relation with the train operating speed. Both deterioration mechanisms will follow a law of exponential acceleration with the time due to the feedback of the mechanisms that generate them.

Firstly, because high speeds produce higher dynamic loads, an important impact will be generated in the ballast increasing its stiffness and causing the appearance of the first settlements in the track. Due to these two factors, during the next load cycles that are generated when passing the trains, the dynamic loads generated by them will be increased, due to the increases of stiffness and the increase of the imperfections existing in the track due to the settlements, and there will again be a greater stiffening of the ballast layer and larger settlements in the track. This mechanism will continue to occur at a higher speed until the track failure occurs.

In conclusion, in two tracks that present the same characteristics and the same number of load cycles, it is observed that the one in which the trains circulate at a higher speed, bigger settlements and a greater deterioration of the ballast will be produced, as shown in Nguyen et al. (2011).

3.2 Analysis of rail superstructure in High-speed lines around the world

For this reason, in spite of the multitude of economic and technical factors that can define the use of a track typology or another in High-speed lines, it would be necessary to establish an approximate limit from which a ballasted track can constitute a technical restriction for a certain line can reach the maximum speeds for which it was designed, without prejudice to reduce the useful life of the infrastructure.

This fact implies that due to the high proliferation of High-speed lines being built all over the world, which are continuously seeking speed limits higher than those established as conventional, such as 300 to 320 km/h, the implementation of slab track systems will be increased, because they allow to guarantee a high performance and continuous behavior of the superstructure when the train reach the maximum speeds.

An extract of the main High-speed lines that are in operation and the type of track that they present will be made, in order to observe the possible trends that are occurring, according to the data provided by Yokoyama (2010), SSF (2010), Sugrue (2013) and UIC (2017).

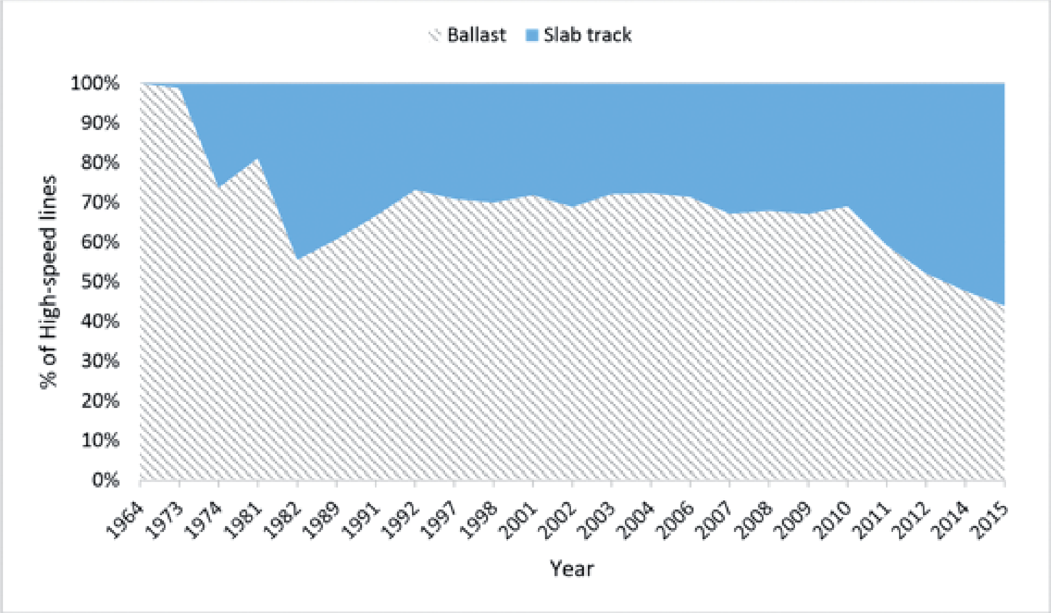


Figure 1. Historical evolution of ballasted track vs. slab track

From the results shown in figure 1 it is observed a significant increase in the use of the slab track in High-speed lines during the last 5 years. This increase is mainly due to the fast expansion of High Speed lines in China, which is using slab track in the main lines of its Highspeed network and has developed more than 4,000 km during the last years. The use of the slab track is, among other reasons, due to the high design speed that has been set in values between 350 to 380 km/h for some of the main lines. This is the case of the Beijing - Wuhan line with more than 1,200 km of slab track designed for a maximum speed of 380 km/h, the case of the Beijing - Schanghai line with 1,318 km of slab track, also, for a maximum speed of 380 km/h or, the case of the lines between Changsa - Hangzhou with 840 km and Heifei - Fuzhou with 810 km of slab track respectively, both designed for maximum speeds in the range of 350 km/h.

Secondly, this growth in the installation of ballastless track includes countries that have traditionally used it, such as Germany or Japan, which continue to expand their already consolidated High-speed networks. In the case of Germany, the construction of the Leipzig / Halle - Erfurt line, in which more than 123 km of slab track was installed for a speed of 300 km/h and, in the case of Japan, construction of the Nagano - Kanazawa line, with the installation of 228 km of ballastless track for a maximum speed of 260 km/h.

3.3 Spanish High-speed model

By contrast, Spain and France continue to have the ballasted track as the base of their Highspeed network, using the slab track in specific cases, such as long tunnels. In Spain, this philosophy is motivated by the abrupt orography that presents the Spanish territory that forces the construction of embankments with a height higher than the 10 meters.

The settlement caused by rail traffic is completed over the course of a few months, accounting for approximately 0.1% of the embankment height, as estimated by Pérez-Romero et al. (2016). Hence, the requirements to build slab track generally set the maximum embankment height around 10 meters, as shown in DB Netze AG (2014), in order to control the postconstructive settlement because the ballastless track allows only a maximum settlement in the range of 30 to 50 mm.

Due to this fact, the implementation of the slab track in Spain would require a significant reduction of the height of the longitudinal profile of the infrastructure, in order to avoid having such high embankments, generating a massive increase of the length of tunnels to be constructed in all the High-speed network, with its consequent increase in the total construction costs of these lines.

In addition, the reduction of height in the longitudinal profile would imply in the second instance a greater concern with other factors, such as the drainage systems to be built, this being another of the most important technical issues to be solved, since Spain has a large amount of areas where total precipitation is concentrated in a few days throughout the year.

On the other hand, because the Spanish High-speed model does not reach operating speeds above 300 km/h currently, it has not become indispensable to think about the possibility of installing a slab track on the main sections of the network.

In conclusion, these important constraints together with the budgetary constraints justify the use of the ballasted track instead of the slab track in the Spanish High-speed model, with 2,938 km of length.

3.4 Relationship between the rail superstructure and the speed

It can be observed in figure 2 a relation of the type of track used regarding to the maximum design speed. For the realization of this figure, only the projects developed in the last 10 years have been used.

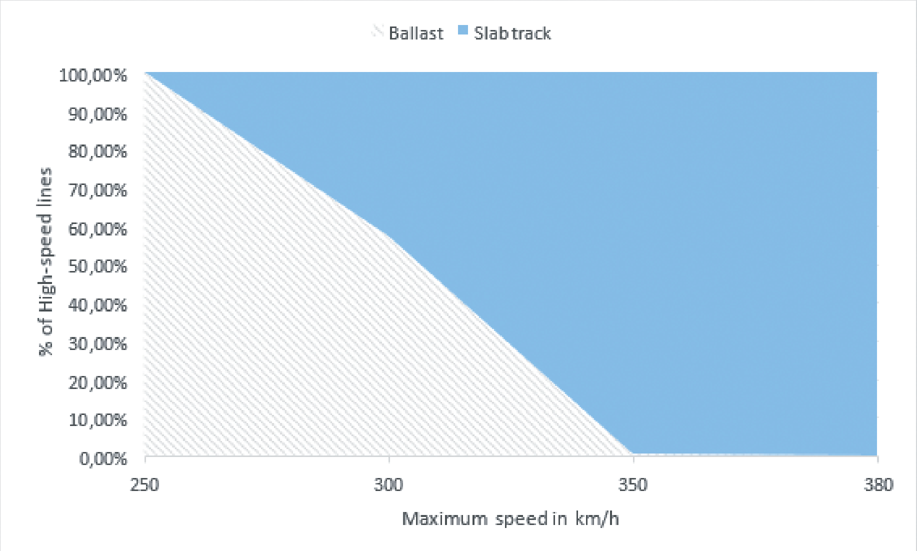


Figure 2. Track typology depending on the maximum design speed

As shown in figure 2, there is a relationship between maximum speed and type of track to be installed. Therefore, in the first instance it could be assumed that as long as the trend in the construction of High-speed railway is to increase the maximum speed of the line, the type of track that will predominantly be used will be that of slab track.

3.5 High-speed trends around the world

Therefore, it is observed that there is a clear tendency to justify the use of one type of track against the other, for High-speed projects. In addition, in recent years the implementation of High-speed networks has been accelerated around a multitude of countries throughout the world, generating a greater technical demand in the design of the new lines and searching for new challenges that allow improving the existing service, as shown in Campos et al. (2009).

In the case of the United Kingdom, this fact has clearly been reflected in the design requirements established for the development of the HS2 project, which will provide a Highspeed line connection among the cities of London, the West Midlands, Leeds and Manchester, as explained by Greening (2012). These requirements set the future maximum design speed at 400 km/h, as shown in Cornet et al. (2017), and 120 years of infrastructure design service life. Therefore, the planned railway superstructure will be of slab track type due to the important condition of the maximum design speed of the project.

On the other hand, in the case of Saudi Arabia, with the project of the Haramain High-speed Line, there have been important problems such as the fight against the desert sand, as shown in Sesma et al. (2012), and resistance to the extreme temperatures, as explained by Durand et al. (2012). The line has been designed for a top speed of 320 km/h, with 758.6 km of ballasted track and 128.88 km of slab track. This decision regarding the track typology is completely reasonable because of speed and country factor. According to the data shown above the design speed of this project is in the range in which there is a mixed use of track typologies and, on the other hand, Spain, which is one of the countries above defined as main user of ballasted track in its High-speed lines, is the designer of the railway superstructure.

In the case of Sweden, the East Link project, called Ostlanken in Swedish, is being developed for a speed of 320 km/h, and it is expected that a slab track type will be implemented, as per Trafikverket (2016, January). This decision is taking into account the two main factors again. In first place, the design speed is in the range of using slab track as shown figure 2 and the companies which are designing most of the track sections are from Germany, which is one of the countries above defined as user of slab track systems in its High-speed lines.

Finally, in the case of the United States, the High-speed line of California is still under development and has been designed to reach top speeds of 350 km/h in the section between San Jose and Los Angeles. The technical specifications leave open the door to the application of slab track or ballasted track, although the latest reports of cost estimation suggest that it will be used almost exclusively the ballasted track, reserving the slab track only for sections of viaduct or tunnel.

4. Conclusions

Throughout the article a historical journey has been made on the different types of railway superstructure installed since the appearance of the first High-speed line and have been described two of the main factors that influence it: the maximum speed and the tradition of design of the installer.

As it has been observed in the previous sections, in general, the typology of slab track is experiencing a greater implementation in the High-speed lines that are being constructed at the moment. This fact is mainly based on that the designs are being made for speeds higher than the current ones and trying to reach the limit of 400 km/h. However, this trend has important exceptions, such as the Haramain line or the California High-speed line that has installed ballasted tracks, reserving the slab track only for those sections where it is specifically needed. In the case of California, only in the areas of greater seismicity and in the sections of tunnel or viaduct, and in the case of the project of Haramain for those section that need a greater protection against the attack of the desert sand.

Finally, it has been shown that the requirements regarding the design of High-speed rail are changing to be more restrictive, taking into account not only the achievement of higher speeds, but the achievement of longer infrastructure service life.

5. References

- AL-SHAER, A., DUHAMEL, D., SAB, K., FORÊT, G., and SCHMITT, L. (2008). Experimental settlement and dynamic behavior of a portion of ballasted railway track under high speed trains. Journal of Sound and Vibration, Elsevier, 2008, Vol.316, pp.211-233
- ANDO, K., SUNAGA, M., AOKI, H., and HAGA, O., (2001). Development of Slab Tracks for Hokuriku Shinkansen Line. Quarterly Report of RTRI Vol.42 (1). pp 35-41
- BLANCO-LORENZO, J., SANTAMARIA, J., VADILLO, E. G., OYARZABAL, O. (2011). Dynamic comparison of different types of slab track and ballasted track using a flexible track model. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 225(6), 574-592.
- BECK, A., (2006). Railway Infrastructure in Germany. DB AG. 2006. (Non-published report).
- CAMPOS, J. and DE RUS, G., (2009). Some stylized facts about high-speed rail: A review of HSR experiences around the world. Transport Policy, Elsevier, vol. 16(1), pp 19-28
- CEDEX (2008). Estudio del comportamiento a medio y largo plazo de las estructuras ferroviarias de balasto y placa. (SP 5-5.1 pp. 38-55 and SP 5-5.2 pp. 13-51).
- CORNET, Y., DUDLEY, G., and BANISTER, D. (2017). High Speed Rail: Implications for carbon emissions and biodiversity. Case Studies on Transport Policy.
- DB NETZE AG (2014). RIL 836 Erdbauwerke und geotechnische Bauwerke planen, bauen und instand halten [Earthworks and geotechnical structures design, construction and maintainance]. Code. Frankfurt am Main, Germany: DB Netze AG. (530)
- DURAND, S., BILLAUDEAU, E., LUBINEAU, G., and CHAPELLE, H., (2012). Design analysis of GRP panels for the roof of the Haramain high speed railway Jeddah station. In Proceedings of the 10th international conference on sandwich structures (ICSS-10). Nantes, France.
- EBELING, K. (2005). High Speed Railway in Germany. Japan Railway & Transport Review, Vol.40.
- ESVELD, C. (1999). Slab Track: A Competitive Solution. Faculty of Civil Engineering, Section of Roads & Railways, Delft University of Technology, the Netherlands.
- ESVELD, C. and MARKINE, V.L. (2003). Developments in High-Speed Track Design. Short paper Proceedings of International Association for Bridge and Structural Engineering (IABSE) Symposium on Structures for High-Speed Railway Transport, Antwerp, Belgium, August 27-29, 2003. pp. 14-15
- GIANNAKOS, K., LOIZOS A., and PLATI, C., (2012). Railway Ballast Requirements for High Speed and Heavy Haul Lines: Hardness, Fouling, Life Cycle. 2nd International Conference on Transportation Geotechnics, Hokkaido, Japan.
- GIVONI, M., (2006). Development and Impact of the Modern High Speed Train: A Review. Transport Reviews, Vol. 26 (5), pp. 593-611.
- GIUNTA, M. and PRATICÒ, F. (2017). Design and maintenance of high-speed rail tracks: A comparison between ballasted and ballast-less solutions based on life cycle cost analysis. International congress on transport infrastructure and systems, Rome.
- GONZÁLEZ-CANCELAS, N., SOLER-FLORES, F., CAMARERO ORIVE, A., LÓPEZ ANSORENA, Í. (2012). Treatment of Outliers to Study Railway Vibrations Transmission. Ingeniería y Ciencia, 8(16), 191-219.

- GREENING, J. (2012). Britain to have new national high-speed rail network. Department for Transport. 10 January 2012. www.dft.gov.uk./news/pressreleases/dft-press-20120110/
- GUTIERREZ PUEBLA, J. (2004). El tren de alta velocidad y sus efectos espaciales. Investigaciones Regionales, (5).
- HUGHES, M., (2015). The Second Age of Rail: A History of High Speed Trains. Book, History Press Limited, 2015.
- KIM, K.S., (2000). High-speed rail developments and spatial restructuring: A case study of the Capital region in South Korea. Cities, Vol. 17(4). pp 251-262
- KOLLO, S.A., PUSKÁS, A., and KOLLO, G., (2015). Ballasted Track versus Ballastless Track. Key Engineering Materials, Vol. 660. pp 219-224
- KORIATH, H., HAMPRECHT, A., HUESMANN, H., and ABLINGER, P., (2003). Bringing Objectivity into System Decisions between Ballasted Track and Slab Track at Deutsche Bahn. Railway Technical Review, 2-3, pp 9-15
- LI, X., HUANG, B., LI, R., and ZHANG, Y., (2016). Exploring the impact of high speed railways on the spatial redistribution of economic activities - Yangtze River Delta urban agglomeration as a case study. Journal of Transport Geography, Vol. 57, pp 194-206
- LÓPEZ PITA, A. (2001). Ballast vibration makes new designs for high-speed lines advisable. In Proceedings of the 5th World Congress on Railways Research, Cologne, Germany.
- NGUYEN, K., GOICOLEA, J.M., and GALBADÓN, F. (2011). Dynamic effect of high speed railway traffic loads on the ballast track settlement. Congresso de Métodos Numéricos em Engenharia, Coimbra, Portugal
- PÉREZ-ROMERO, J., CIANTIA, M. O., ARROYO, M., VAUNAT, J. (2016). Impermeable Membranes for Slab-Track Settlement Mitigation. Environmental Geotechnics, Special Issue in Soil-Atmosphere Interaction. Inst. of Civil Engineers, London.
- REN, J., LECHNER, B., and LIU, X., (2009). Economical evaluating the ballastless railway track with consideration of LCC viewpoint. Eighth International Conference of Chinese Logistics and Transportation Professionals (ICCLTP).
- RIESSBERGER, A. (2006). Ballast track for high speeds. Proceedings "Tracks for HighSpeed Railways", pp 23-44, Porto, Portugal, October 2006
- SESMA, J., LOMBARDERO, M., and RODRÍGUEZ DE LEMA, P. (2012). Sand and wind: an outline of the study of aeolian action on infrastructure with reference to Haramain High Speed Railway, Makkah– Al-Madinah. Revista de Obras Públicas, Vol. 159(3537). pp 7-36.
- SSF INGENIEURE (2010). Ballastless track on high-speed lines. Company Report. 2010.
- SUGRUE, W. (2013). Permanent Way for High Speed Lines. UIC, 9th Training on High Speed Systems, Paris.
- TRAFIKVERKET(2016, January).Retrieved from Trafikverket.se: http://www.trafikverket.se/nara-dig/Sodermanland/projekt-isodermanlandslan/Ostlanken/Om-projektet-Ostlanken-/
- UIC (2015). High Speed Rail. Fast Track to Sustainable mobility. UIC Brochure. Paris.
- UIC (2016). High Speed System in the World. UIC Presentation. Madrid 2016.
- UIC (2017). High Speed Lines in the World. UIC Report. UIC Passenger Department.
- UREÑA, J. M., MENÉNDEZ, J. M., GUIRAO, B., ESCOBEDO, F., RODRÍGUEZ, F., CORONADO, J. M., RIBALAYGUA, C., RIVAS, A. & MARTÍNEZ, Á. (2005). Alta velocidad ferroviaria e integración metropolitana en España: el caso de Ciudad Real y Puertollano. EURE (Santiago), 31(92), 87-104.
- YOKOYAMA, A. (2010). Infrastructure for high speed lines in Japan. International Practicum on Implementing High-Speed Rail in the United States, APTA and UIC.

High Speed Rail Infrastructure

360.revista de alta velocidad
número 5 - junio - 2018. Pág 229 - 235

Stabilization techniques in railway track maintenance

López-Bachiller Fernández, Miguel
Sampedro Rodríguez, Ángel
Díaz Minguela, Jesús

[1]
ANTER
Alfonso X el Sabio School
IECA

Abstract

Stabilizing or recycling soils has been a solution when you have problems with the quality or the behaviour of the built platform. It is an easy way to repair, quick and efficient, and there is in the market enough number of companies that can do this work with a big experience repairing soil platforms.

Keywords: punctuality, stabilizing, recycling, efficiency, quick repair

---

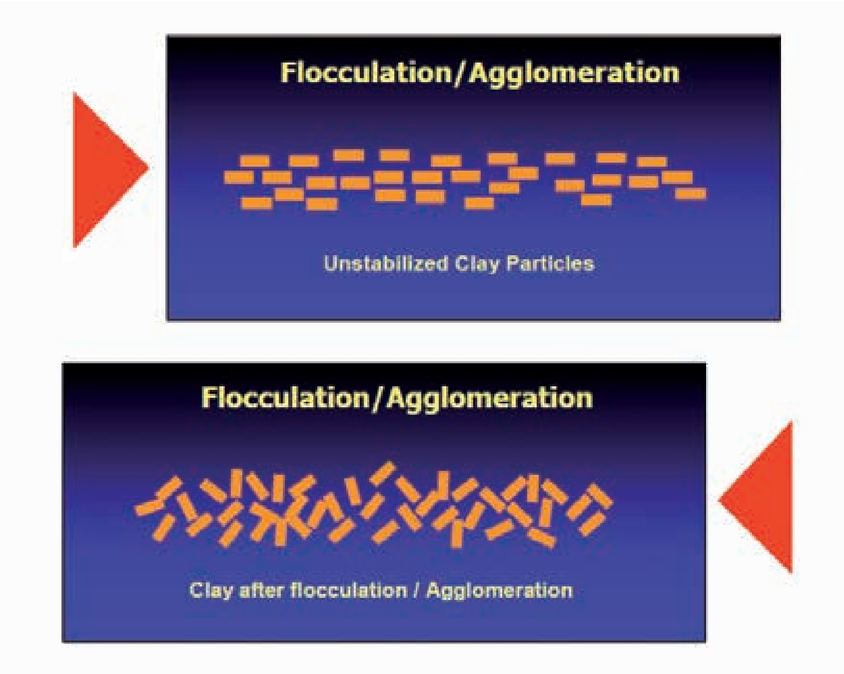
[1] *López-Bachiller Fernández, Miguel. ANTER, Spanish technical association for stabilizing and recycling platforms. miguel@firmesecologicossoltec.com. (Corresponding author)*
Sampedro Rodríguez, Ángel. Alfonso X el Sabio School. sampedro@uax.es. (Corresponding author)
Diaz Minguela, Jesús. IECA. jdiaz@ieca.es

1. Introduction

Much of the success of the high speed train is due to its reliability, along with a large number of frequencies in the lines that are already exploited. Renfe announcement since the beginning of its operation in Spain on the refund of the ticket if the delay is more than a quarter of an hour breaks radically with the idea that associated with the carriage of passengers by rail with delays and uncertainty of the time of arrival and at this time the AVE (Alta Velocidad Española, Spanish High Speed) is associated with punctuality and safety arrived at the scheduled time, well above the air transport.

This is possible thanks to a minimal disruption in maintenance, so it highlights the enormous importance of having an infrastructure in which the maintenance, although essential, is never reason to cause delays in the operation of the track.

Using more reliable building materials than the merely natural, together with maintenance carried out with these same materials facilitates this work.



Flocculation/Agglomeration

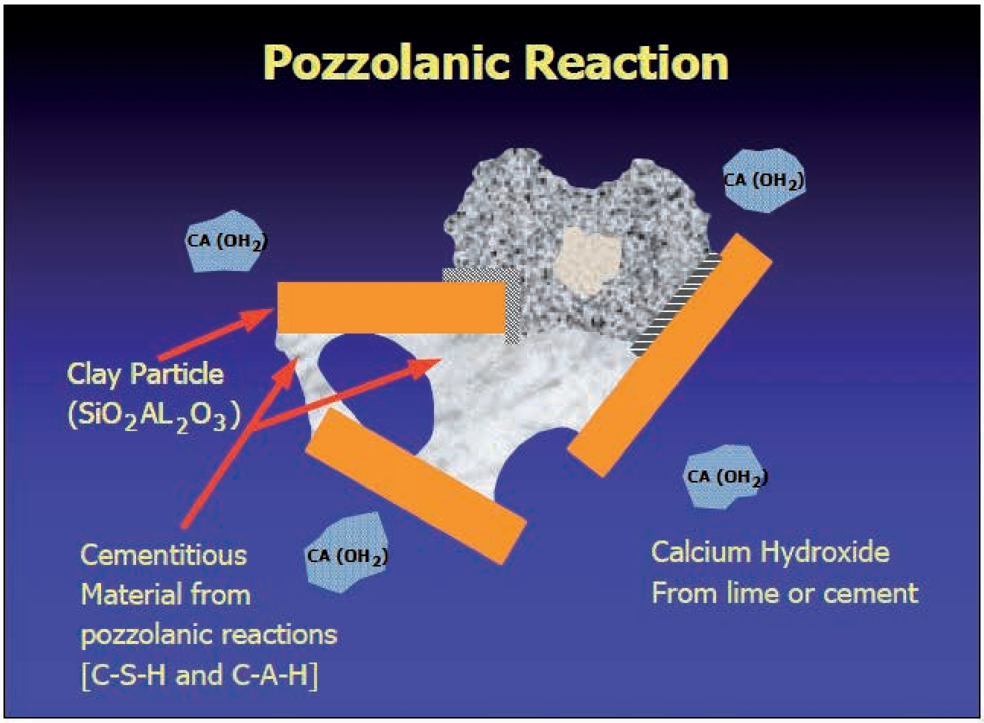
Unstabilized Clay Particles

Flocculation/Agglomeration

Clay after flocculation / Agglomeration

We currently have new forms of reparation of the platform of the sections on service in cases which detects the presence of materials not suitable to the function for which were intended, transforming them through the process of stabilization, which gives an immediate improvement, improving even more their characteristics as time goes on, without having to take them out and replace them with others in the short time that we use to have for repairs.

In Spain we have a group of companies with a park of machinery with capacity to stabilize layers of 0.5 m depth at one time without having to remove them, which allows a great capacity for repair of damaged layers. These powerful machines are normally accompanied by vibratory rollers compactor can densify these thicknesses of layer with guarantees, capturing dynamic plate module reached through compactometers at the same time and keeping these records in easily readable files which can be used to observe the evolution of the affected area with the passage of time.



Pozzolanic Reaction

Clay Particle
$(SiO_2AL_2O_3)$

Cementitious
Material from
pozzolanic reactions
[C-S-H and C-A-H]

Calcium Hydroxide
From lime or cement

You can reach rigidities of the platform depending on the amount and the type of cement used so that you get an immediate commissioning. Even if it were essential to achieve tremendous rigidity in the first hours of operation, also setting accelerators may be used.

In the works in which by the nature of the soil, you must use lime as conglomerate, commissioning can be immediate since lime acts quickly transforming the soil, while the strength gain can be increased with the passage of time producing a cementing of this soil, slow but inexorably.

The same occurs in relation to the maintenance of service paths, although these are not critical for the correct operation of the track or have restrictions as demanding like this available for repair periods. If the goal is to build so that maintenance is cheap and durable, it is clear that this is a field in which we have to take in account without any doubt with stabilized materials, without having to reach the employment of more expensive materials such as concrete or asphalt mixtures.

2. The technique of in-situ recycling - stabilizing

The technique is very simple and it is running since antiquity, although currently with media that make the cost-effectiveness and speed, very low.

It mainly consists of mixing the existing soils, whatever the quality they are, with the right binder in each case and the amount that is prescribed to achieve a certain result, with very few restrictions that will be described later.

This way or repairing demands some previous work more than traditional solutions replacing material by a harder one, but the results are well worth. Rigidities reached can be 10 times higher than the existing soil, or bigger depending on the circumstances, and therefore durability may also be much greater.

The binders used are usually Lime and Portland Cement, that are sufficiently known in the world of the construction of the public works with regard to its functioning, characteristic works, improvements that provide, cost, availability in the environment of the works, tests that should be carried out to control quality, etc, what is without a doubt a great advantage over other possible new materials based on polymers from petroleum or other sources.

Photo 1 Drying the platform

Each binder works in a certain way, has different advantages and their employment is not normally function of its price at each site, but the quality of soil which have to stabilize.

Briefly, when we operate with lime we produce a chemical reaction with clays, so that it breaks the laminated structure, which provides a great absorption of water, "soapy" appearance to glide particles on other, turning it in something more Sandy in appearance, which it decreases absorption of water, structure pass from SOAP to granular and increase significantly the capacity to support, obtaining CBR values from to 2 or 3, to values 20 or 30 with small additions of lime, usually around 2% lime, usually in the form of quicklime.

Lime cannot react with granular materials and not plastic materials so that its effect is clearly diminished and therefore is not recommended in these materials.

However cement works sticking mineral particles as stable and long-lasting glue, inferring materials treated with a greater resistance to getting excellent rigidity in time relatively short but also controllable through the type of cement, the endowment or additives that could be added.

In any way at any time one could speak of rigid or flexible solutions, but simply solutions to the type of soil that can be used on platforms already built and different ways of adapting them to the function to be fulfilled according to the dimension or the site where this soils have been placed.

The only restrictions you have to use this technique are to avoid any chemical reaction in the soil with the binders added that may produce some material that evolves with the passage of time in a bad way. It is known that this effect may occur when there is enough ion sulfate (SO3) that can react with free lime (Ca) added with the lime or the cement.

It forms a molecule (Ettringite) which stable form has 24 molecules of water and it has a very big swelling pressure during its formation.

Another bad circumstance may be the presence of stones bigger than the size recommended by the administration, 80 mm. bigger stones can cause great damage in the machine with a big cost.

3. Repairing areas with problems with soft soils or inadequate materials

It is not necessary to look for works in which similar loads be supported, or even higher, such as aircraft landing platforms or containers platforms in a port to be able to extrapolate the benefit of this technique to other fields.

In the following works, we will briefly describe the existing problem and the solution adopted the form of execution carried out.

Photo 2 Drying the platform in Olmedo

As an example, we have selected the following works:

- Repair by stabilization with cement on the platform already constructed the LAV Zaragoza to Lérida as solution to the emergence of subsidence on the platform at the time of the installation of the track.
- Drying the platform in the high speed line Madrid – Valladolid, in Olmedo, to accelerate the delivery of the work

Description of each one of these works:

- Repair by stabilization with cement on the platform already constructed the LAV of Zaragoza to Lérida

The problem in this section of the Madrid - Barcelona line was the emergence of soft soils, which originated many collapses in the platform, built two years earlier, at the time of the Subballast layer prior to placement of the track.

Photo 3 Stabilizing the platform in Zaragoza

The solution adopted was stabilization with cement, dry way, through the upper 30 cm for this platform providing the rigidity that lacked the floor of the platform only in the sections where the soil failed.

Before, they made some tests assuring that there were no SO3 problems in the soil.

The reparation was made very quickly without disturbing the rest of the works.

- Drying the final platform in the high speed line Madrid - Valladolid, in Olmedo

The 2002 winter rains soaked the entire platform of the work in this, which put into question the delivery from the stretch of corresponding work. That was the reason why it was decided to dry this platform by applying a small amount of quicklime in the most humid areas in the upper 50 cm platform getting the goal within a few days.

The only control they had to do was to check the final humidity and the compaction of the layer before setting the last one.

4. Conclusions

Soil stabilization is a solution to be considered in areas where for any reason the soil has failed or prevents to proceed with the work under the contracted deadlines.

The reason to choose the binder in each case is not its cost but the greater or less affinity with the type of soil that we have in our jobsite. The action of each binder is different in hardness and the speed is achieved. In soils with plasticity index greater than 15 we should better us quick lime as Binder but soils with Ip rates below 20 is usually recommended to use cement. (For materials with Ip between 15 and 20, would have to see other parameters).

The machinery necessary to carry out these stabilizations is usual machinery existing in the market, making it easy to work with that technique. The thickness that can be reaching by this machinery becomes 50 centimeters for mixing, which can be compacted with heavy rollers, more than 16-18 tons, usual in this type of work.

It is advisable to make prior to the execution analysis to predict the evolution of the area repaired in terms of hardness achieved platform and to see the amount of binder to be applied.

5. References

- ADIF. Pliego de Prescripciones Técnicas Tipo para los Proyectos de Plataforma PGP-2008 v.11. Artículo G10U "Tratamiento in situ de suelos con cal".
- ANCADE. "Manual de Estabilización de suelos con cal". Madrid, 1997.
- ANCADE y ANTER. "La Estabilización de suelos con cal en carreteras y líneas de alta velocidad". Varias Ponencias. Jornada Técnica SMOPYC-2008. Zaragoza 2008.
- ANCADE, ANTER e IECA. "Manual de estabilización de suelos con cemento o cal". Madrid, 2008.
- ANTER, ANCADE e IECA. Jornada Técnica "Estabilización de suelos en obras ferroviarias". Varias Ponencias. Madrid 2009.
- ANTER. "Guía de Soluciones para obras de estabilización de suelos, ejecución de suelocemento in situ y reciclado de firmes". Madrid, 2010.
- CEMOSA. Jornada Técnica "Estabilización de suelos". Varias Ponencias. Sevilla 2010.
- INTEVÍA. "Los suelos marginales en la construcción de Obras Lineales". Varias Ponencias. Sevilla 2002.
- INTEVÍA. "Mezclas con Cemento en Explanadas y Firmes". Varias Ponencias. Madrid, 2004.
- Junta de Andalucía – GIASA. "Recomendaciones para la redacción de Pliegos de Especificaciones Técnicas Generales para el Tratamiento de los suelos con cal". Sevilla, 2003.
- Laboratoire Central des Ponts et Chaussées (LCPC) – Service d'Études Techniques des Routes et Autoroutes (SETRA). "Traitement des sols á la chaux et/ou aux liants hydrauliques". Guide technique. Paris, 2000.
- Little, Dallas N. "Handbook for Stabilization of Pavement Subgrades and Base Courses with Lime". USA, 1995.
- Ministerio de Fomento. Pliego de Prescripciones Técnicas Generales para Obras de Carreteras y Puentes (PG-3). España.
- Ministerio de Fomento. Pliego de Prescripciones Técnicas Generales de Materiales Ferroviarios (PF). España.
- Sampedro, A. "Estabilización de suelos en Alta Velocidad Ferroviaria". Universidad Politécnica de Madrid. Departamento de Ingeniería Civil – Transportes. Madrid, 2006.
- Sampedro, A. "Innovaciones en las estabilizaciones de suelos con cal". Ponencia del VIII Congreso Nacional de Firmes. AEC y Junta de Castilla y León. Valladolid, 2008.
- TRATCICA 2008. "Seminario de Tratamientos de materiales con cal y cemento para Infraestructuras de Transporte. Casos de obras portugueses". Varios autores. Facultad de Ingeniería. Universidad de Oporto. Portugal, 2008.

High Speed Rail Infrastructure

360.revista de alta velocidad
número 5 - junio - 2018. Pág 237 - 249

Soil stabilization in new railway construction

López-Bachiller Fernández, Miguel
Sampedro Rodríguez, Ángel
Díaz Minguela, Jesús

[1]
ANTER

Alfonso X el Sabio School

IECA

Abstract

Since ancient times, all kind of soils have been improving in order to obtain a better use of them, getting more support and better resistance to the passage of time and the weather effects. These soils were modified with small additions of different materials, improving or stabilizing such soils.

Clay stabilized with lime has been used for the construction of the Shersi's pyramids in Tibet more than 5,000 years ago. Later on, already in the modern era, cement Portland was used to give resistance to the soils in cobble roads. In addition, it was used as well to stabilize a network of rapidly building airfields during World War II, giving way to what today we know as the modern roads.

Currently, soils stabilization is a technique of application widespread in many public and private works as factories, airports, ports, and all kinds of roads, from rural roads to modern motorways.

The purpose of this communication is to show the use of the technique in the construction of new railway works, achievements and targets set for the near future.

Keywords: punctuality, stabilizing, recycling, efficiency, quick repair

---

[1] *López-Bachiller Fernández, Miguel. ANTER, Spanish technical association for stabilizing and recycling platforms. Email: miguel@firmesecologicossoltec.com. (Corresponding author)*
Sampedro Rodríguez, Ángel. Alfonso X el Sabio School. Email: sampedro@uax.es. (Corresponding author)
Diaz Minguela, Jesús. IECA. Email: jdiaz@ieca.es

1. Object of this communication

Since ancient times, all kind of soils have been improving in order to obtain a better use of them, getting more support and better resistance to the passage of time and the weather effects. These soils were modified with small additions of different materials, improving or stabilizing such soils.

Clay stabilized with lime has been used for the construction of the Shersi's pyramids in Tibet more than 5,000 years ago. Later on, already in the modern era, cement Portland was used to give resistance to the soils in cobble roads. In addition, it was used as well to stabilize a network of rapidly building airfields during World War II, giving way to what today we know as the modern roads.

Currently, soils stabilization is a technique of application widespread in many public and private works as factories, airports, ports, and all kinds of roads, from rural roads to modern motorways.

The purpose of this communication is to show the use of the technique in the construction of new railway works, achievements and targets set for the near future.

2. The technique

The load-carrying capacity of the platform and the tensions that comes to it through dynamic loads of rail traffic is critical for determinate thickness of layers in any superstructure seat in earthwork projects. One of the goals of a correct dimensioning about railway's superstructure is to obtain vertical tension on the top side of the platform results clearly lower than its bearing capacity.

To improve the ability to increase the support capacity of the embankments in areas where clays or other lands with low resistance are, lime or cement can be used in order to improve the results obtained with these soils and to minimize the transportation of large masses of earth.

In addition to other objectives, minimize the construction costs, both economic and environmental, and the tasks of maintenance are the targets, reducing the technical and mechanical damages that maintenance work involves.

The rail traffic planned for the future and stresses that it will generate must be taken in account as in any type of project. If it is expected that this traffic to overcome to the admissible, it must be intervened to reduce tensions, by means of the modification of the railway seat structure, the increase of the platform capacity bearing or both actions at the same time.

There are still various uncertainties we can found regarding to:

- dynamic actions in the case of the high speed railway,
- the fatigue of materials behavior that are more requested
- the lateral resistance of the track and
- the maintenance of the mechanical properties obtained in the construction of the structure.

Consequently, this is one of the aspects that most should worry. International research and standard organizations are continuously studying alternatives to the traditional solution that they confer greater reliability and stability to the high performance railway movement.

It is in this context where an alternative to the classic structure is proposed, based on the use of layers treated with hydraulic binders. The use of both binders, lime or cement, applied and mixed with the soil leave a structure more resistant. They reduce the tensions on the platform and increase the transverse strength, both in the vertical loaded via (which allows better performance in the railway operation), as in the discharged via (improving stability of the track against the efforts of thermal origin). The desired stiffness of the soil can be obtained, and even, this rigidity can change in the path depending on the soil characteristics and the dosage of binder. (modulus variable in the material).

Other advantages that these solutions based on the use of layers treated with hydraulic binders can provide are:

- Use of materials of the working path what allows economies in the building phase. It avoids the quarries (inside and outside the working unit) and the need for the supply by other agents (that lets all internal tasks coordination and an important improvement in the speed of execution).

- It is a more environmentally friendly solution: avoiding opening or quarrying and vehicles of heavy transport by roads, with the damage that they do. In addition also this technique prevents the need for landfill sites where to send the soils that otherwise should be removed from the work.

- Use of manufacturing equipments that exist normally in the works. It is not necessary to incorporate special machinery.

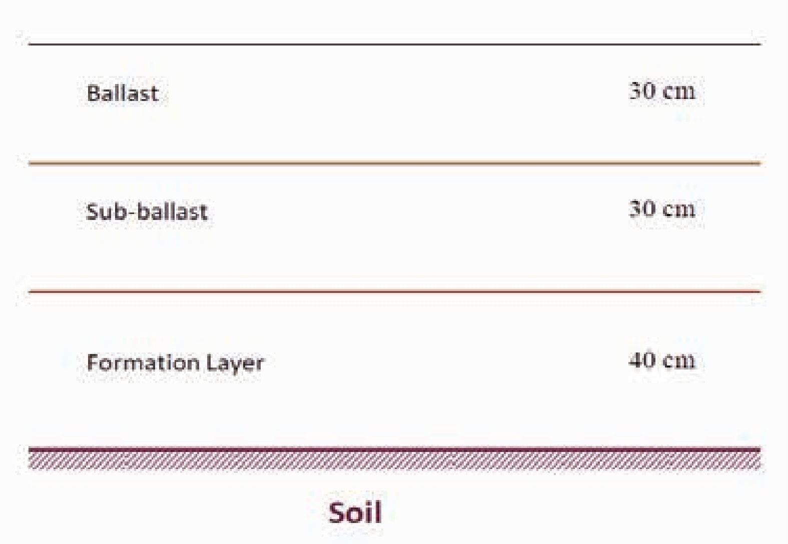
- Improves the impermeability of the base layer, obtaining greater stability in time (durability) and better performance against water. This increase of tightness is in addition to the 2% pumping that must be given to drain the water that has seeped through the ballast.

3. Traditional starting situation

Designers used to have abacuses to determine the thickness of the layers in the traditional railway (up to 200 km/h speed), according to the following parameters:

- Characteristics of the soils that constitute the platform on the stretch of railway considered.
- Characteristics of the platform as a whole.
- Climatic conditions of the location of the platform area.
- Characteristics of the traffic in the considered section.
- Characteristics of the superstructure.

Designers have been made various adaptations and extrapolations in the high speed train case in Spain. The structure that is used is reflected in the figure below:



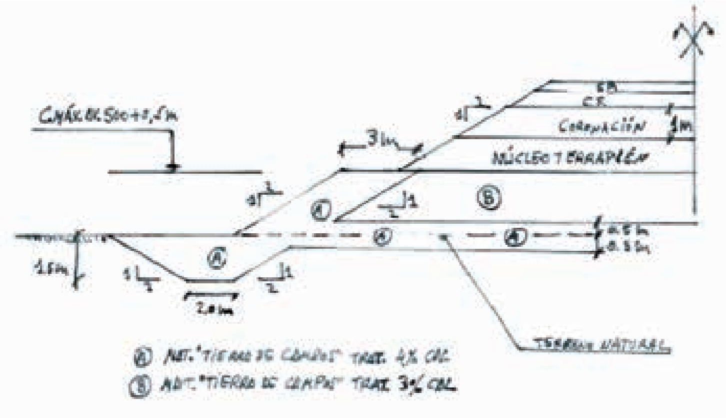
The demands of each layer are those listed below:

	Ballast	subballast	Track Bed	Soil	Platform
CBR	>22		>10	>5	10<CBR<20
E2 Modulus Plate-bearing	>1.700 kp/cm^2	>1.200 kp/cm^2	>800 kp/cm^2		
Permeability		< 10-6			
LA desgaste		< 24	<30		
Wet Micro-Deval		< 16	≤25		
Density (% MP)		> 98% PM	>95% PM	>1,75 kg/m3	
Max. size			≤10 mm		
Fines content			≤5%		
Atterberg limits				LL<50 If 0,35<LL<50, then Ip>0,73(LL-20)	
Type of soil					QS3, or QS2+40 cm QS3 or QS1+60 cm QS3

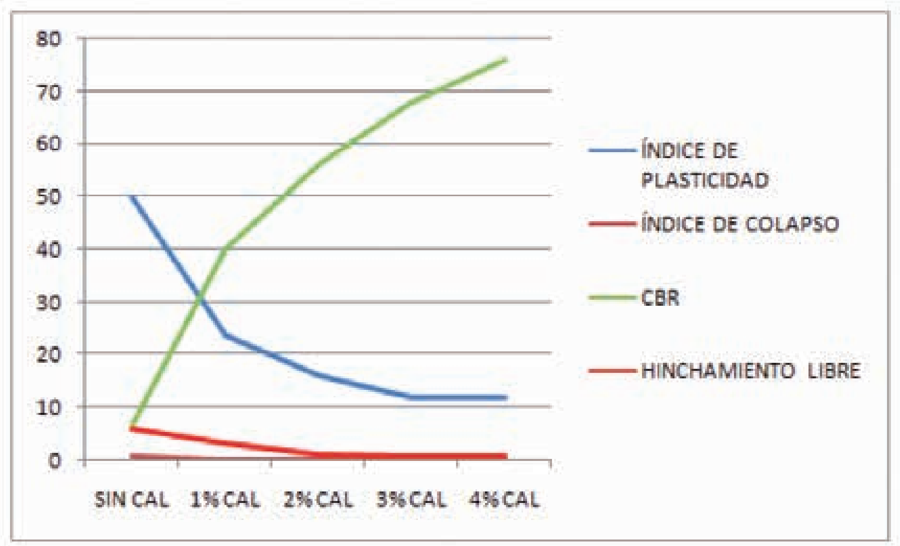
4. Alternative proposals

Stabilized materials make possible those soils of the own track can be used in different areas of the work that has to be built as the following:

- **Landfills:** treatment with a binder allows the reuse of the soils of the track, when their geotechnical properties are poor, and improving its support capacity.
- **Foundations for embankments or structures:** stabilization of soft soils that support embankments improves its ability to support and protect these areas of moisture, reducing the settlements of the embankments.
- **Zones of transition:** the stabilization of materials in the areas beside the engineering structures or hard points of the platform allows a gradual transition of behavior and rigidity of the materials.
- **Track bed layer:** any soil may be used provided that their geotechnical characteristics are improved. Thus, this layer will better support the solicitations during its service life, protecting the lower ones.
- **Special areas:** in certain cases, the compromised situation in some areas of the fillings, as flood areas, can be designed with some fillings stabilized with binder.



Modified of the soils characteristics in the case of using lime:



5. Potentially adverse factors

Three are the aspects which most negatively influence soil stabilization: its content in organic material, the soluble sulphates and the soils particles bigger than 80 mm.

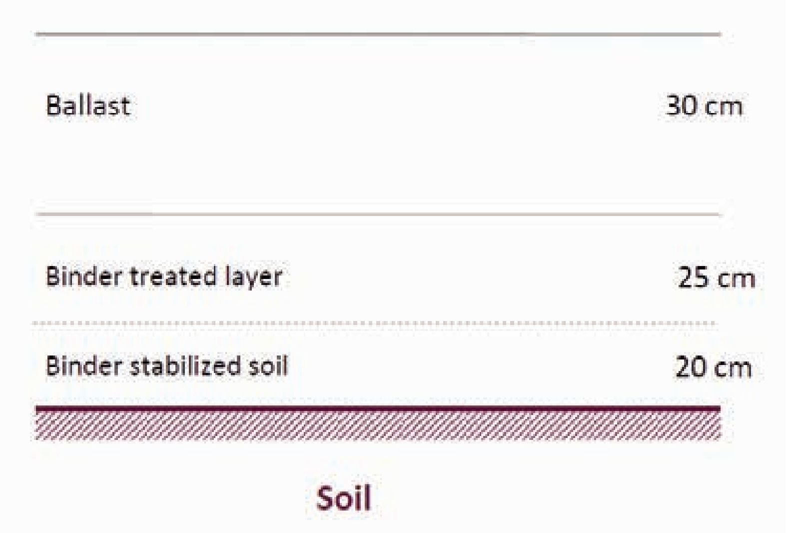
- The content in organic matter can inhibit the Pozzolanic reactions and retarder the effects of the binder on the ground. In case of soils with excessive rates of this, you should conduct studies seeing the evolution of the improvement to more long-term than usual, to dose correctly the necessary binder (lime can remove part of this organic matter).
- The content of soluble sulphates can form ettringite (trisulfoaluminato calcium) and other similar substances (thaumasite) produced by the soil and lime in a Pozzolanic reaction. They are very expansive and it can break the already mixed and compacted layers. The content of soluble sulphates can exist in the soil, or can be provided by existing groundwater that can affect stabilization through the reaction of sulphates solubilized in water with calcium aluminates hydrated.

  Lime needed for this reaction can come from lime added when this is used as a stabilizing agent, or from cement when it is released in the process of setting and curing.

  The use of sulfo-resistent cements does not exempt from the attention that must be paid in the works in which appear the sulfates. Normally cements that resist the attack of sulfur have a low content on alumina, therefore avoiding the formation of the product S - Al - Ca, whose stable form can contain up to 24 water molecules. However, in soil stabilization may be sufficient alumina content to form part of the damaging formula.
- The presence of big amount of big stones (bigger than 80-100 mm), especially if there is no continuous soil grading, makes impossible to achieve a good mixture and can cause problems in the machine.

6. Some example performed and comparative analysis

In order to improve the performance of the track structure and minimize implementation and maintenance costs, alternative solutions were used whose scheme responds to the following figure:



As you can see, 30 cm thick sub-ballast layer is replaced by a 25 cm thick treated material with a hydraulic binder layer. It allows the use of a soil of the working zone that does not require as stringent conditions as the one of sub-ballast features (for instance, many quarries don't meet the Los Angeles index in Northwest Spain). Also the thickness of the formation layer is considerably reduced (from 40 to 20 cm) when it is stabilized with cement and the required soil characteristics are much more permissive.

The soil may be the same used in the construction of any suitable layer (the formation layer for instance) or other which comply with some minimum requirements. The hydraulic binder could be medium resistant category cement or hydraulic lime. A 3-4% by weight of aggregates in case of cement and 1,5 – 2,5 % for lime jobs is usually used. The dosage depends on the nature of the soil to be treated, that is way previous tests with different provisions must be carried out.

The minimum requirements to the soil must be:

	UPPER LAYERS	FORMATION LAYER
Sieve size	< 80 mm Pass 2 mm 20 – 55 % Pass 0,063 mm 2-20 %	< 80 mm
Plasticity	LL < 30, IP < 10	LL ≤ 40, IP ≤15
Organic material	< 1 %	< 1%
SO3	< 0,5 %	
Salt water soluble		< 0,2 %
Minimum binder content by mass	3 % cement	3 % cement or 2 % Quick lime
Compression resistance 7 days specimens	2,5 MPa	1,5 MPa

As for the permeability of this layer, check that its value is less than 10 - 6 cm/s, which should be achieved easily.

Having one or some layers stabilized with hydraulic binder increases the degree and speed of consolidation of the structural package, which improves the lateral resistance of track.

In order to assess the behaviour of vertical stresses the structure of project and the different alternatives, a comparative analysis has been done. The program makes a multilayer analysis, commonly used for the assessment of different packages in different surfaces.

To use this program there are the following assumptions:

- Layers are horizontally infinite and finite vertically, except the bottom layer which is considered as a bottomless.
- Layers consist of homogeneous, isotropic and elastic materials and they are characterized by its modulus of elasticity E and its coefficient of Poisson.

Boundary conditions considered were as follows:

(a) On the surface:

- Tangential stresses are null, except in the case that horizontal loads are applied.

(b) At finite depth:

- In case of adhesion between two layers: stresses and deformations are equal on both layers at the level of the interface.

- In case of sliding between two layers: vertical normal stresses and vertical displacements are the same and a tangential stress is zero in both layers at the interface level.

(c) At infinite depth:

- Stresses and deformations are nil.

To determine the stresses and the resulting deflection in the layers with this program, the layers characteristics and the charge of a wheel type have been supposed. So, it has been obtained an estimate of the variation of the coefficient of ballast between the solution of project and the different alternatives and the value function of the dynamic load in each structure.

For estimation of stresses in the platform, the following elastic characteristics in each layer have been considered:

	E (MPa)	C.Poisson
Soil	50	0,35
Formation layer	100	0,35
Sub-ballast	150	0,35
Ballast	400	0,35
Soil stabilized	250	0,25
Soil cement	4.000 / 6.000	0,25

The determination of dynamic loads has been a rough, by the method of the SNCF estimate, based on a static 17.2 t axle load and have been considered two cases, both for speed of 350 km/h, obtaining the following values:

- With sub-ballast layer and a 50 kN/mm stiffness was obtained a dynamic loading of 12,68 t/wheel.
- With stabilized layers and a 62.5 kN/mm stiffness was obtained a dynamic loading of 14,14 t/wheel.

The output of the program provides graphically the maximum values of the stresses and deformations of each layer at the level of the different interfaces. It also indicates the place where these maximums are obtained and determines the value of deflection under the wheel and the radius of curvature of the deformed.

The following table is a summary of the results where deflection (Dk) and vertical stress on the top side of the platform (σ_z) are showed:

Summary of results:

	D_k 10^{-2} mm	σ_z platform
traditional section	84,11	$1{,}701 \cdot 10^{-2}$
20 S-EST + 25 SC (6000 MPa)	55,18	$1{,}364 \cdot 10^{-2}$
20 S-EST + 25 SC (4000 MPa)	57,96	$1{,}550 \cdot 10^{-2}$
25 S-EST + 20 SC (6000 MPa)	59,77	$1{,}662 \cdot 10^{-2}$
25 S-EST + 20 SC (4000 MPa)	62,21	$1{,}838 \cdot 10^{-2}$

The results of the following solutions can be highlighted:

- The solution with a layer treated with cement of 25 cm thickness and 6,000 MPa modulus of elasticity on a 20 cm of soil stabilized with cement, reduces a 34.4% deflections and a 19.8% the value of vertical stress on the upper face of the platform with respect to the natural soil solution.
- The solution with a layer treated with cement of 25 cm thick and 4,000 MPa modulus of elasticity on a 20 cm layer of soil stabilized with cement, reduces a 31.1% deflections and 8.88 % the value of vertical stress on the upper face of the platform with respect to the natural soil solution.
- In case of modifying the thickness, the solution with two layers treated with cement, one of 20 cm and 6000 MPa module over another stabilized 25 cm thick and 250 MPa module, decreases a 28.9% deflections and a 2.29% the vertical stress value on the top side of the platform with respect to the natural soil solution.

7. Conclusions

To carry out a practical assessment, it has been established the comparison of the different alternatives studied with regard to the usual section in railway superstructure with ballast (reference solution):

- 30 cm ballast (E = 600 MPa; v = 0.35)
- 30 cm sub-ballast (E = 200 MPa; v = 0.35)
- 40 cm formation layer (E = 100 MPa; v = 0.35)
- CBR soil > 5 (E = 50 MPa; v = 0.35)

In the absence of availability of quality aggregates in many areas of Spain, the possibility to modify this structure by replacing some of its layers made with natural soils by others treated with lime or cement has been studied. Different thickness and modulus of elasticity has been analyzed. In any case, the target is that aggregates and soils can ensure the stability of the superstructure over the passage of an important number of TBR and not be obliged to frequent maintenance tasks and high cost maintenance.

Acceptable results in comparative terms used a program prepared for road traffic have been obtained (vertical stress σ_z in the upper part of the Esplanade and the deflection of the Dk are the parameters achieved).

The following alternative among the studied ones is selected:

- 30 cm ballast (E = 600 MPa)
- 25 cm of soil treated with cement (E = 4,000 MPa)
- 20 cm of soil stabilized with cement (E = 250 MPa)

This solution has been implemented in a test section with good results. Several years later, different values were measured and compared with the ones obtained in a traditional section. The seats of the railroad ties varies between 0,40 and 0,50 mm. The results are in similar orders of magnitude in both sections. The stiffness increase measured is 17% higher in the case of stabilized soils. Probably a longer term project will be necessary in order to obtain more and best results.

Its advantages are:

- A reduction of 31.10 % in deflection.
- A reduction of 8.88 % in tension.
- An important height reduction.
- A 17% higher stiffness increase.
- A substantial reduction of the condition of the environment to avoid transport and manipulation of materials extracted in other areas and the poured eventually to landfill.
- Use of local soils of the working place (excavation and materials from the tunnels, with reduced technical limitations).
- Obtaining layers with more stable materials and greater life cycle.
- Reduction in the requirement for maintenance and repair work and its high cost that they hinder and restrict the exploitation of the railway traffic.
- It is possible to obtain the desired stiffness and even, this rigidity can change in the path depending on the soil characteristics and the dosage of binder (modulus variable in the material). It is more, the rigidity is more homogeneous than in the case of use only soil without binder.

8. Appendix 1: stabilizations in the spanish high speed railway

8.1 Stabilization with cement

- High speed Atlantic axis. Cerceda-Meirama, access Sogama branch section. Stabilization of formation layer and replacement of the Sub-ballast by a soilcement done on-site with the same equipment or larger percentage of cement stabilized soil. Company Ferrovial
- New railway access to the North and Northwest of Spain, Madrid - Valladolid.
- Tramo: Valdestillas-River Duero, stabilized with lime formation layer and stabilized with cement service roads, and stabilization of technical structures in Valdestillas blocks. Company: AZVI
- Tramo: Segovia-Valladolid. Work: Stabilized in slabs beside structures with cement and lime. Thickness 30 cm, April 2002. Company: Ute Valdestillla(Collosa-Comsa). Property: Entity manager of railway infrastructure
- Work: Ave Segovia source of this cross. Area 6,000 m2. Work: Stabilized with cement through the dry. Thickness 30 cm, July 2005. Company: Vias y Construcciones.
- Work: Segovia AVE station. Stabilization of the platform of the track shoulders.
- New railway access to the Northest of Spain, Madrid - Zaragoza - Barcelona:
- Ave L' Arboç (Tarragona). Work: Stabilized with cement. Thickness 30 cm, September 2002. Company: Ute Ave L' Arboç (Dragdos-Tecsa).
- Zaragoza - Lerida. Stabilization of a layer of the platform in deforestation between NII motorway towards Lleida to increase some capacity after spending the winter.
- AVE 350, Tr. Zaragoza. Area 165.000 m2. Work: Stabilized platform. Thickness 30 cm, March 2001. Company: Ute AVE350 (Tecsa, Guinovart, Coprosa).

- High speed Antequera-Granada line. Section: Pinos Puente-Granada. work: stabilized service roads and embankments of structures. UTE Vimac-Isolux Corsan
- Rail access to the Terminal del Prat (Barcelona port). Surface 3.100 m2. Work: Stabilized with cement, wet way. Thickness 40 cm, June 2012. Enterprise: Infesa (engineering and rail services)
- Ave El Casar (Cáceres). 200,000 m2 surface area... Extraction of material that fill the gaps between granite massifs, mixed and stabilized with cement with a bulldozer (could not be with mixing machine because of the large amount of big stones) and filling again. Thickness up to 40cm, October 2011. Company: Copisa
- Seville container terminal. 64,000 m2 surface area. Work: Stabilized with cement. Thickness 30 cm, June 2009. Enterprise: Rover Alcisa
- Salamanca. Work: Removal step level small roads in Pedroso de la Armuña (Salamanca). 20,000 m2 surface area. Work: Stabilized with cement. Thickness 25 cm, November 2006. Company: Extraco
- Ávila. Overpasses in the railway from Ávila. 20,000 m2 surface area. Work: Stabilized with cement. Thickness 20 cm, June 2005. Company: Intersa

8.2 Stabilization with lime

- New railway access to the North and Northwest of Spain, Madrid - Valladolid - León:
  - Section Madrid - Valladolid. Stabilization of formation layer, drying of the platform and stabilization with lime in Olmedo.
  - Section Madrid - Valladolid. Platform in Valdestillas drying with quick lime and stabilization.
  - Villada working platform: stabilized with lime formation layer. Company Azvi.
  - Ave Segovia Valladolid. 287,000 m2 surface area. Work: Stabilized with lime. Thickness 30 cm, September 2002. Company: Constructions Sanchez-Dominguez (Sando).
  - Property: G.I.F. (railway infrastructure manager).
  - AVE Palencia - León. Stabilization of the formation layer in Sahagun de Campos.
  - AVE Sahagún. Surface 300,000 m2. Work: Recycling with lime. Thickness 30 cm, may, August and September 2010. Company: Constructora San Jose.
  - LAV UTE Pozal in Medina del Campo (Valladolid)
  - LAV stretch Santas Martas-Bercianos del Real Camino (León) LAV stretch Bercianos de Real road - Rio Cea (León).
- High speed Bobadilla-Granada line. Section: Peña de Los Enamorados – Archidona. Work:
  - Stabilized with lime nuclei of embankment. UTE Dragados-Azvi.
  - LAV stretch stump-Valderrubio (Granada)
  - LAV stretch pine Puente-Granada (Granada) LAV stretch Vera - Los Gallardos (Almería)

- Perpignan. Stabilization of the platform with lime slurry.
- Railway connection of the LAV Madrid with the Mediterranean, C-14 stretch - runner Constantí (Tarragona)
- Salamanca. Work: Suppression of three level crossings in Peñaranda Bracamonte . 17,000 m2 surface area. Work: Stabilized with lime dry. Thickness 20 cm, January 2004. Company: Azvi. Property: Renfe.

9. Appendix 2: references

- ADIF. Pliego de Prescripciones Técnicas Tipo para los Proyectos de Plataforma PGP-2011 v.11. Artículos G-0104 sobre “Rellenos” y G-0112 “Tratamientos in situ de suelos con cal”.
- Ministerio de Fomento. Pliego de Prescripciones Técnicas Generales para Obras de Carreteras y Puentes (PG-3). España.
- Ministerio de Fomento. Pliego de Prescripciones Técnicas Generales de Materiales Ferroviarios (PF). España.
- IECA. Guía Técnica Estabilización de suelos con cemento en obras ferroviarias. Marzo de 2013.
- ADIF y CEDEX. Jornada sobre “Estabilización de terrenos con cal”. Varias Ponencias. Madrid, 30 de noviembre de 2012.
- Quereda Laviña, José y Díaz Minguela, Jesús. Aplicación de capas tratadas con cemento en la superestructura ferroviaria. Propuesta de alternativa de infraestructura de la vía. Determinación de la carga dinámica y tensiones de la plataforma del tramo de ensayo en Meirama. Octubre de 2011.
- ANTER. “Guía de Soluciones para obras de estabilización de suelos, ejecución de suelocemento in situ y reciclado de firmes”. Madrid, 2010.
- CEMOSA. Jornada Técnica “Estabilización de suelos”. Varias Ponencias. Sevilla 2010.
- ANTER, ANCADE e IECA. Jornada Técnica “Estabilización de suelos en obras ferroviarias”. Varias Ponencias. Madrid 2009.
- TRATCICA 2008. “Seminario de Tratamientos de materiales con cal y cemento para Infraestructuras de Transporte. Casos de obras portugueses”. Varios autores. Facultad de Ingeniería. Universidad de Oporto. Portugal, 2008.
- Sampedro, A. “Innovaciones en las estabilizaciones de suelos con cal”. Ponencia del VIII Congreso Nacional de Firmes. AEC y Junta de Castilla y León. Valladolid, 2008
- ANCADE y ANTER. “La Estabilización de suelos con cal en carreteras y líneas de alta velocidad”. Varias Ponencias. Jornada Técnica SMOPYC-2008. Zaragoza 2008.
- ANCADE, ANTER e IECA. “Manual de estabilización de suelos con cemento o cal”. Madrid, 2008.
- Sampedro, A. “Estabilización de suelos en Alta Velocidad Ferroviaria”. Universidad Politécnica de Madrid. Departamento de Ingeniería Civil – Transportes. Madrid, 2006.INTEVÍA. “Mezclas con Cemento en Explanadas y Firmes”. Varias Ponencias. Madrid, 2004.
- Junta de Andalucía – GIASA. “Recomendaciones para la redacción de Pliegos de Especificaciones Técnicas Generales para el Tratamiento de los suelos con cal”. Sevilla, 2003.

- INTEVÍA. “Los suelos marginales en la construcción de Obras Lineales”. Varias Ponencias. Sevilla 2002.
- Laboratoire Central des Ponts et Chaussées (LCPC) – Service d’Études Techniques des Routes et Autoroutes (SETRA). “Traitement des sols á la chaux et/ou aux liants hydrauliques”. Guide technique. Paris, 2000.
- ANCADE. “Manual de Estabilización de suelos con cal”. Madrid, 1997.
- Little, Dallas N. “Handbook for Stabilization of Pavement Subgrades and Base Courses with Lime”. USA, 1995.

High Speed Rail Infrastructure

360.revista de alta velocidad
número 5 - junio - 2018. Pág 251 - 261

Maintenance. From asset management to direct cost calculation. A key issue for the future of the HS Railways System.

Jiménez, Pablo Lorenzo
Zamorano, Clara Isabel
Escobar, Adrián
Escobar, Jorge

European University Madrid[1]
Technical University of Madrid

Abstract

The separation between infrastructure managers and transport operations is one of the basic principles of European railway regulations. The relationship between railway undertakings and infrastructure managers is based, among other considerations, on the payment of a charge for the use of the infrastructure.

The recast of the first railway package (Directive 2012/34/EU) states that “the charges for the minimum access package and for access to infrastructure connecting service facilities shall be set at the cost that is directly incurred as a result of operating the train service” (Article 31.3). This is what is known as the direct cost principle, whereby the infrastructure manager will apply a charge so that, if an additional train passes, the charge paid by the latter will allow no losses to occur. The current formulation of the charge is based on different parameters, among which are the number of trains-km, speeds or train types.

To enable the establishment of these charges, infrastructure managers need to have a comprehensive understanding of all the elements that determine the cost of operating the infrastructure. Amongst them, maintenance costs caused by normal operation of train services might be considered.

The purpose of this presentation is to discuss a model that could be used to improve the implementation of the direct cost principle, in terms of maintenance costs. Firstly, the formulation of the fees is analysed, using the Spanish case as an example. A brief overview of railway maintenance and asset management is given below. Different studies are then considered on the relationship between tear and wear, as well as the costs of railway maintenance. Finally, a proposal is made for a new methodology.

Keywords: Directive Recast, Infrastructure charges, direct costs, asset management, maintenance

---

[1] *Jiménez, Pablo Lorenzo. Civil Engineering Department European University Madrid. (corresponding author). Email: pablolorenzo.jimenez@universidadeuropea.es.*
Zamorano, Clara Isabel. Technical University of Madrid. Email: clara.zamorano@upm.es
Escobar, Adrián. Technical University of Madrid. Email: adrian.escobar.pastor@alumnos.upm.es

1. Introduction

Since the beginning of the 1990s, European railway legislation has sought to revitalize the sector by separating the functions of provision of transport services and infrastructure management. An essential aspect of this separation is to correctly regulate the relations between the two functions, ensuring equitable and non-discriminatory access for undertakings wishing to provide their services on railway infrastructure. Access to the infrastructure and services of the infrastructure manager must be paid by the levying of charges.

The regulation of these charges is a very important aspect of European railway legislation. Section 2 of Chapter IV of Directive 2012/34/EU (Recast) deals with fees and the principles to be followed in setting, applying and charging them. One of these principles, set in article 31.3, states that *"the charges for the minimum access package and for access to infrastructure connecting service facilities shall be set at the cost that is directly incurred as a result of operating the train service"*

The determination of the direct costs is a difficult aspect, since it is a matter of discerning which part of the costs incurred by the administrator in his activity is reportable to each circulation. One of the main costs is due to the maintenance of infrastructure, which is one of the fundamental functions of the managers, as set out in Article 7 of Directive 2012/34 / EU, as amended by Directive 2016 / 2370.

The Rail Sector Law (Law 38/2015) reflects the Recast Directive, and in fact, the PITVI 2012-2024 establishes three major action programs. One of them is the regulation, control and supervision program. Within this, one of the points is the modification of the system of charges, as a consequence of the application of the Recast, explicitly stating that *"for the Conventional network, the charge per use will be equivalent to the directly attributable cost to the operation of the rail service, and for the High-Speed network, whose objective will be the recovery of costs, the charge will include surcharges based on principles of efficiency, transparency and nondiscrimination".*

This task is fundamental for the rail system. The rail system, in search of greater speeds in the transport of passengers, and of greater loads in the transport of goods, needs a high level of maintenance, to achieve a high-quality infrastructure. Therefore, for a correct pricing of the charges, it is essential to allocate to each circulation the cost that it entails within the maintenance.

The purpose of this paper is to analyse the methodology of direct costs, and how it can be used for the costs stemming from maintaining the railway infrastructure, and, more specifically, the platform and track. For this, first, we analyse the charges and fees, and the methodology of direct costs. Next, we deal with rail maintenance. Subsequently, the existing methodologies for the determination of the direct costs of railway maintenance are analysed. Finally, a proposal is made to improve these methodologies, based on the condition data of the different lines.

2. Charges and european directives. Direct costs

2.1 Concept of charges

Charges constitute an important part of the income of railway infrastructure managers. They cover the operating costs of the infrastructure manager, and, therefore, a high part of the price paid by the passenger for his ticket is intended for the levying of charges (Fernández Arévalo, 2013). Thus, charges are an important variable for the economic balance of the system.

Directive 1991/440 marked a milestone for European railways, changing totally their management. The main idea is the separation of the activities of infrastructure management and provision of transport services. The access rights to railway infrastructure must be granted in a uniform and non-discriminatory manner. The payment for this access shall be done with charges; the determination of these is further developed later different normative texts, among which the following can be found:

- Council Directive 95/19 / EC of 19 June 1995 on the allocation of railway infrastructure capacity and the charging of infrastructure fees
- Directive 2001/14/EC of the European Parliament and of the Council of 26 February 2001 on the allocation of railway infrastructure capacity and the levying of charges for the use of railway infrastructure and safety certification.

In Spain, its development is done through the Rail Sector Law. The first law of such name, Law

39/2003, includes the concept of infrastructure charges in section IV of title V. Subsequently, in Law 38/2015 (second Rail Sector Law), the structure of the charges is modified, contemplating the payment of a charge for use of the railway lines, called Mode B. This charge is divided into two parts: a taxable base, to be determined by the infrastructure manager, and an additional, to be determined in the Law of General State Budgets[1]. Each year, the Network Statement, published by the Infrastructure Manager, sets out the general rules, deadlines, procedures and criteria that will govern the fees.

As indicated in the previous section, the European Union sets the principles to be followed by national regulations. A basic principle is that the infrastructure manager may only include in the calculation those costs for which it can objectively and rigorously demonstrate that they derive directly from the operation of the different services. This means that elements whose wear does not vary because of traffic, such as wear of track signs or interlockings, should not be included in the charge.

This is known as the principle of direct costs, whereby the infrastructure manager will use a charge so that, if an additional train passes, the amount paid will cause no loss in the system. The formulation currently used for the charges is based on different parameters, among which is the number of trains · km, speeds or types of train.

The Commission Implementing Regulation (EU) 2015/909 of 12 June 2015 for the calculation of the cost that is directly incurred of the cost that is directly incurred as a result of operating the train service sets out a number of ideas for infrastructure managers, or bodies responsible for the determination of charges, so they can calculate these direct costs, and thus determine which parameters should be used.

It should be noted that, according to Fernández Arévalo (2013), railway infrastructure is characterized by very low marginal costs, which means that a very low percentage of the total costs can be recovered with infrastructure charges. As Fernández Arévalo states, this can result in a lack of incentives for investment in new infrastructures to address capacity problems, as this would contribute to increase the manager's deficit. However, raising the charge above average costs can result in the expulsion of operators with lower market resources, reducing competition.

---

[1] *Law 3/2017, of June 27, of General State Budgets for the year 2017, which includes these aspects in articles 71 and 72.*

2.2 Charges in Spain

The charges' structure currently used in Spain can be seen in the Network Statement 2017, last published by Adif and Adif AV. The charges are applied for the use of the lines, as well as the services rendered on them. Currently, there are the following modalities of rail charges and fees (Network Statement, 2017):

- Fees:
  - Safety Fee in Passenger Transport
  - Other Fees: to provide the necessary services for the granting of approvals, certifications, to issue titles to railway staff and grant Licenses or Safety Certificates

- Rail Tariffs:
  - Tariffs for using railway lines:
    - <u>Access tariff</u>: tariff for the intended use of the RFIG[2]. It depends on the type of line, and the estimated traffic volume, measured in km / trainyear.
    - <u>Tariff for reserving capacity</u>: levies the availability of the route requested. It depends on the time period, the type of line, and the type of service / train. It is measured in € / Train-km booked.
    - <u>Tariff for running</u>: it is similar to the previous one, but it governs effective use of the reserved capacity. It is measured in € / km-circled.
    - <u>Traffic tariff</u>: levy for the traffic produced on rail infrastructure. It is measured in € / 100 seats-km.
  - Tariffs for the use of stations and other railway facilities:
    - <u>Tariff for using stations</u>: it is applied to passengers using the rail transport service, depending on the distance covered and the station category in which the journey begins or ends. It is measured in € / passenger.
    - <u>Tariffs for stabling and using platforms at stations</u>: this tariff is calculated considering the period of train stabling, track change operations performed upon request of the operator and the category of the station. It is measured in € / Train.
    - <u>Tariffs for passing through gauge changers</u>: it is applied at every passing of a train by a gauge changer. It is measured in €/Train.
    - <u>Tariffs for using sidings</u>: they are set according to the type of lines at the station where used siding track belongs, to track occupancy time and type of service/ train. It is measured in € / Train.
    - <u>Tariffs for providing services that require authorization to use public rail system</u>: it is determined according to the area occupied. It is measured in € / m^2- month

It should be noted that the railway charges cited in the Network Statement correspond to the concepts form Law 39/2003 (previous Rail Sector Law), and have not been developed in accordance with Law 38/2015 (current Rail Sector Law). However, in the preamble of Law 38/2015, it is stated the new structure of the charges is simpler, and it is intended to stimulate the traffic better than the previous one. The charges for access are abolished,

---

[2] *RFIG: Rail Network of General Interest. A definition can be found at Adif's Network Statement.*

as they constituted a barrier to the entry of new operators to the market. Besides, new criteria for the classification of lines are implemented

The tariff structure for the use of facilities is also modified. The classification criteria for passenger transport stations are modified and expanded to take into account the economic capacity of associated services to determine the amount of the fee.

As a conclusion, neither in the current structure of charges, nor in the one to come, costs stemming from infrastructure maintenance are directly considered. In the current structure, the tariff for running would cover maintenance costs, as well as other variable costs due to the passage of trains.

3. Rail maintenance and asset management

3.1 Main railway maintenance activities

Spanish national standard UNE-EN 13306: 2011 defines maintenance as *"the combination of all technical, administrative and management actions during the life cycle of an element, intended to preserve or return it to a state in which it can perform the function required”* (UNE-EN 13306:2011).

In any element, we can distinguish two main types of maintenance, according to the occurrence of failures: preventive maintenance, when maintenance aims to avoid failure, and corrective maintenance, which is done when the failure has already occurred. The standard defines them as follows:

- Preventive maintenance: *"maintenance that is performed at predetermined intervals or according to established criteria, and is intended to reduce the probability of failure or degradation of the functioning of an element"*.
- Corrective maintenance: *"maintenance that is performed after the recognition of a fault and that is destined to put an element in a state in which it can perform a required function"*. That is to say, it is the one destined to correct or to repair the defects already produced, with potential harm to the service.

Preventive maintenance operations aim to minimize the likelihood of a failure in the elements, due to operating risks, and the higher cost that corrective maintenance entails. Within preventive maintenance, we can distinguish predetermined maintenance, which is performed according to established time intervals or operating units, and condition based maintenance, where the condition of the element is monitored or inspected to determine the actions to be undertaken.

In railway infrastructure, the different elements that integrate it have different associated maintenance methodologies. In the track, preventive maintenance operations that are carried out are based on the previous knowledge of the state of the elements, and, therefore, it is mainly a condition based maintenance.

To determine the state of the track, different systems are used. The status of track elements can be checked, or geometric quality of the track can be analysed. For this purpose, different types of inspections can be used: walking tours on the infrastructure, cabin rides with specialized staff, or train measurements, where we can distinguish several types: geometric, dynamic, ultrasonic for the condition of the rails, etc.

The geometric quality of the track measures the deviations between the theoretical track and the real track with a series of parameters. The definition of these parameters, and the methodology for measuring them, is defined in the standard UNE-EN 13.848.

The parameters used to determine the geometric quality of the track are the following: gauge irregularities, longitudinal and cross level irregularities, alignment irregularities and twist irregularities.

3.2 Asset Management

A basic tool for managing railway infrastructures is the asset management policy they follow. Efficient management results in an increase in the life cycle of the different elements, besides of a reduction in the costs associated with that life cycle. This type of management should foresee maintenance needs, to cover them before they limit circulation or pose any risk. Therefore, in order to optimize maintenance, it is necessary to have tools that allow to anticipate the needs.

In 2010, the UIC issued a guideline document for the application of asset management policies (available in Adif, 2011). This document explains briefly the steps to follow for the implementation of the system. Some aspects to watch are the definition of an asset strategy, and the data that will feed it, which is the asset information. The basic data that should be available for each asset include the following:

- Instalation date
- State
- Failure history and performance / security impact
- Maintenance and renewal history and plans
- Maintenance and renovation unit costs

These data are introduced in LCC (Life Cycle Cost) tools, which aim to optimize maintenance and renewal decisions, providing a forecast of all activities to be carried out annually.

3.3 Maintenance costs

The costs associated to the maintenance are necessary to feed the LCC tools. Obviously, knowing these costs correctly will allow us to fine-tune the optimization process and obtain a correct forecast.

On the subject of costs, it is worth noting that more than 50% of the maintenance costs due to traffic are allocated to track, according to Larsson & Gunnarsson (2001), in their analysis of the Swedish network.

It is clear that an improvement in the procedures followed in maintenance, seeking a better knowledge of the infrastructure, and the prognosis of the problems that will occur, should result in a saving in the costs incurred, and therefore in the canons passed. This is in line with Article 30.1 of Directive 2012/34/EU, which states that "managers shall be encouraged to reduce the costs of making infrastructure available and the amount of access".

4. Methodologies for estimating direct costs

The levying of charges and the principle of direct costs have led to a need for research in order to develop a methodology to calculate these direct costs, for each type of traffic.

4.1 Previous experiences

Before presenting these methods, let's take a brief look at the development of this problem. The first country to establish a separation between operation and infrastructure management was Sweden in 1988, when it separated its former state railway authority into two parts:

Statens Järnvägar (service operation), and Banverket (infrastructure management)[3]. For this reason, it is not surprising that the first investigation in this field was made in that country. Subsequently, the number of studies has expanded, covering other networks, such as Wheat and Smith (2008) for the English network. All approaches until 2015 follow the so-called econometric model; in 2015, Smith et al. present a model in which the damage caused in the infrastructure by different types of vehicles is analysed.

4.2 Methodologies

In order to assess correctly the amount of the fee, it is necessary to develop a methodology to determine what expenses are incurred. The problem is therefore to establish a link between the damage generated by a train, and the cost this entails for infrastructure. For this, we can use different models, for which we will need to have a series of inputs that can come from different types of analysis.

Wheat and Smith (2008) identify three different methodologies for determining directly incurred costs: top-down approach, bottom-up approach, and cost allocation method. These methodologies are explained in the following sections.

4.2.1 Econometric approach

The econometric approach, also called "top-down", is the determination of a global cost function for the maintenance of infrastructure. In this perspective, the concept of direct costs corresponds to the marginal cost, that is to say, to the increase of cost that is produced in the maintenance when one train runs over it. To obtain this, a global cost function must be established, which is done by statistical methods. Once this function is generated, the average cost of maintenance caused by the passage of a train can be approximated as the marginal cost, which is equivalent to that derived from the cost with respect to the number of trains running. This method has been the most used to the date.

4.2.2 Engineering approach

The engineering approach, also called "bottom-up", is based on the application of any method that allows to determine the concrete damages caused by a train and the subsequent maintenance needs. By valuing the cost of these, the cost to pay for each circulation can be calculated.

Different types of trains will produce different types of wear on the infrastructure. For example, the damage caused by a high-speed train will not be the same as by a freight train with a high axle load. Therefore, in this approach, the damages produced by each type of train are determined.

---

[3] *Subsequently, there have been more changes in the organization of the Swedish railways. Statens Järnvägar has been separated into seven companies, of which four have been sold. The provision of passenger services (SJ AB) and goods (Green Cargo) remains public. In terms of infrastructure management, the inspection functions (Järnvägsstyrelsen) have been separated from those of administration, which in turn has merged with the Trafikverket. The maintenance of the infrastructure is the responsibility of a subsidiary company, Infranord.*

Within their analyses, Smith et al. (2016), and Smith et al. (2017) use these methods to determine the damage produced by each type of train traveling through the study sections.

4.2.3 Cost allocation

The cost allocation method can be defined as a hybrid between both methods (Nielsen et al., 2016). Smith et al. (2016), and Smith et al. (2017) employ hybrid methods, with both economics and engineering analysis. These studies start from the idea that it is difficult to implement methods that, by means of a single step, econometric or engineer, is useful to determine the correspondence between costs and damages. For this reason, it is necessary to use methods with both econometric and engineering approaches.

It should be noted that, in these methods, the amount paid for maintenance in previous years, is considered to be produced by the different sources of damages. This has a problem, as it implicitly assumes the hypothesis that it covers the maintenance needs of the period. It is normal that, in situations of budget shortages, a certain maintenance deficit is generated, which in turn causes losses of the patrimonial value of the infrastructures.

5. Proposal for a new methodology

As it has been stated the models developed for the moment may have an econometric approach, an engineering approach, or a combination of both. The general idea from this methodology proposal is that it would be advisable to integrate asset management systems with charges calculation. This proposal, based on asset management systems, could be classified as a mixed proposal, as seen in the previous section, as it encompasses an engineering and economics approach.

Beginning with an initial estimate of maintenance costs based on the experience of infrastructure managers, the continuous use of these systems over several years would allow these costs to be fine-tuned until a good correlation between the forecast and the actual cost incurred is found.

To do this, the data of train measurements and asset management systems would be used to determine the condition of the track and the maintenance needs generated between two successive passes of the measurement trains. The steps would be:

1. The measurement data would allow to evaluate the damages that have occurred between t two successive passes. To do this, the asset management strategy should define what would be the minimum frequency of passing for the measurement trains.

2. The asset management system uses these data, and cross-matches them with data on expected traffic, age of elements, etc., to generate an estimate of the annual maintenance costs, as well as the replacement rates. Here, it is important to consider that we must separate the costs associated with the elements that present a wear and tear on the traffic (susceptible, therefore, of a direct cost), of those that do not have wear associated with traffic directly.

3. A fundamental aspect is to make a division of the infrastructure in homogeneous sections depending on its characteristics and traffic, to obtain comparable data

that can be extrapolated. The analyses to separate the sections by homogeneous types must consider the following parameters: type of infrastructure, type of track weaponry, climatology, layout in plan and elevation, etc.

These data would be cross-matched with data of traffic to obtain an estimation of costs by train and characteristics of the same in each specific section, producing a system of equations with as many equations as sections, and as many unknowns as types of trains, and characteristics of traffic. This would be easier in High Speed tracks, as the rolling stock would be more homogeneous.

From our point of view, this method would present the following advantages:

1. Starting from actual damage to the track, not from simulations, whose results would need to be calibrated.
2. With many track sections, a good amount of different cases would be available.
3. For different levels of track quality, the accumulated damages can be assessed, making possible to extrapolate in other cases with a different initial quality of the track.
4. It would be possible to calculate the real cost per type of train for each section of track.
5. This methodology, consisting of integrating the cost model with the asset management model, has the challenge of gathering data from both the financial department and the asset management department.

6. Conclusions

The integration of European Directives leads infrastructure managers to the use of the methodology of direct costs for the assessment of charges. One of the main costs incurred by railway infrastructure managers is maintenance and part of it is directly attributable to railway operation. Therefore, it must be reflected in the calculation of the tariffs. However, the cost of maintenance resulting from the passage of a train is not easily assessed. Therefore, different approaches have been developed, such as the econometric one, where a generalized cost function is determined to obtain the marginal cost, or the engineering approach, based on dynamic simulations, in which a correspondence is made between the behaviour of the trains and the damages produced on the track.

It should be noted that the methodology of direct costs has several limitations:

- Tracks with lesser quality, and thus, with grater maintenance needs, will have higher marginal costs. However, the major maintenance needs may be due to non-trafficdependent factors:
  - By the presence of worse layouts, with more curves and steeper slopes.
  - Greater deterioration of elements. The higher the wear, the higher the wear rate, and the higher maintenance requirements. Tracks in worse state are likely to cause greater damage to trains, due to the presence of defects which

influence the dynamic behaviour of trains, and therefore, on their elements, such as wheels (producing, for example, flat wheels). Thus, tracks in worse condition, which are going to cause greater damage on the trains, would have higher charges, which is a contradiction.

- The wear of the elements is considered completely attributable to the passage of trains, and does not consider the influence of time.

In this paper, an idea of a methodological proposal for the calculation of infrastructure maintenance direct costs has been posed, based on the use of asset management tools. This proposal, correctly calibrated, would allow a realistic approximation to the costs directly incurred by the passage of a train.

7. References

- ADIF (2011) Gestión de Activos en Administradores de Infraestructura Ferroviaria.
- Estrategias Ferroviarias Europeas. Número 50. Marzo 2011.
- ADIF (2017). Network Statement
- ADIF AV (2017). Network Statement
- ANDERSSON, M.; SMITH, A., WIKBERG, A. & WHEAT, P. (2011). Estimating the marginal cost of railway track renewals using corner solution models. (No. TRV 2010/92179). Stockholm. Retrieved from

  http://urn.kb.se/resolve?urn=urn:nbn:se:vti:diva-5361
- FERNÁNDEZ ARÉVALO, FRANCISCO JAVIER. (2013). Bases económicas para la regulación del canon por uso de la infraestructura ferroviaria. Revista de Derecho de la Unión Europea. Nº 25. Pp. 205-220.
- LARSSON, D. & GUNNARSSON, J. (2001). A Model to predict Track Degradation Costs. 7th International Heavy Haul Conference.
- NIELSEN, D., SPIRYAGIN, M., & COLE, C. (2016). A model to estimate infrastructure damage costs for different train types. In CORE 2016, Maintaining the Momentum, Conference on Railway Excellence, Melbourne, Victoria, 16-18 May 2016.
- SMITH, A. S., KAUSHAL, A., ODOLINSKI, K., IWNICKI, S., & WHEAT, P. (2014). Estimating the damage and marginal cost of different vehicle types on rail infrastructure: combining economic and engineering approaches.
- SMITH, A., IWNICKI, S., KAUSHAL, A., ODOLINSKI, K., & WHEAT, P. (2017). Estimating the relative cost of track damage mechanisms: combining economic and engineering approaches. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit. 231(5). Pp. 620-636.
- UNE-EN 13306:2011. Mantenimiento. Terminología del mantenimiento.
- UNE-EN 13848-1:2004+A1:2009. Aplicaciones ferroviarias. Vía. Calidad de la geometría de vía. Parte 1: Caracterización de la geometría de vía.

- UNE-EN 13848-2:2007. Aplicaciones ferroviarias. Vía. Calidad de la geometría de vía. Parte 2: Sistemas de medición. Vehículos de registro de la vía.

- UNE-EN 13848-5:2009+A1:2011. Aplicaciones ferroviarias. Vía. Calidad de la geometría de la vía. Parte 5: Niveles de calidad geométrica. Plena vía.

- WHEAT, P. & SMITH, A.S.J. (2008). Assessing the marginal infrastructure wear and tear costs for Great Britain's railway network. Journal of Transport Economics and Policy. Vol. 42. Pp. 189–224.

High Speed Rail Infrastructure

360.revista de alta velocidad
número 5 - junio - 2018. Pág 263 - 279

Calculation and rational dimensioning of railway infrastructure materials using numerical modelling

Álvarez, Fernando
Balmaseda, Lucía
Gallego, Inmaculada
Rivas , Ana
Sánchez-Cambronero, Santos

University of Castilla-LaMancha (UCLM)[1]

Abstract

In railway engineering, track-vehicle interaction is key to ensuring a certain level of track quality. In order to achieve this quality, it is necessary to carry out an exhaustive control of the deterioration of the railway, this being a consequence of the traffic and the vertical stiffness of the track. In order to obtain an optimum value of stiffness, together with the need to rationalize the sizing of the track, it is useful to study the behaviour of the railway cross-section. The study of this has forced the designers to use different tools, such as numerical models that seek to model the behaviour of the railway platform so that its design is such that the requirements of safety and comfort that are required to the rail transport. This paper shows the realization and proposal of a numerical model that seeks to reproduce as realistically as possible the behaviour of the railway platform, with a dual purpose: to become a calculation tool for the designer and that its use allows the own elaboration of design recommendations.

Keywords: High-Speed railway, Infrastructure, Numerical modelling, Elastic behaviour, Elastoplastic behaviour, Design.

---

[1] Álvarez, Fernando. University of Castilla-La Mancha (UCLM). Email: Fernando.Alvarez1@alu.uclm.es
Balmaseda, Lucía. University of Castilla-La Mancha (UCLM). Email: Lucia.balmaseda@uclm.es
Gallego, Inmaculada. University of Castilla-La Mancha (UCLM). Email: Inmaculada.gallego@uclm.es
Rivas, Ana. University of Castilla-La Mancha (UCLM). Email: Ana.rivas@uclm.es

1. Introduction

In recent years, rail transport has experienced great expansion throughout the world, becoming an efficient and competitive transport system for countries with thousands of kilometres of track. In the coming decades, it is envisaged that new and ambitious railway projects will be developed, whose technical difficulties in the interaction with the environment by which it intends to circulate the route and the reach of greater speeds on this one, make it a real challenge for civil engineering; such as the expansion of the High-Speed network in China (Abu Sayeed and Shahin, 2016) and the expected development of the High-Speed line in the United States (Fort and Fort, 2016).

With regard to High-Speed networks, the demographic and economic growth of many countries, which have such a rail network with optimum operating conditions and study experience, have led to higher speeds in order to comply the demand and / or economic objectives required of this transportation system. The increase of the speed could generate an increase in the values of the loads that are transmitted to the railway platform, being the speed proportional to the value of this one. This is related to the effect of the weight of the non-suspended masses of the train, since, if this weight does not change but increases the speed value, the dynamic effects that occur in the vehicle-track interaction, generates a dynamic overload that increases the values of the loads to be supported by the track. If we now refer to rail freight transport, future commercial expansion of several countries may lead to increased freight and / or transport capacity by trains, either by increasing the number of cars or the axle load being transmitted to the railway platform (Li, et.al., 2016). These increases in the magnitudes of the loads show that, if it is desired that this transport system correctly fulfils its function, the train must circulate by a means that guarantees safety and stability to the infrastructure and comfort for the passengers and loads to be transported.

On the other hand, awareness of the economic and landscape impact generated by the design and construction of a High-Speed lines has led to an awareness of the importance of valuing and treating the local materials of the trace, avoiding the massive waste of these through their disposal can consider the possibility of subjecting them to a treatment that improves their geomechanical behaviour (Gomes Correia, et.al., 2016), and make it fit to be usable in the construction of the railway platform.

So far, the sizing of the railway platforms has been carried out from the point of view of the experience obtained in the realization of projects, being in the case of the High-Speed, a very conservative sizing for the seat layers that form the railway platform. In the current context of economic crisis, it is necessary, in the field of exploitation and design, to introduce certain design criteria so that the cross-section of the railway is defined according to criteria of efficiency and rationalization of the same.

These three aspects; increase of the loads that request the railway platform, geomechanical characterization of recycled materials that form it, and introduction of new criteria that realize a rational and efficient design of the platform; are the key reasons for carrying out a study that has, as a final objective, the design of the railway cross-section in accordance with criteria of technical efficiency, being stable and functional in the face of high load conditions and economic efficiency, avoiding an oversizing of the same .

The purpose of this paper is to describe the elaboration of design recommendations, for which it has previously been necessary to make a numerical model from the refined other existing model, and can also serve as a future phase in the performance of a dynamic analysis.

Finally, the structure of the following article is described, being organized as follows:

First, the description the methodology adopted in the realization of said numerical model, followed by a description of the characteristics and cases of analysis considered in the mime; then the main results obtained from the resolution of the calculation by the numerical model will be presented, being these mainly values of seats and vertical tensions in the different layers.

2. Methodology

The study of the geotechnical behaviour of the railway platform is one of the most difficult problems in the field of civil engineering in terms of its resolution and interpretation of the results obtained. From the outset, this problem has been approached through the use of classical analytical solutions obtained from the use of Elasticity Theory hypotheses (Winkler, 1867; Poulos and Davis, 1974; Jimenez Salas, 1981) until the and the development of new mathematical and numerical tools (Indraratna, 2016), such as the Finite Element Method (MEF) and Discrete Elements Method (MED) (McWilliams, et.al., 2000; Huang and Tutumluer, 2011) has made it possible to solve this problem in a more precise way, allowing the addition of new variables that try to reproduce more accurately the behaviour of the railway platform.

One of the effects that computers and numerical methods had on the calculation of railway platforms was the possibility of modelling the behaviour of the materials of the platform in its elastic or plastic form together, thus integrating the model of elastoplastic behaviour; On the other hand, it was also improved the capacity of calculation before cases of loads that considered the variation in the space and time of this one, giving place to the dynamic analyses of load in railway platforms with a wide development in our days. However, these two aspects are not yet fully integrated into current numerical models, since the use of the Plasticity Theory hypotheses requires the study of complex constitutive laws and the use of additional mechanical parameters that are not required if It is assumed an elastic behaviour of the materials and therefore, many current numerical models still assume elastic or derivative behaviour models as the hyperbolic model (Gallego, et al., 2013; Shih, et.al., 2017). With respect to dynamic analysis, most of the commercial software that operate with the use of finite elements have tools that allow the modelling of mobile load cases and dynamic properties of materials, thus having a greater difficulty in adequately characterizing the geomechanical behaviour of these.

The numerical model presented here allows to model the geometry and the elements that form the section of railway platform used in the Spanish High-Speed. Different recommendations and modelling methodologies used by different have been used for its elaboration, being the most outstanding recommendations in normative series (Ministerio de Fomento, 1999) and, as numerical models, those made by Gallego (2012) and others (Gallego, et.al., 2013).

With the orientation of these existing models, it has been refined or improved in the same as will be seen, by eliminating the ballast material that confines, in the transverse and longitudinal direction, the sleepers of the platform, generating a model with confining ballast and without it. In addition to this, a deepening in the interpretation of the results that obtain numerical resolution is made, comparing these assuming both a constitutive model of elastic behaviour and elastoplastic for the granular materials that form the seat layers of the platform, all for a simple static load case.

Finally, for the calculation of the numerical model and the interpretation of the results obtained, a sensitivity analysis was carried out by varying the value in one of the resistant parameters of the granular materials that form the railway platform, these being in particular the Formation layer and Subballast.

3. Numerical model

3.1 Description

The numerical model presented here consists of a three-dimensional Finite Element model and fully parameterized, that is, it allows the user to directly modify the geometric properties that define the railway track section as well as the mechanical and geomechanical characteristics that define the behaviour of the elements of the superstructure of the track and the granular layers of the infrastructure. The tool used for the numerical modelling is commercial software ANSYS® Mechanical APDL 17.2, which requires the use of external programming codes or through the direct use of the interface.

With regard to the same, only the most relevant elements of the superstructure are modelled: rail, bearing plate, sleepers and ballast; and infrastructure layers: subballast, formation layer and embankment or subgrade. In this way, the finite element model acquires the form shown in Fig. 1, simulating half of the section of track as a single track.

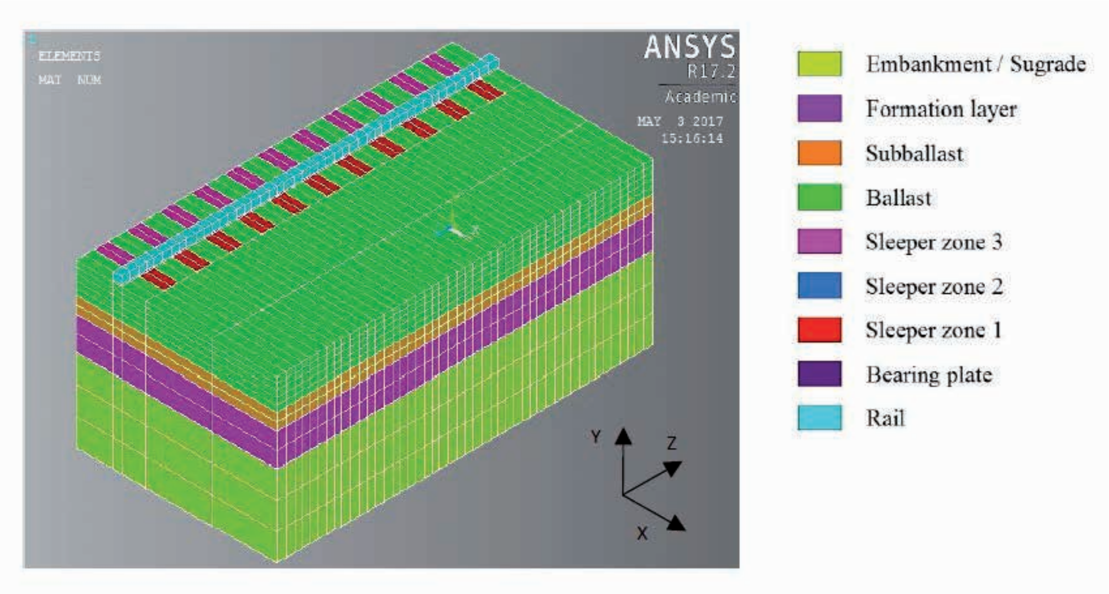


Fig. 1. General view of the numerical model and its mesh.

3.2 Geometry and material properties

For the definition of the geometry and the elements that form the section of railway platform have been used the different recommendations and methodologies of modelling collected by the different mentioned above.

With regard to the elements of the track superstructure: UIC60 rail, bearing plate and sleepers; These have been modelled with equivalent geometries of parallelepiped volumes by threedimensional elements of 20 nodes and that, together with the adjustment of the mechanical properties of materials (Gallego, 2012), it is possible to model with more accuracy the structural behaviour of these in the different situations of load.

In relation to the seat layers, these are re-modelled with parallelepiped volumes assuming that each of these behaves as a continuous medium. Further:

1. It has been assumed that all granular materials of the granular layers are homogeneous and isotropic materials where the geomechanical properties are constant throughout the volume and there is no change in the value of these.

2. In the case of ballast, by its rheology and interaction between particles, it would be more advisable to use discrete elements for its numerical modelling, in this numerical model it is modelled as a continuous medium by finite elements, paying close attention to that, if a model is adopted sufficiently adequate, it will be able to reproduce a more accurate tenso-deformational response than using simpler models (Ishikawa, et.al., 2014).

3. The geometry of the cross section has been defined following the dimensions proposed by ADIF (2006) in its different regulations for High-Speed lines, as shown in Fig. 2. In this figure you can see how it would also be configured the section of eliminate the confining ballast by refining the numerical model.

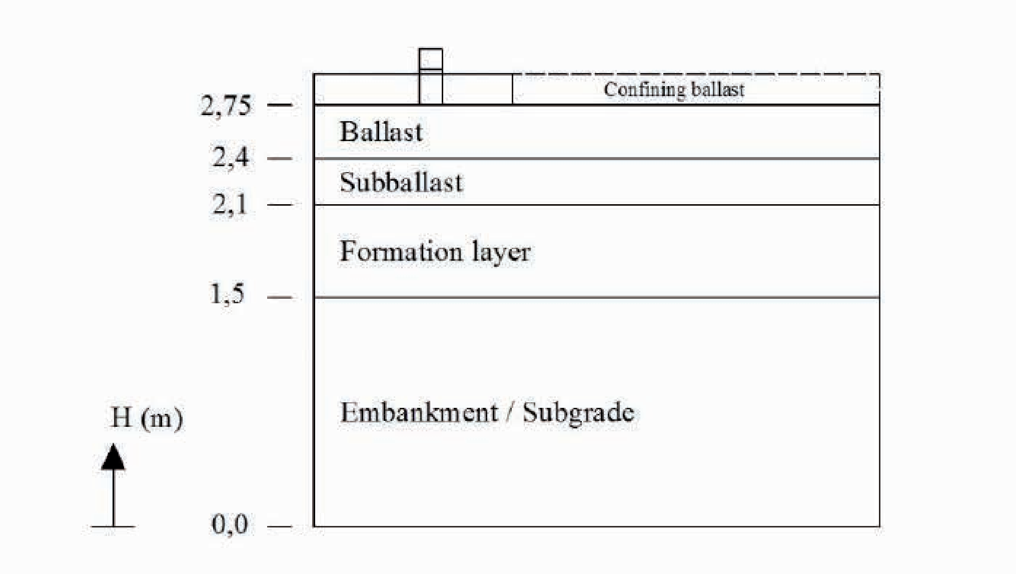


Fig. 2. Layer dimensions for the High-Speed cross section in Spain.

The mechanical and geomechanical properties of the different materials are listed in Table 1. In this it can be seen how, for the sensitivity analysis, the value of the internal friction angle has different values for two of the granular materials shown. In the case of the Formation layer there will be a decrease in the resistance and, in the Subballast the opposite. It has been assumed that all granular materials are in perfectly drained conditions, so that the tensions obtained will be in terms of effective coincident with the values of total stresses.

Table 1.	E (Pa)	v	ρ (N/m^3)	c (Pa)	Φ (°)
Rail (Steel)	2.10E+11	0.3	7500	-	-
Bearing plate	6.91E+07	0.4	1	-	-
Sleeper zone 1 (Concrete)	7.83E+10	0.25	2500	-	-
Sleeper zone 2 (Concrete)	4.90E+10	0.25	2500	-	-
Sleeper zone 3 (Concrete)	3.59E+10	0.25	2500	-	-
Ballast	1.30E+08	0.2	1900	0	45
Subballast	1.20E+08	0.3	1900	0	Φ_{SB1}, Φ_{SB2}
Formation layer	8.00E+07	0.3	2000	0	Φ_{FL1}, Φ_{FL2}
Embankment/ Sugrade	2.50E+07	0.3	2000	1.00E+04	20

3.3 Treatment of interfaces

To better model the behaviour and interaction of the different elements, especially the contact between the sleeper and the ballast, a numerical tool has been used to solve local problems in contact areas, such as the high concentration of stresses due to the existence of two materials with very different stiffness. The tool used in such software is to equalize the displacements by duplicating a node for both sides of the interface through the coupling and nodes (ANSYS, 2008), which allows to solve satisfactorily the problem of the discontinuity in the stresses and deformations and also, does not raise the computational cost in the resolution of the model.

3.4 Analysis cases

With respect to the behavioural laws that govern the behaviour of the materials, we have chosen to model and calculate the model assuming two ways: one in which it considers that all the materials of the railway platform are governed by a linear elastic behaviour and another in which, the granular materials of the platform are governed by a non-linear behaviour of elastoplastic type.

For the elastoplastic behaviour model, a perfect plasticity model has been assumed (ANSYS, 2013), where the hardening and softening effects of the material defined by the hardening parameter H' are not considered in the following expression:

$$E^{ep} = E \cdot \frac{H'}{E + H'} \tag{1}$$

where E^{ep} is the elastoplastic deformation modulus and E the Young's modulus. If we consider a null value for H', the stress-strain curve will have the form shown in Fig. 3 (Oliver and Agelet, 2002; Chaves, 2013), in which it is represented that when the material reaches the stress threshold value σ_e, its tensional states will only move within the horizontal elastoplastic branch, but if it is not reached, it will continue to behave as a linear elastic material.

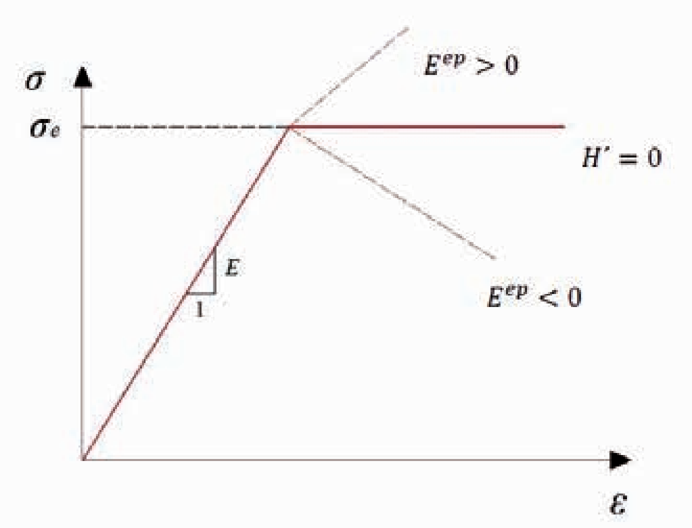


Fig. 3. Elastoplastic model assuming perfect plasticity.

For the elastoplastic model, a rupture criterion o has also to be used, namely the Drucker-Prager criterion, a generalized version of Von Mises's perfect plasticity model and also a smoothed approximation of the Mohr-Coulomb model (Jiménez Salas, 1981; Chaves, 2013), see Fig. 4.

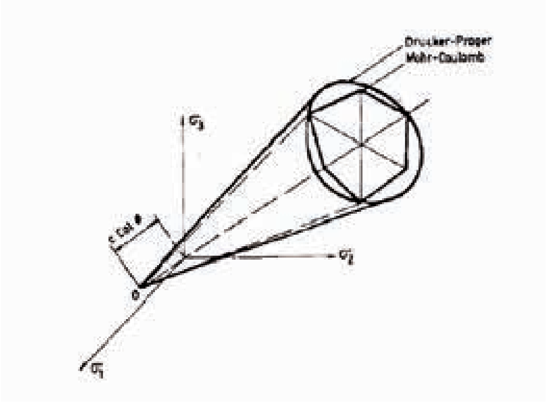


Fig. 4. Conceptual representation of the creep surface for the frictional elastoplastic models of Mohr Coulomb and Drucker-Prager (Gens and Potts, 1988). Note: The correct value of c·cotf is c $\sqrt{3}$ cotf (Chaves, 2013).

The load to be applied in the numerical model consists of a static load obtained from the weight of the load values per axis corresponding to Spanish train models used in High-Speed, which reaches an equivalent value of 186.4 kN per axis (Gallego, 2012). With respect to the purely elastic model, only one loading state corresponding to the application of the load □ on the platform will be applied, whereas for the elastoplastic model, depending on the tensions and deformations of the load history, three states of load: one corresponding to the self-weight (LS1), another one due to the application of the load Q (LS2) as in the elastic model and, last one where the discharge (LS3) is considered.

With all these aspects, the models under study have been elaborated according to the following criteria:

1. All models will maintain the constitutive model of elastic behaviour for the Rail, Bearing plates and Sleepers due to their high stiffness.
2. The comparison between elastic and elastoplastic behaviour models will only refer to granular materials of the railway section.
3. The sensitivity analysis will be performed by varying the value of the internal friction angle corresponding to the Subballast and the Formation layer.

Table 2 shows, in summary, all the models or cases of analysis to be analysed, according to the behaviour of the granular materials, the existence or not of confining ballast and the episodes of load to be considered.

Table 2. Summary of characteristics of the models to be studied.

Model	Constitutive model-Infrastructure later	Confining ballast	Load step
#1	Elastic	Yes	LS2
#2	Elastic	No	LS2
#3	Elastoplastic (Φ_{FL1}, Φ_{SB1})	Yes	LS1+LS2+LS3
#4	Elastoplastic (Φ_{FL1},Φ_{SB1})	No	LS1+LS2+LS3
#5	Elastoplastic (Φ_{FL2}, Φ_{SB1})	Yes	LS1+LS2+LS3
#6	Elastoplastic (Φ_{FL2}, Φ_{SB1})	No	LS1+LS2+LS3
#7	Elastoplastic (Φ_{FL1}, Φ_{SB2})	No	LS1+LS2+LS3
#8	Elastoplastic (Φ_{FL2}, Φ_{SB2})	No	LS1+LS2+LS3

4. Results

4.1 Adjustment model

During the resolution of the numerical model that considered the linear elastic behaviour of all the materials of the platform, it was detected that the vertical compression stresses in the ballast under the loaded sleeper reach values well below the usual ones of 100 to 120 kPa, obtained in experimental observations performed for High-Speed sections in Spain (Gallego, et. al., 2013). This could be because, considering the ballast as a linear elastic material, makes the ballast elements located above the support plane of the crossbeam oppose the compressions that occur in said plane, generating tensions that They decrease the value of the compressions that are given in the sleeper-ballast interface. To solve the problem we defined the following strategies to consider for the numerical model:

1. Eliminate the ballast that confines the sleepers in the longitudinal and transverse direction of the track, see Fig. 2.
2. Assume in the ballast a law of elastoplastic behaviour, without cohesion, to avoid the appearance of tractions in the same.

The second of the strategies did not make sense since it was contemplated to compare the elastic model with the elastoplastic. Therefore, it was decided to apply the first of them to check if really, removing the top layer would influence the vertical stresses that occurred in the model. Fig. 5 shows the different steps that were performed in such a strategy, where compression stresses has negative sign and tractions stresses have positive sign: Fig. 5a shows the platform model with confining ballast that assumes linear elastic behaviour for all materials, being able to observe traction stresses in the area next to the sleeper where the load is applied and the adjacent ones, above the support plane with the ballast, and in areas close to the limits in the numerical model domain; If we eliminate the confining ballast and continue assuming an elastic behaviour, Fig. 5b, the traction stresses in the area next to the loaded sleeper disappear above the support plane of the sleeper, where they no longer oppose the compressions and they increase their value below the loaded sleeper but, in the lower and top zone to the support plane of the adjacent sleepers, the tractions continue to persist and continue to the limit of the model, possibly by a “distortion” of the numerical model; Considering now to assume an elastoplastic behaviour for the granular materials together with the presence of confining ballast in the sleepers, we have what is shown in Fig. 5c, where now next to the loaded sleeper, despite having considered a cohesion value (c = 0), tractions appear above the contact plane of the sleeper, possibly due to the ballast that confines the sleepers, whereas in the adjacent sleepers, the tractions persist below and above the contact plane due to arrange nodes coupled to the cross-ballast contact and are eliminated completely in the limits of the model when considering a value of cohesion null for the ballast; Finally, the elimination of the confining ballast now results in what is shown in Fig. 5d, where now the tractions disappear completely next to the sleeper where the load is applied and in the limits of the model, nevertheless they continue present In the plane of contact of the sleepers adjacent to the loaded one due to the coupled nodes.

In view of these results, the following can be stated:

1. The occurrence of tractions alongside the loaded sleeper, above the support plane, is present in all cases where confining ballast is available in the sleepers, decreasing the value of the compressions in the underlying layers. These disappear by mistrusting the sleepers in all cases.
2. The occurrence of tractions under the contact plane in the sleepers adjacent to the loaded sleeper is a consequence of the use of the coupled nodes. Considering a zero cohesion for the ballast and that there is no ballast that confines the sleepers, makes them decrease in the elastoplastic models with respect to the elastics.
3. The tractions in the limits of the numerical model are due to a "distortion" of the results as a result of finite element modelling and, they disappear only when considering null cohesion in the ballast.

Based on these observations, it was decided to analyse and compare the results obtained by the elastic and elastoplastic models that better simulated the tenso-deformational behaviour, which were the ones that did not have confining ballast, Models #2, #4, #6, #7 and #8 defined in Table 1, remaining the others discarded when not fulfilling this condition.

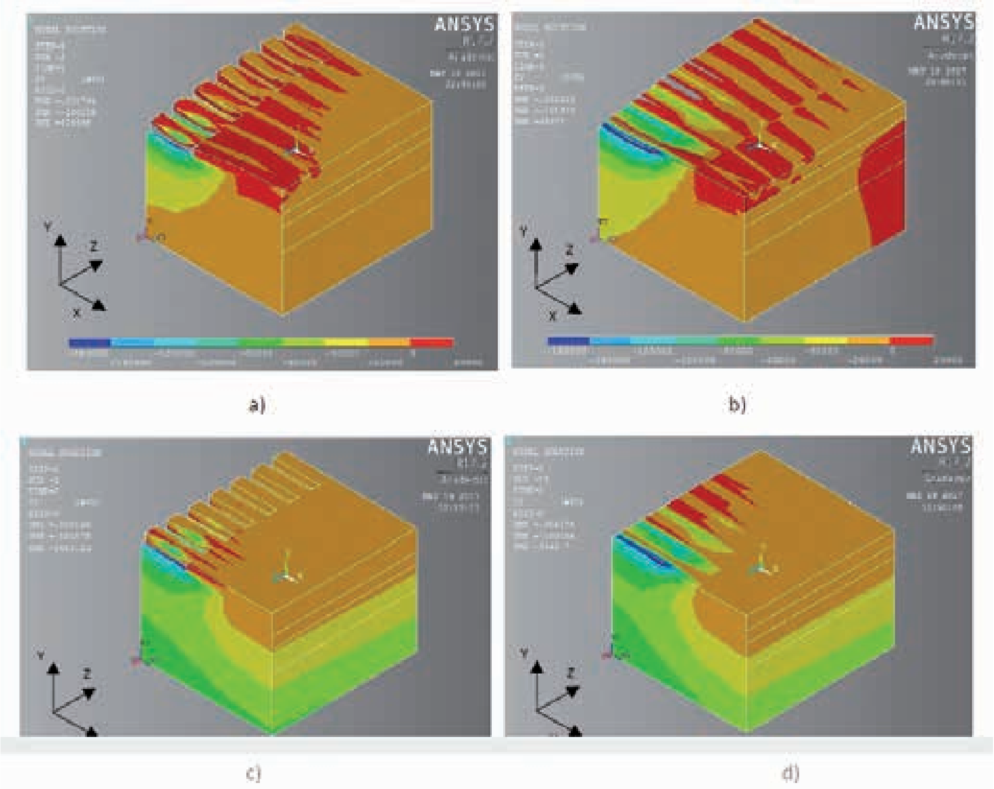


Fig. 5. Evolution of the numerical model readjustment. (a) Elastic model with confining ballast; (b) Elastic model without confining ballast; (c) Elastoplastic model with confining ballast; (d) Elastoplastic model without confining ballast. Values in Pa.

4.2 Compressions in each layer

Table 3 shows the vertical compressive values below the sleeper where the load is applied, for each of the layers of the model assuming a linear elastic behaviour for all materials (Model #2), generating a total seat δ of 2.38 mm obtained from the sum of the compressions in each layer ρ_i, such that according to Fig. 6:

$$\delta = \sum_{i=1}^{n} \rho_i \qquad (2)$$

Table 3. Compressions in each assembly and layer in the elastic model (Model #2).

	ρ_i (mm)	% respect to δ
Rail-Bearing plate-Sleeper	-0.48	20.04 %
Ballast	-0.24	10.22 %
Subballast	-0.15	6.16 %
Formation layer	-0.31	12.97 %
Embankment/Subgrade	-1.21	50.61 %

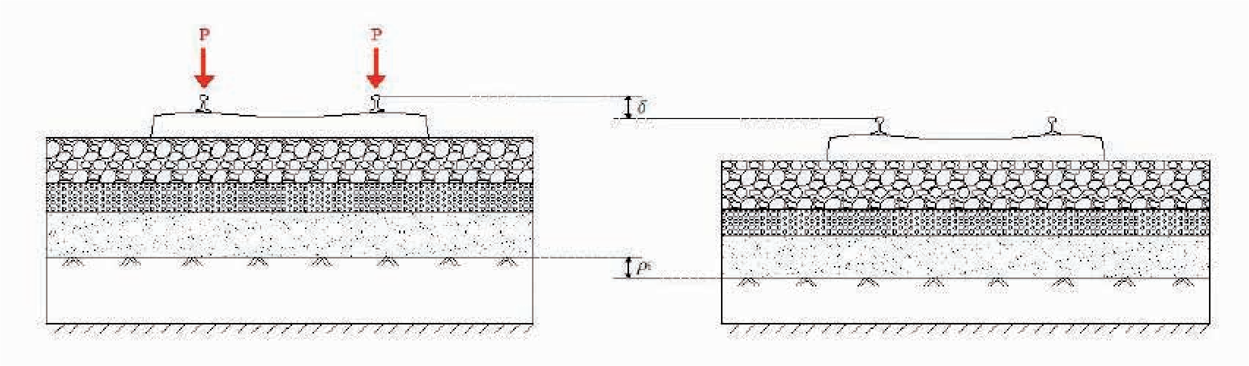


Fig. 6. Vertical settlement of the track and compression of each layer.

As expected, the Embankment or Subgrade that serves as the base to the rest of layers and elements, is the layer that has greater compression due to its smaller modulus of deformation compared to the rest of layers. After this, the Rail-Sleeper-Bearing plate assembly is the one with the greatest seat, because the seat plate has a much smaller modulus of elasticity than those corresponding to the rail and sleeper.

If we now compare the compression values in each of the layers that form the railway platform for all the elastoplastic models studied, see Table 4, we see that there is a redistribution of the values of vertical compressions and in the percentages with respect to the total seat. The RailSleeper-Bearing plate assembly and Ballast, even improving the Subballast resistance or decreasing that of the Formation layer, continue to have the same behaviour when returning to have the same values of compressions. As for the Subballast, we see that clearly, varying its angle of friction does not influence significantly since the values of compressions between the four elastoplastic models are very similar.

One of the main differences if we compare the results between the elastic model and the elastoplastics is to see that the settlement values in the elastoplastics are higher. This is due to the consideration of the sum of the elastic and plastic components that can now be given, which causes the total deformation to be defined by:

$$\varepsilon = \varepsilon^{e} + \varepsilon^{p} \qquad (3)$$

where ^{e} is the elastic deformation of the element and ^{p} the plastic deformation of the element. This is important to consider since the railroad is subject to numerous load cycles during its useful life, which can result in an increase of the plastic or irreversible deformation component and generate permanent deformations that can compromise the safety of the track.

Table 4. Compressions and vertical settlement in each assembly and layer in the each elastoplastic model.

	Model #4 (Φ_{FL1} = 35° ; Φ_{SB1} = 35°)		Model #6 (Φ_{FL2} = 20° ; Φ_{SB1} = 35°)	
	ρ_i (mm)	% respect to δ_1	ρ_i (mm)	% respect to δ_2
Rail-Bearing plate-Sleeper	-0.46	18.11 %	-0.45	16.61 %
Ballast	-0.24	9.45 %	-0.22	8.12 %
Subballast	-0.16	6.30 %	-0.17	6.27 %
Formation layer	-0.34	13.39 %	-0.48	17.71 %
Embankment/Subgrade	-1.34	52.76 %	-1.40	51.66 %
	Model #7 (Φ_{FL1} = 35° ;Φ_{SB2}= 45°)		Model #8 (Φ_{FL2} = 20° ; Φ_{SB2} = 45°)	
	ρ_i (mm)	% respect to δ_1	ρ_i (mm)	% respect to δ_4
Rail-Bearing plate-Sleeper	-0.46	18.25 %	-0.45	16.79 %
Ballast	-0.24	9.52 %	-0.22	8.21 %
Subballast	-0.15	5.95 %	-0.16	5.97 %
Formation layer	-0.34	13.49 %	-0.47	17.54 %
Embankment/Subgrade	-1.33	52.78 %	-1.39	51.87 %

δ_1 = 2.54 mm in Model #4
δ_2 = 2.71 mm in Model #6
δ_3 = 2.52 mm in Model #7
δ_4 = 2.68 mm in Model #8

The vertical settlement of the track is a parameter that allows the calculation of the vertical stiffness of the track K, defined by the following expression (López Pita. 2001):

$$\mathrm{K} = \frac{Q}{\delta} \tag{4}$$

This parameter is important to consider by the designer in any project of design and maintenance of the railway platform as this function of the mechanical properties that occur in the elements of the superstructure and layers of railway infrastructure. A low vertical stiffness value would mean having a flexible path with little dissipation energy and a bending deformation of the ballast that would generate its degradation by abrasion whereas, a high value of vertical stiffness would lead to increase the dynamic force at the wheel-track interface, transmitting a greater load to the sleeper-ballast interface and may cause deterioration and fatigue of the track elements (Woodward, et.al., 2014). Table 5 shows the vertical stiffness values obtained by applying the above expression for the value of the load defined in this analysis, in which we can see how the elastic model (Model #2) has a higher value than the others. For the fact of having a lower seat value with respect to the others, leading to a value greater than could occur in reality if we assume that an elastoplastic model reproduces reality better.

Table 5. Vertical stiffness values for each studied model.

Model	Pointed load (kN)	Vertical settlement δ (mm)	Vertical stiffness K (kN/mm)
#2	186.4	2.38	78.32
#4	186.4	2.54	73.39
#6	186.4	2.71	68.78
#7	186.4	2.52	73.88
#8	186.4	2.68	69.43

4.3 Vertical stresses in each layer

After calculating the displacements for the elements of the model, the values of the stresses in them can be calculated from the following constitutive relation (Chaves, 2013):

$$\{\sigma|(x,y,z)\} = [\mathrm{D}]\{\varepsilon\,(x,y,z)\} \qquad (5)$$

where $\{\sigma\}$ is the stress vector, $\{\ \}$ is the strain vector and $[\mathrm{D}]$ is the constitutive matrix which contain the mechanical properties of the material. In the case that we consider a calculation with elastic behaviour of the material, matrix $[\mathrm{D}]$ will only consider this behaviour model through the parameters that define it; On the other hand, if we consider an elastoplastic behaviour, this matrix is modified by considering the plastic contribution in the calculation of the solution, which according to Jiménez Salas (1980), the new constitutive matrix $[\mathrm{D}']$ is given as the difference:

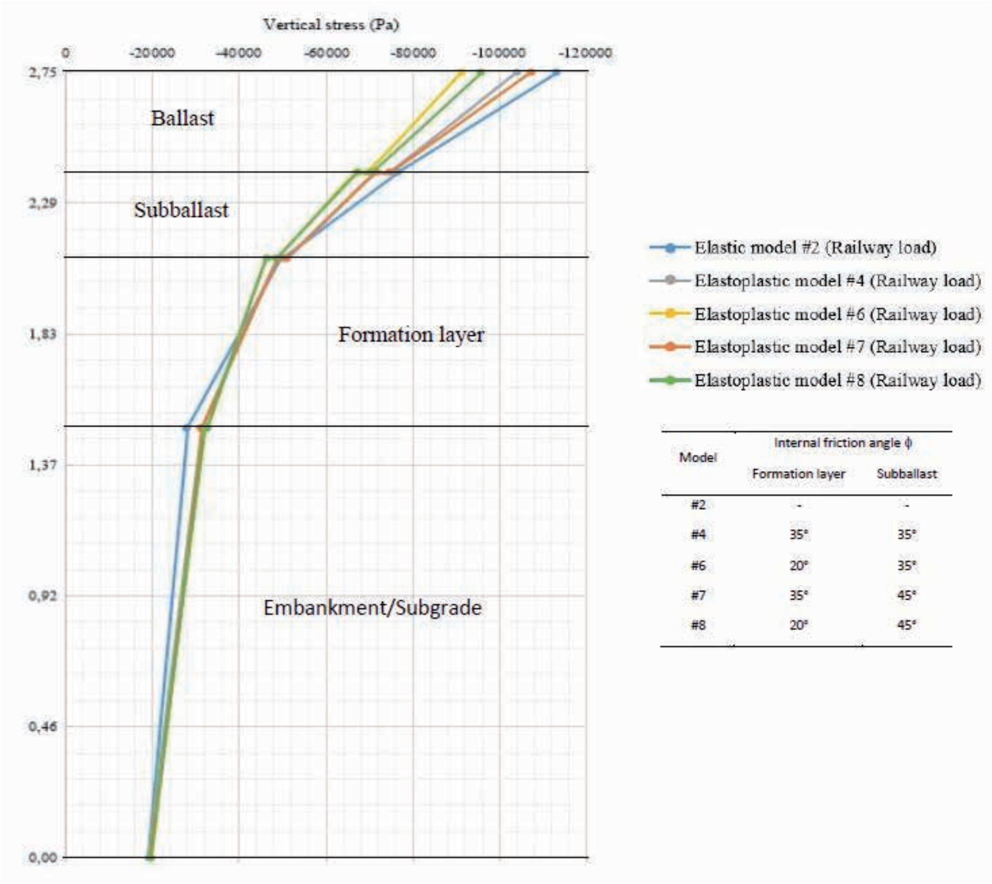
$$[\mathrm{D}'] = [\mathrm{D}^{\mathrm{e}}] - [\mathrm{D}^{\mathrm{p}}] \qquad (6)$$

where $[\mathrm{D}^{\mathrm{e}}]$ is the elastic constitutive matrix, $[\mathrm{D}^{\mathrm{p}}]$ is the plastic constitutive matrix, and $[\mathrm{D}']$ is the elastoplastic constitutive matrix. The different calculation methodology between the elastic and elastoplastic analysis has been seen in a significant way in the calculation of the previous displacements. This difference will also affect the calculation of tensions according to the type of calculation that is used according to the type of behaviour defined for the material.

From the point of view of design, the knowledge and determination of the profile of vertical stresses in the railway section is important to know the values that are achieved in the different materials of the section. These values of stresses obtained by the calculation allow us to see if the material reaches the maximum admissible stress, obtained experimentally, that if reached, could give rise to some of the most common geotechnical problems in railways (Li, et.al., 2016), possibly due to an inefficient dimensioning of the section or the use of inappropriate materials

Fig. 7 shows the profiles of vertical stresses that develop throughout the entire height of the railway platform below the sleeper where the load Q is applied.

1. Firstly, it is verified that by making the Subballast stiffen by increasing the value of its friction angle ¤ (Models #7 and #8), it leads to a greater stress absorption and, therefore, the stress profile moves to the right, towards higher stress values.
2. With regard to the Formation layer:
   - If it has a friction angle of 35° constant and the Subballast of 35° (Model #4) and it changes to 45° (Model #7), the differences in the vertical stress level are negligible.
   - In the same way, it happens in the case of having a constant value 20° in the Formation layer of form and to re-vary that of the Subballast (Models #6 and #8).
3. On the other hand, a constant value of the internal friction angle of 35° in the Subballast and with a variation in the Formation layer (Models #4 and #6) makes a significant influence of vertical stress difference in the layers of Ballast and Subballast. The same occurs when a constant value of friction angle is set in the Subballast of 45° and the shape layer (Models #7 and #8) varies with the same difference of vertical stresses in the two upper layers.
4. In the contact Formation layer-Embankment at the height of 1.5 m. it is observed that the vertical stresses are equal for all elastoplastic models until reaching the 0 or base of the model.



Model	Internal friction angle ϕ	
	Formation layer	Subballast
#2	-	-
#4	35°	35°
#6	20°	35°
#7	35°	45°
#8	20°	45°

Fig. 7. Profile of vertical stresses due to the application of the railway load for all the models studied.

5. Conclusions

From the point of view of numerical modelling, it is appropriate to model the materials of the superstructure with a linear, homogeneous and isotropic elastic behaviour as these have a high modulus of elasticity compared to that of the granular layers. On the other hand, it is more advisable to carry out analyses that consider the elastoplastic behaviour of the materials, especially if they are granular, since the remaining or irreversible deformations are important because they can cause seats that can affect the longitudinal and transversal profile of the track.

As for the results obtained with the numerical model we can clarify that regarding the use of behavioural models. The use of an elastoplastic model leads to higher values of displacement due to the appearance of the plastic component of deformation besides that the use of an elastic model leads to an overestimation of the tensions with respect to those that would have in reality if an elastoplastic behaviour were assumed.

Finally, based on the analysis of all the previous results, some design recommendations can be used that can be used by the designer in the design of the railway platform:

1. If the Embankment or Subgrade has a high height, it is advisable to use an elastoplastic behaviour model due to the high numerical error that would be in the calculation of seats when considering an elastic model. If it has a smaller height, it is useful to use a model of elastic behaviour and. more if Formation layer is not available.
2. With poor materials or low resistance it is advisable to use an elastoplastic analysis against an elastic one due to the influence that would have in terms of seats and tensions, the layers of the infrastructure in their joint interaction.
3. The use of an elastic model will lead to an underestimation of seats that would be in reality, assuming that the railroad is more rigid than it really is.
4. Increasing the angle of friction of the Subballast from 35° to 45° does not imply any relevant changes in the values of the tensions and vertical seats for the rest of materials.

6. Notation

The following symbols are used in this paper	
H' = hardening parameter	ε^p = plastic deformation
E^{ep} = modulus of elastoplastic deformation	K = vertical track stiffness
E = modulus of elasticity	$\{\sigma\}$ = stress vector
c = material cohesion	{ } = strain vector
Φ = internal friction angle	[D] = constitutive matrix
ρ_i = vertical compression of each layer	$[D^e]$ = elastic constitutive matrix
ε = deformation of the element	$[D^p]$ = plastic constitutive matrix
ε^e = elastic deformation	$[D']$ = elastoplastic constitutive matrix

7. References

- ABU SAYEED, M. and SHAHIN, M.A. (2016). *Three-dimensional numerical modelling of ballasted railway track foundations for High-Speed trains with special reference to critical speed. J Trans*p Geotech 6(8), pp. 55-65.
- ADIF (Administrador de Infraestructuras Ferroviarias). (2006). *Instrucciones y Recomendaciones para redaccion de Proyectos de plataforma-Geotecnia y Obras de tierra.* Madrid, Spain.
- ANSYS INC. (2008). ANSYS user´s manual. Canonsburg. PA. USA.
- ANSYS INC. (2013). ANSYS Mechanical APDL Theory Reference .Canonsburg. PA. USA.
- CHAVES, E.W.V. (2013). *Notes on Continuum Mechanics. International Center for Numerical Methods in Engineering (CIMNE),* Barcelona, Spain.
- FORT, L. and FORT, C. (2016). *España y la red de alta velocidad en Estados Unidos.* Revista de Obras Públicas 163(3580), pp. 36-53.
- GALLEGO, I. (2012). *Heterogeneidad resistente de las vías de alta velocidad: Transición terrapléneestructura.* Editorial Académica Española, Madrid, Spain.
- GALLEGO, I., MUÑOZ, J., SANCHEZ-CAMBRONERO, S., RIVAS, A. (2013). *Recommendations for Numerical Rail Substructure. Modeling Considering Nonlinear Elastic Behavior.* J Transp Eng 139(8), pp. 848–858.
- GENS, A. and POTTS, D.M. (1988). *Modelos elastoplásticos de estado crítico para análisis numéricos de problemas geotécnicos. I- Formulación básica y principales modificaciones*. Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería 4, pp. 497-522.
- GOMES CORREIRA, A., WINTER, M.G., PUPPALA, A.J. (2016). *A review of suitainable approaches in transport infrastructure geotechnics*. J Transp Geotech 7(12), pp. 21-28.
- HUANG, Y.H. and TUTUMLUER, E. (2011). *Discrete element modeling for fouled railroad ballast.* J Constr Build Mater. 25(20), pp. 3306-12.
- INDRARATNA, B. (2016). 1st Proctor Lecture of ISSMGE: *Railroad performance with special reference to ballast and substructure characteristics.* J Transp Geotech 7(4), pp. 74-114.
- ISHIKAWA, T., MIURA, S., SEKINE.,E. (2014). *Simple plastic deformation analysis of ballasted track under repeated moving-wheel loads by cumulative damage model.* J Transp Geotech 1(4), pp. 157-170.
- JIMÉNEZ SALAS, J.A. (1981). Geotecnia y Cimientos II. Rueda, Madrid, Spain.
- JIMÉNEZ SALAS, J.A. (1980). Geotecnia y Cimientos III. Rueda, Madrid, Spain.
- LI, D., HYSLIP, J., SUSSMANN, T., CHRISMER, S. (2016). *Railway Geotechnics.* CRC Press/ Taylor & Francis Group, Boca Raton, FL, USA.
- LÓPEZ PITA, A. (2001). *La rigidez vertical de la vía y el deterioro de las líneas de Alta Velocidad.* Revista de Obras Públicas 148(3415), pp. 7-26.
- MCWILLIAMS, P., FERNÁNDEZ MERODO, J.A., PASTOR PÉREZ, M., NASARRE GOICOECHEA, J., CARRILLO, J. (2000). *Aplicaciones del Método de Elementos Finitos a la ingeniería ferroviaria.* Revista de Ingeniería Civil 118, pp. 71-82.
- MINISTERIO DE FOMENTO. (1999). *Recomendaciones para el proyecto de plataformas ferroviarias*. Madrid, Spain.

- OLIVER, X. and AGELET, C. (2002). Mecánica de medios continuos para ingenieros. UPC publications, Barcelona, Spain.
- POULOS, H.G. and DAVIS, E.H. (1974). *Elastic Solutions for Soil and Rock Mechanics.* John Wiley and Sons, New York, USA.
- SHIH, J.Y., THOMPSON, D.J., ZERVOS. A. (2017). *The influence of soil nonlinear properties on the track/ground vibration induced by trains running on soft ground.* J Transp Geotech 1(11). pp. 116.
- WINKLER, E. (1867). *Die Lehre von Elastizität und Festigkeit.* Prague.
- WOODWARD, P.K., KENNEDY, J., LAGHROUCHE, O., CONNOLLY, D.P., MEDERO, G. (2014). *Study of railway track stiffness modification by polyurethane reinforcement of the ballast.* J Transp Geotech 1(4), pp. 214-224.

8. Acknowledgements

Author would like to acknowledge the support of teachers and staff of the Department of civil engineering and building in the School of Civil Engineering of the University of Castilla-La Mancha as well as the Enrique Castillo Institute for the financial support necessary for the realization of this work.

High Speed Rail Infrastructure

360.revista de alta velocidad
número 5 - junio - 2018. Pág 281 - 299

Reducing High-Speed Rail Costs by Combined Double-Single Tracks

Grande, Zacarías
Torralbo, Julia
Lobera, José Manuel
Sánchez-Cambronero, Santos
Castillo, Enrique

University of Cantabria [1]
Universidad de Castilla-La Mancha

Abstract

This paper aims to demonstrate the advantages of the Alternate Double-Single track (ADST) solution with respect to the traditional double track alternative for railway line design. The paper starts with an introduction and a summary along with a description of the ADST approach and its main advantages. For illustration purposes two real cases, which use this procedure, have been described.

The Santander-Bilbao and the Vitoria-Zaragoza line proposals are analysed in some detail showing the important savings and performances when compared with the double and single track solutions. Finally, some conclusions and recommendations are given.

Keywords: Railway line design, rational investment analysis, optimization, rail capacity, relative travel time

---

[1] Grande, Zacarías. University of Cantabria. Email: Zgandrade@gmail.com
Torralbo, Julia. University of Cantabria. Email: julia.torralbo@alumnos.unican.es
Lobera, José Manuel.University of Cantabria. Email: jose-manuel.lobera@alumnos.unican.es
Sánchez-Cambronero, Santos. Universidad de Castilla-La Mancha, UCLM. Santos. Email: Sanchez@uclm.es
Castillo, Enrique. University of Cantabria. Email: castie@unican.es

1. Introduction and motivation

The design of high-speed railway lines could have a new alternative thanks to new methodologies that offer important savings in construction and maintenance costs with no practical losses in travel times.

This case study focuses on an alternate double-single track (ADST) methodology (Castillo, et al., 2015). The main idea behind ADST consists of using single track where the infrastructure is very expensive (tunnels and viaducts) and double track where it is cheaper (smooth orography), if it is necessary.

The ADST methodology is especially suitable for peripheral sections where demand forecast is low or intermediate. A double-track solution in these situations could lead to oversized lines with inefficiency in exploitation and negative social impact investment financial returns. On the other hand, the single-track would not satisfy passenger demand. The ADST performance in peripheral lines is much closer to double- than to single-track performance, whereas the ADST cost is much closer to single- than to double-track cost.

From the previous paragraph, the primacy of considering traffic volumes in rail design could be deduced. An estimation of the demand can determine which the right solution for each case study is: double-track line (high demand) or ADST line (low or intermediate demand).

The main tool required to develop an ADST line is an optimization program that allows us to compare different track combinations and permits us to find the optimal sequence of single- and double-track segments. Thus, construction costs are drastically reduced (up to 40%) and maintenance costs are also substantially reduced (up to 50%) (Castillo, et al., 2015). Some interesting publications related to the optimization of timetables are: (Amit & Goldfarb, 1971) , (Burdett & Kozan, 2010), (Cacchiani & Toth, 2012), (Caprara, et al., 2002), (Carey & Crawford, 2007), (Castillo, et al., 2015), (Castillo, et al., 2011), (Castillo, et al., 2009), (Castillo, et al., 2016), (D'Ariano, et al., 2007), (Pachl, 2014), (Sahin, 1999), etc.

Train routing and other optimization problems have been dealt with in (Assad, 1980), (Carey, 1994), (Carey & Lockwood, 1995), (Cordeau, et al., 1998), (D'Ariano & Pranzo, 2004), (Haghani, 1987), (Hellström, 1998), (Lin & Ku, 2013), (Ouyang, et al., 2009), (Petersen, et al., 1986), (Yang & Hayashi, 2002), etc.

In the same way, the timetable must be optimized in order to reduce travel time. Since travel times of different trains circulating along the network or line could be very different, and the impact of a five-minute delay on a one-hour trip is not the same than on a three-hour trip, the program uses relative travel times.

The relative travel time is the quotient:

$$Relative\ travel\ time\ (RT) = \frac{Travel\ time}{Travel\ time\ at\ maximum\ speed} \qquad (1)$$

Then, a relative time 1 means that we travel at maximum speed; contrary a relative time value of 1.10 or 1.20 means that we have been used for the trip a 10% or a 20% more time, respectively.

Delaying or advancing the departure or arrival time without changing the total travel times is achieved forcing the trains to cross inside double-tracked segments.

The design and management of an ADST line is complex, because it requires:

1. Deciding which segments should be constructed in single track and which in double track.
2. Satisfying the safety and timetable constraints of the different services with the aim of obtaining small travel times when we have a single track in some segments.
3. Minimizing costs and travel times and optimize the infrastructure usage.
4. Obtaining all rail timetables of the whole network.

Due to the complexity of the problem, the use of an optimization program is necessary in order to satisfy all the imposed safety and service conditions.

This paper aims to introduce two case studies that clearly show the benefits of using this methodology. Finally, some conclusions will be drawn.

2. Case studies

The case studies used in this paper correspond to the corridors Santander-Bilbao and VitoriaZaragoza (Spain). For each case, the following procedure will be applied: (1) a diagnosis of the current situation will be described; (2) the inputs and outputs of the program will be outlined; and (3) the adopted solution will be justified by carrying out a multi-criteria analysis.

2.1 Santander-Bilbao case

2.1.1 Current line

The existing rail line between the cities of Santander and Bilbao (see Figure 1), in northern Spain, is obsolete compared to modern transportation. Due to the inefficiency of this means of transport, displacements between these two cities are mainly done by private vehicle or bus.

Figure 1. Actual road and railway connections between Santander and Bilbao.

The population of the metropolitan areas of these two capital cities combined comes to more than one million citizens, being Bilbao the most populated one.

This High speed railway line could be part of a possible 'Cantabrian Corridor', from Galicia to the French border, improving connections with Europe. This fact addresses the significance of this infrastructure.

However, the complex orography, characteristic of the North of Spain, makes the construction of a high-speed rail line difficult. In spite of the huge social benefits that a high-speed train could mean, construction costs could be excessively high.

For this reason, there is a strong need for a new solution that offers a considerable reduction in construction and maintenance costs without a great impact on travel times.

After conducting a demand analysis, it has been estimated a demand of approximately 1.1 million passengers in the first year travelling between these two cities.

2.1.2 High speed line proposal

All the factors discussed in the previous section make the alternate double-single track (ADST) line the best solution for the Santander-Bilbao corridor. The features of the proposed line include mixed traffic, Iberian-gauge track (1,668 mm) and a design speed of 250 km/h.

The layout of the line, as depicted in Figure 2, consists on an inland itinerary pursuing a straight line. In order to reduce costs, the existing infrastructure nearby the cities of Santander and Bilbao is used.

Figure 2. Existing railway line and initially proposed HSR line between Santander and Bilbao.

Trying to minimize the environmental impact of this infrastructure, the initial layout has been modified to avoid (see Figure 3): (1) inappropriate land uses; (2) human settlements with more than 100 inhabitants; (3) Site of Community Importance (SCI) in a radius of 1 km; and (4) Protected Areas.

Hence, the ADST line proposal meets the criteria in most of this path as depicted in Figure 4 except for:

- Santander and Bilbao accesses, where the use of the current railway platform is proposed.
- In the vicinity of Bilbao, where potential environmental impacts emerge to coniferous forests. Consequently, to mitigate its impact, tunnels in this area are suggested.

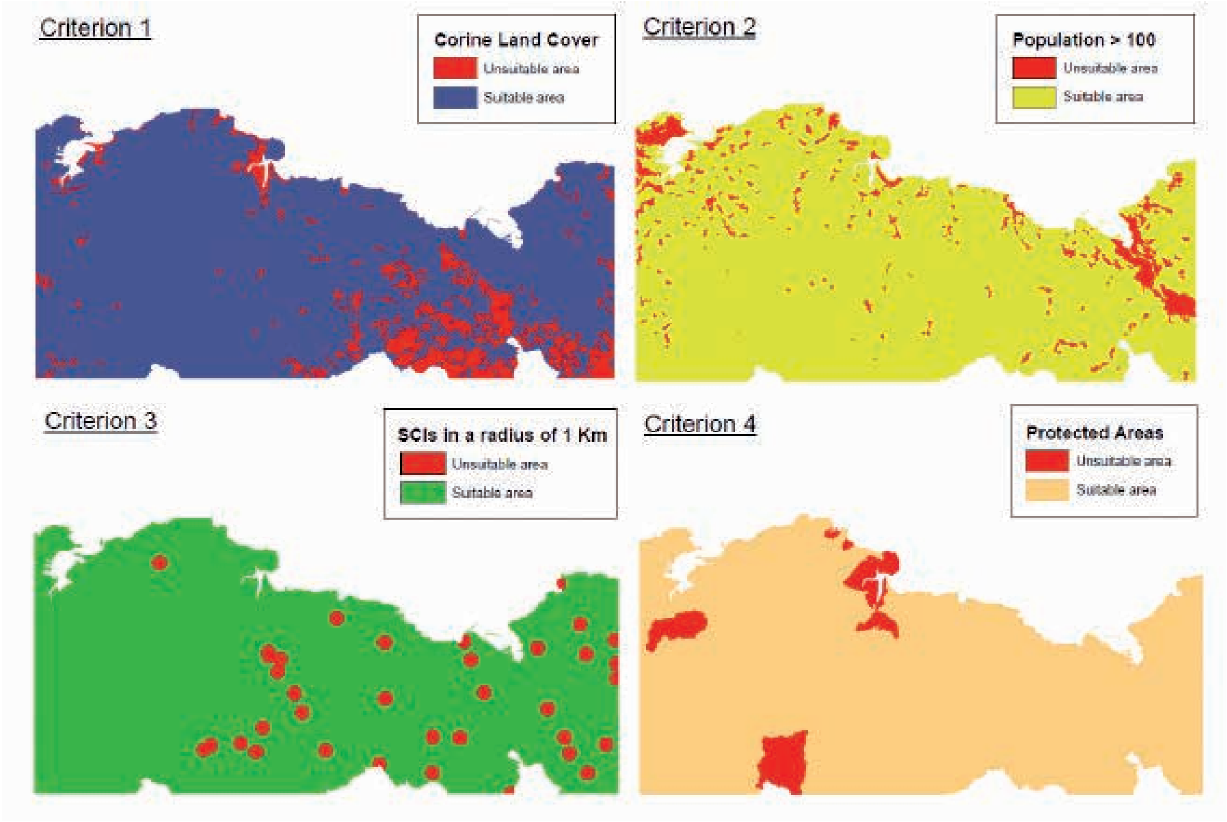


Figure 3. Restricted areas imposed by environmental and other type of constraints.

Figure 4. Final proposal for the Santander-Bilbao line after considerations of required constraints.

The final layout has been modelled in Autocad Civil 3D, meeting the Spanish technical requirements for high-speed railway tracks. The final result consists of a total length of 90.9 km whose 68.18 km are brand new. Due to the complex terrain and the stringent geometrical requirements of this type of lines, the layout includes 28.3% of the total length in tunnel and 13% in viaduct (see Figure 5 and Figure 6).

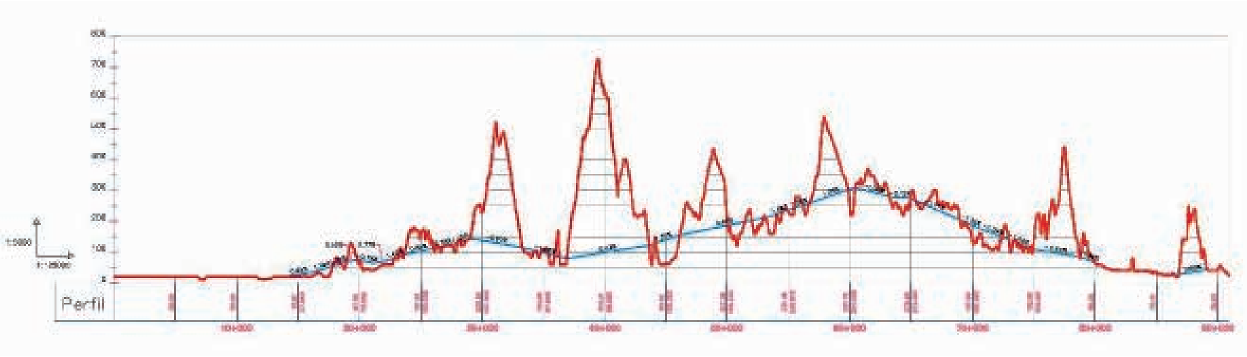


Figure 5. Elevation profile of the proposed HSR line.

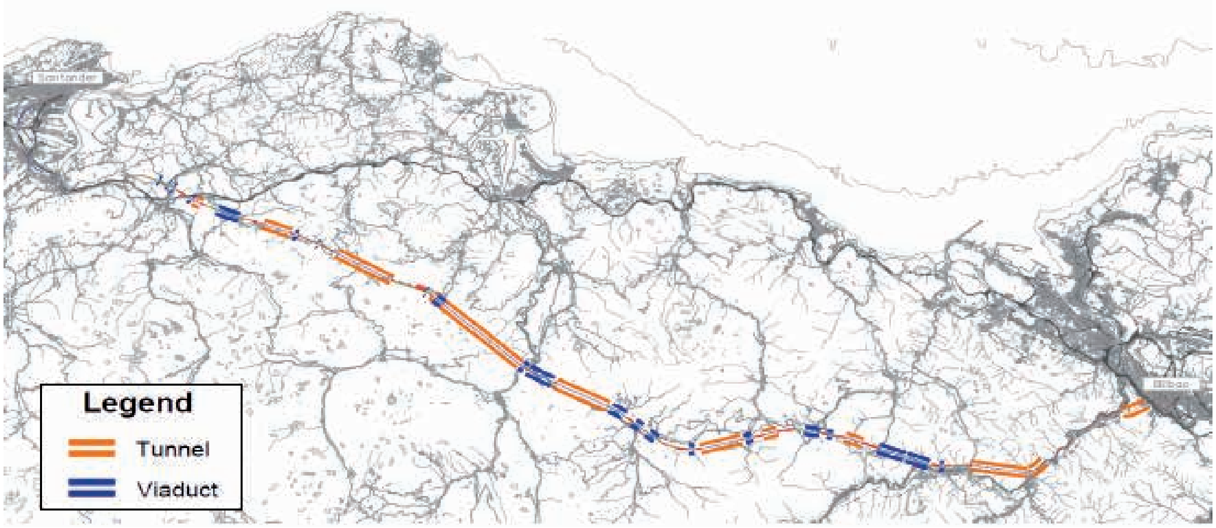


Figure 6. Detailed line layout with the exact tunnel and viaduct locations.

2.1.3 Railway line optimization

To optimize the railway network, the first step for the exploitation study consists of dividing the line into segments of similar characteristics. This is needed in order to obtain an estimation of the cost per kilometre of each segment, in both single- and double-track (see Table 1).

Figure 7. Segments considered in the optimization program. Santander-Bilbao Case

Table 1 Data used for the Santander-Bilbao HSR line, including the segment definition, their lengths and costs in single and double track.

	Segment		Length		Cost (M€/km M£/mi)			
Id	**Origin**	**Destination**	**Km**	**Miles**	**Simple**		**Double**	
1	Santander	Heras	14.0	8.7	1.29 €	£1.82	4.16 €	£5.89
2	Heras	Entrambasaguas	6.7	4.2	5.59 €	£7.92	7.73 €	£10.95
3	Entrambasaguas	Riaño	8.4	5.2	10.21 €	£14.45	17.72 €	£25.10
4	Riaño	San Miguel de Aras	7.8	4.8	13.73 €	£19.45	23.53 €	£33.33
5	San Miguel de Aras	Gibaja	6.6	4.1	16.54 €	£23.42	31.33 €	£44.37
6	Gibaja	La Cadena	6.7	4.2	17.68 €	£25.04	31.52 €	£44.64
7	La Cadena	Mollinedo	9.4	5.8	12.59 €	£17.83	21.87 €	£30.98
8	Mollinedo	Mimetiz	10.8	6.7	13.73 €	£19.44	22.98 €	£32.54
9	Mimetiz	Sodupe	8.9	5.5	13.24 €	£18.75	23.82 €	£33.73
10	Sodupe	Bilbao	11.5	7.1	3.67 €	£5.20	9.77 €	£13.83

The number of services estimated for the new line has been established based on the analysis of the demand, summarized in Table 2. It has been considered a passenger train with a capacity of 238 seats and an average occupancy of 80% (varying from 58% to 95% in 40 years of useful life). This results in 44 daily services, 32 for passengers and 12 for freight.

Table 2 Estimated demand and travel time of the different means of transport between Santander and Bilbao before and after the construction of the HSR line.

Means of transport		Before HSR line		After HSR line	
		Demand	**Travel time**	**Demand**	**Travel time**
Automobile		18,302 p/d/d	1 h 20 min	15,556 p/d/d	1 h 20 min
Bus		1,063 p/d/d	1 h 30 min	956 p/d/d	1 h 30 min
Train	Passengers	0	3 hours	**3,137 p/d/d**	40 min
	Fright	3 trains	-	**6 trains**	-

The segment cost and the estimated services, along with the maximum segment speeds and a proposed schedule, are the inputs to the program that allows us to obtain the optimal combination of single- and double-tracked segments together with the exploitation graphs.

There have been seven alternatives considered. Alternatives 1-5 differ from each other in the maximum relative travel time, varying from 1.20 to 1.00 in a 0.5 interval. In addition, each of these alternatives contemplates two solutions: (1) terminate all the existing train services of FEVE (2) terminate all services except the commuter line between Santander and Liérganes, which will run parallel to the new line in Segment 1. The first solution confers an economic advantage, while the second one a social advantage. Finally, Alternative 6 refers to no action needed and Alternative 7 represents double-track solution.

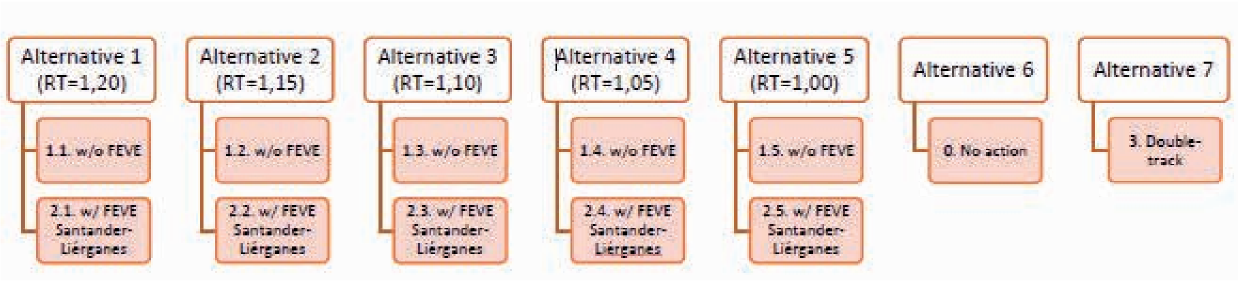


Figure 8. Illustration of the seven alternatives considered.

The program not only minimizes costs and travel times, but also defines the optimal sequence of single- and double-track segments for each alternative. This makes it possible to estimate the construction costs of the infrastructure, as well as the savings compared to the double-track solution.

Table 3 Cost comparison of the 7 alternatives considered. Santander-Bilbao Case

Alternative	**Case**	**Segment**										**Track Typology**		**Budget (Mill)**		**Construction Saving**
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**HS Double**	**HS Simple**	**M€**	**M£**	
1	1.1	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	0%	100%	904.20	795.69	45%
	2.1	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	0%	100%	944.39	831.07	42%
2	1.2	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	0%	100%	904.20	795.69	45%
	2.2	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	0%	100%	944.39	831.07	42%
3	1.3	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	20%	80%	958.80	843.75	42%
	2.3	🟩	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟥	20%	80%	1,028.79	905.33	37%
4	1.4	🟥	🟩	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	20%	80%	1,007.77	886.83	39%
	2.4	🟩	🟩	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟥	20%	80%	1,077.75	948.42	34%
5	1.5	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟥	20%	80%	1,014.38	892.65	38%
	2.5	🟩	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟥	30%	70%	1,092.16	961.10	33%
6	0											-	-	0.00	0.00	100%
7	3	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	100%	0%	1,641.18	1,444.24	0%

Single HS Track	🟩	Double HS Track	🟥

2.1.4 Rational investment analysis

Once all possibilities have been determined, a multi-criteria analysis will decide which one is the best alternative, not only from the economic perspective (50% weight) but also in regard to the quality from a technical standpoint (50% weigh).

To evaluate the economic perspective, a cost-benefit analysis (CBA) is performed on a long-term scale, with a greater focus on their social benefits. The indicators used for this aim are the following:

(a) Benefit-cost ratio (BCR):

$$BCR = \frac{Discounted\ value\ of\ incremental\ benefits}{Discounted\ value\ of\ incremental\ costs} \quad (2)$$

(b) Payback period considering a useful life of 40 years (PB).

$$QM = \frac{\sum Services \cdot Weighing}{\sum Services} \quad (3)$$

Regarding to the technical indicators, the following quality measures have been used:

(a) QM index:

Table 4. Weighing used for the definition of the QM index.

RT	1 <RT<1,05	1,05<RT<1,1	1,1<RT<1,15	1,15<RT<1,2	RT>1,2
Weighing	1	0.8	0.6	0.4	0.2

a) Maximum relative travel time, RTmax
b) Mean relative travel time, RTmean
c) Continuation of services: score 1 if the Santander-Liérganes line is continued and 0 if not.

Table 5. Multi-criteria analysis results. Santander-Bilbao Case

		CBA		Quality indicators				Multi-criteria Analysis
Alternative	Case	BCR	PB	QM	RTmax	RTmean	Cont. Serv.	
1	1.1	1.435	21	0.57	1.52	1.15	0	7.291
	2.1	1.401	22	0.59	1.51	1.15	1	8.197
2	1.2	1.435	21	0.60	1.53	1.17	0	7.418
	2.2	1.327	24	0.57	1.50	1.18	1	7.813
3	1.3	1.381	23	0.83	1.45	1.07	0	8.053
	2.3	1.262	26	0.79	1.41	1.08	1	8.399
4	1.4	1.240	27	0.85	1.39	1.07	0	7.521
	2.4	1.233	28	0.84	1.39	1.07	1	8.432
5	1.5	1.342	24	0.92	1.33	1.03	0	8.256
	2.5	1.220	28	0.93	1.22	1.02	1	8.740
6	0	0.000	0	0.20	5.00	5.00	1	3.905
7	3	0.990	42	1.00	1.00	1.00	0	6.653

2.1.5 Description of the adopted solution

The alternative with highest score is Case 2.5, which is the adopted solution. The solution has the following characteristics:

- The configuration results in 30% of double track against 70% of single track.
- The maximum relative travel time is 1.00. This involves a total duration of the journey of 40 minutes between the city centers of Santander and Bilbao.
- It includes 44 daily services, 32 for passengers and 12 for freight. The resulting timetable is shown in Figure 109.
- This solution also chooses the continuation of the commuter line between Santander and Liérganes.
- The construction cost of the adopted solution amounts to 1,092.16 million euros, saving 549.02 million euros with respect to a double-track solution (33%), which would cost 1,641.18 million euros.
- Considering a useful life of 40 years and meeting the benefits regarding the social welfare, the investment becomes profitable starting from the 28th year in operation.

Figure 9. Final layout of the proposed line showing the single- and double-track segments. Santander-Bilbao Case

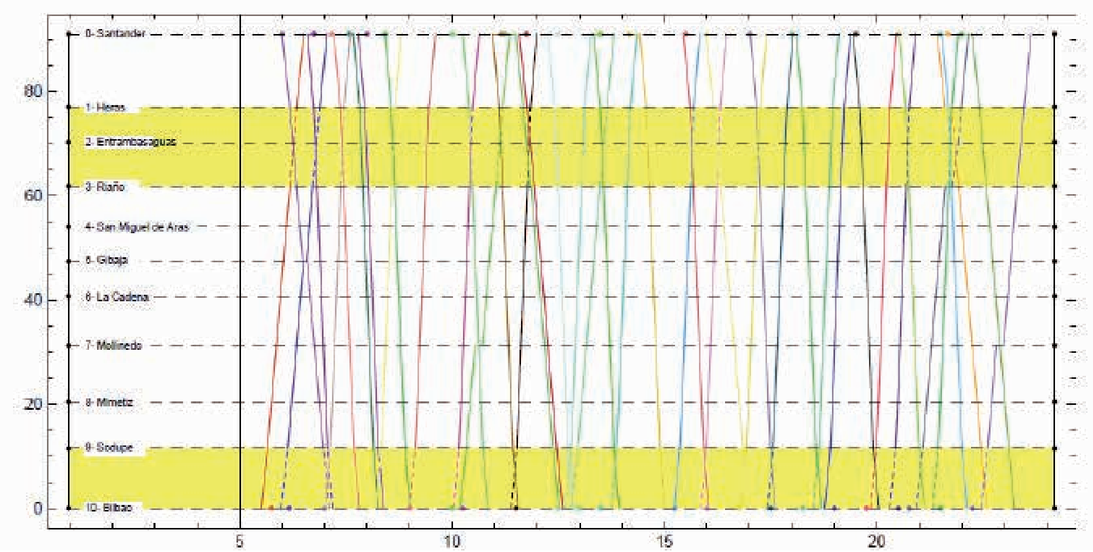


Figure 10. Optimized timetable and selected double- (yellow shade) and single-track segments. Santander-Bilbao Case

2.2 Vitoria-Zaragoza

2.2.1 Current Line

This case describes an ADST proposal between Vitoria and Zaragoza. As Figure 11 shows, in this area there is a railway infrastructure that connects four important cities: Vitoria, Zaragoza, Pamplona and Logroño, which could be considered as four different lines, but for convenience they are assumed to be a unique infrastructure in order to define a global railway improvement for the whole region.

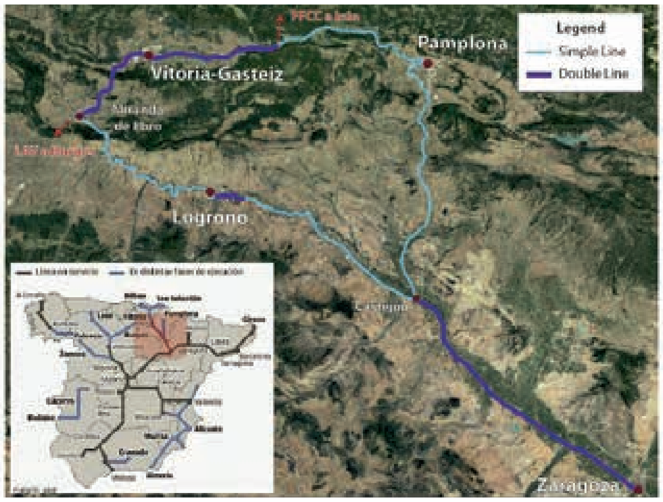


Figure 11. Current line Vitoria-Zaragoza

2.2.2 High Speed line proposal

In this case, similarly to the previous one, a high-speed railway line is projected with:

- Mixed traffic, because this region currently bear an important amount of freight trains and Zaragoza aims to position itself as a major logistic centre.
- Iberian-gauge track (1,668 mm) and a design speed of 250 km/h, in order to reduce the environment impact over this high-value region (La Rioja), to improve the average speed of the current line and to permit the usage of the new line by regional and freight services.

Moreover, due to the high amount of regional services and low-speed freight trains that are currently circulating, it is planned to maintain the conventional line service.

Hence, the proposed line, depicted in Figure 12, mainly consists of a new brand high-speed line together with the rehabilitation of the line through the Logroño metropolitan area and from Castejón (Navarra County) to Zaragoza, which is a segment with a high quality alignment.

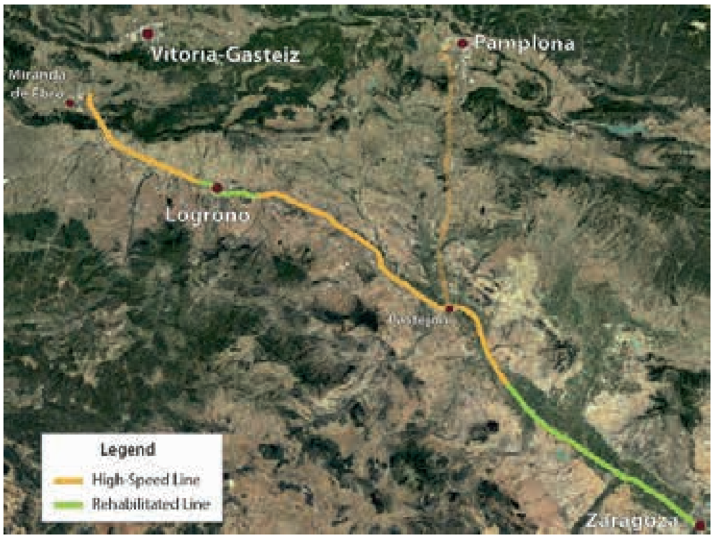


Figure 12. Initial proposal of ADST (Vitoria-Zaragoza Case)

2.2.3 Railway line optimization

In order to define an ADST solution, it is necessary to divide the line in different segments with similar characteristics, see Figure 13. Thus, the 213-km line is divided into fourteen segments with an average cost of 4.6 M€/km and 7.4 M€/Km in single- and double-track respectively, as shown in Figure 13 and Table 6.

Figure 13. Segments proposal of case Vitoria-Zaragoza

Table 6 Data used for the Vitoria-Zaragoza HSR line including the segment definition, their lengths and costs in single and double track.

id	Begin	End	Length (Km)	Cost segment (€)		Cost /km (€/km)		SC/DC (%)
				Single Track	Double Track	Single Track	Double Track	
1	Zaragoza	Alagón	23.0	14,128,988	25,440,536	614,304	1,106,110	55.54
2	Alagón	Gallur	22.0	28,993,623	45,950,485	1,317,892	2,088,658	63.10
3	Gallur	Ribaforada	18.9	11,617,791	20,920,381	614,698	1,106,898	55.53
4	Ribaforada	Murillo	15.1	73,974,914	116,275,615	4,899,001	7,700,372	63.62
5	Murillo	Alfaro	16.0	48,677,810	76,162,378	3,042,363	4,760,149	63.91
6	Alfaro	Rincón de Soto	15.0	34,878,528	56,806,100	2,325,235	3,787,073	61.40
7	Rincón de Soto	Sartaguda	15.0	39,008,833	61,823,412	2,600,589	4,121,561	63.10
8	Sartaguda	Alcanadre	15.0	89,308,075	128,688,296	5,953,872	8,579,220	69.40
9	Alcanadre	Agoncillo	15.0	52,798,000	84,884,756	3,519,867	5,658,984	62.20
10	Agoncillo	Logroño	15.0	9,229,092	79,935,395	615,273	5,329,026	11.55
11	Logroño	Elciego	10.0	135,879,542	211,839,647	13,587,954	21,183,965	64.14
12	Elciego	Bastida	15.0	70,863,223	104,060,948	4,724,215	6,937,397	68.10
13	Bastida	Salinillas	9.0	94,929,137	147,335,037	10,547,682	16,370,560	64.43
14	Salinillas	Armiñón	9.0	93,206,128	144,687,746	9,811,171	15,230,289	64.42
Total			213.5	797,493,684	1,304,810,732	-	-	-
Average values			15.25	56,963,385	93,200,767	**4,583,865**	**7,425,733**	**61.12**

Similarly to the previous case, the costs per km in single- and double-track and mean operating speed of each segment must be used as data for the ADST definition. For this case, 5 main alternatives are defined, which differ from each other in the maximum relative travel time of high speed services, varying from 1.20 to 1.00 in 0.5 intervals.

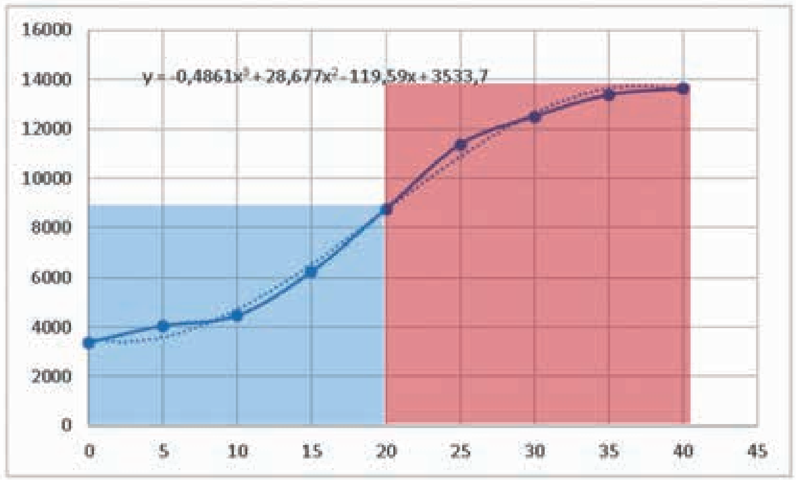


Figure 14. Demand estimate for Vitoria-Zaragoza Case.

Moreover, to estimate the demand of new services for the exploitation period, which is considered of 40 years, a demand analysis is proposed. Thus, as in the previous case, we have considered the current demand, taking into account the travel time and price of each conveyance service. In addition, the passenger redistribution has been estimated with consideration of the existing and the new line, i.e. taking into account not only how many users would maintain their conveyance and how many would change to high-speed services, but also the new generated users.

To calculate the user demand (ordinate axis of Figure 14), two different periods of 20 years have been considered (abscissa axis of Figure 14) in order not to saturate the line unnecessarily and to adapt the rolling stock usage to passenger demand. Consequently, for the initial period (first 20 years), 30 services have been estimated and 60 for the rest of the exploitation period, as

Table 7 shows. Moreover, the current services are added in the first part of the line (ZaragozaCastejón), so that segment will support 116 services (60 high-speed services plus 56 services of conventional lines).

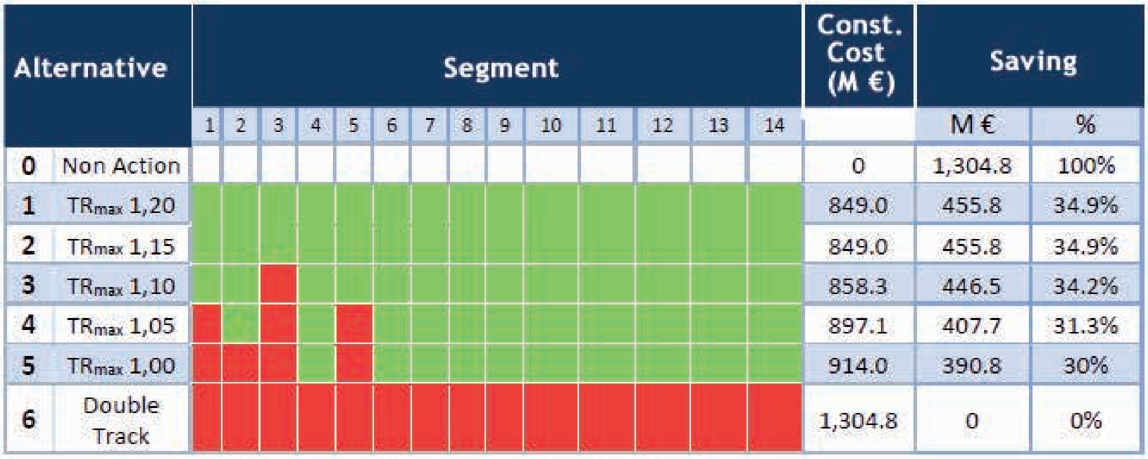
Table 7. High-Speed services estimation

Route	High-Speed Services	
	Year 1-20	Year 21-40
Zaragoza - Logroño - Vitoria	14	28
Logroño - Vitoria	2	4
Pamplona - Zaragoza	12	24
Pamplona - Logroño	2	4
TOTAL	**30**	**60**

Therefore, with all this information, i.e. construction cost, services, speed and proposed schedule, the optimization program calculates which segment should be designed in single- or double- track, in order to define the most cost-effective solution.

The results of the analysis are shown in Table 8 where, in addition to the previously mentioned 5 alternatives, there are two more proposals: Alternative 0, that is, the "No action alternative" that plans to maintain the current line, and Alternative 6, that is, the "Double track alternative", i.e all line in double track. It should be noted that the optimization program has not considered necessary this alternative despite of the high number of 116 circulating services.

Table 8. ADST alternatives for Vitoria-Zaragoza case



Alternative		Segment														Const. Cost (M €)	Saving	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14		M €	%
0	Non Action															0	1,304.8	100%
1	TR_{max} 1,20	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	849.0	455.8	34.9%
2	TR_{max} 1,15	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	849.0	455.8	34.9%
3	TR_{max} 1,10	🟩	🟩	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	858.3	446.5	34.2%
4	TR_{max} 1,05	🟥	🟩	🟥	🟩	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	897.1	407.7	31.3%
5	TR_{max} 1,00	🟥	🟥	🟥	🟩	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	914.0	390.8	30%
6	Double Track	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	1,304.8	0	0%

It is convenient to highlight that, with a relative time of 1.00, which means optimal relative time for the high-speed services, the construction cost of the infrastructure is still a 30% (390,8 M€) cheaper than the double track proposal.

2.2.4 Rational investment analysis

Subsequently, with the aim of defining the optimal multi-criteria solution, a rational investment assessment is developed, in which all the alternatives are considered under an economic, social and quality points of view. The criteria to define the multi-criteria evaluation is defined in subsection 2.1.4, where 50% is the economical approach and the other 50% considers the social benefits. This kind of analysis is essential because these infrastructure constructions have two different perspectives, thus the cost-benefit analysis is fundamental in the short-term economic development, while technical indexes reflect the social benefits which are achieved in a medium and long term.

Therefore, in this case, the same indexes explained in subsection 2.1.4 have been used, except for parameter (d) of the Technical indicator, because in this case it is the Environmental Index, IE.

IE is a quality index whose objective is to benefit in the weighting the alternatives that optimize the route scheme against possible future actions to satisfy higher demands. To achieve this, which fraction of common line between the routes Zaragoza - Vitoria and Zaragoza – Pamplona is measured, that is to say, those with noticeably greater demand, is projected in double-track. In this way, it is possible to satisfy a series of premises:

- Minimal environmental impact, since most of the line would be carried out in an unproductive area because there was previously a railway line and, therefore, it consists of a duplication instead of a line construction.
- Performance optimization, since this segment affects both routes at the same time.
- Reduction of construction costs because most of them are track duplication and not new construction.
- Reduction of maintenance costs, because these segments beards both high speed and conventional services.

Finally, the multi-criteria evaluation, shown in Table 10, indicates that the best economic criteria are obtained by those alternatives that imply cheaper construction costs, while the best quality indices come from those alternatives with a higher percentage of double track. However, the optimal multi-criteria final value, which computes the whole indicator bound to equation (3) and its own weights (see Table 9), corresponds to alternative 5, as shown in Table 10.

Table 9 Indicators weights of Case Vitoria-Zaragoza

Indicator	Weighted
BCR	40%
PB	10%
QM	40%
I_E	10%

Table 10 Multicriteria Assessment for Vitoria-Zaragoza Case

Case	Economic Criteria		Quality indices				Multicriteria Value
	BCR	PB	QM	TR_{max}	TR_{medio}	I_E	
0	0.000	1	0.25	-	-	-	1.975
1	2.829	9	0.69	1.29	1.09	-	7.534
2	2.825	9	0.68	1.28	1.09	-	7.473
3	2.819	9	0.83	1.22	1.05	0.20	8.299
4	2.753	9	0.96	1.16	1.02	0.61	9.118
5	2.716	10	0.97	1.11	1.01	0.84	9.328
6	2.333	13	1.00	1.00	1.00	1.00	8.974

2.2.5 Description of the adopted solution

Finally, the alternative with highest score is Alternative 5, which is the adopted solution. The solution has the following characteristics:

- The configuration results in 37.4% of double track against 62.6% of single track.

- The maximum relative travel time is 1.00 for the High speed services and 1.11 for the conventional services.
- It includes 116 daily services, 60 for passengers and 56 freight services. The resulting timetable is shown in Figure 15.
- The construction cost of the adopted solution amounts to 914.03 million euros, saving 390.8 million euros (30%) with respect to a double-track solution, which would cost 1,304.8 million euros.
- Considering a service life of 40 years and meeting the benefits regarding the social welfare, the investment becomes profitable from the 10th year in operation.

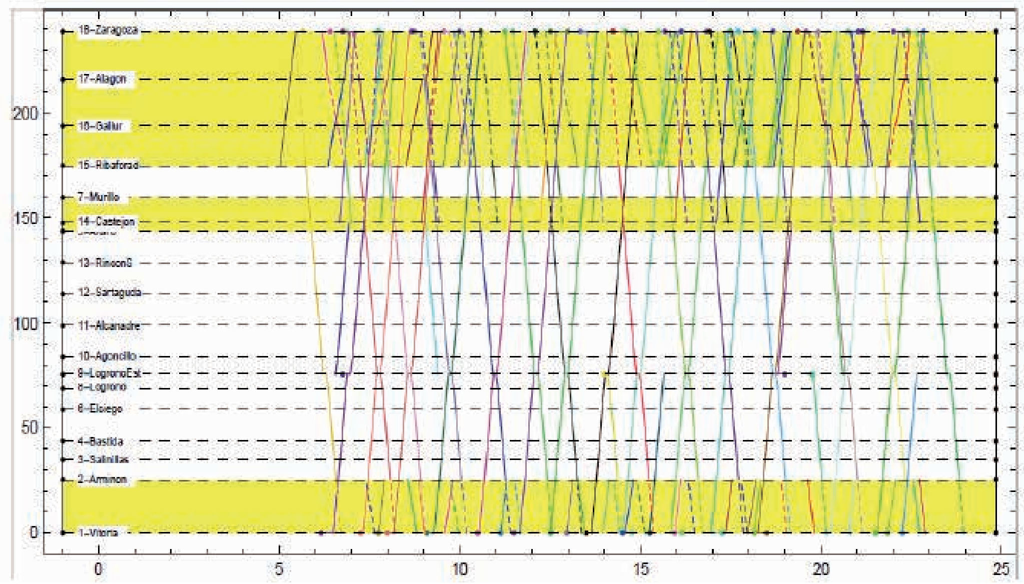


Figure 15. Optimized timetable and selected double- (yellow shade) and single-track segments. Vitoria-Zaragoza Case

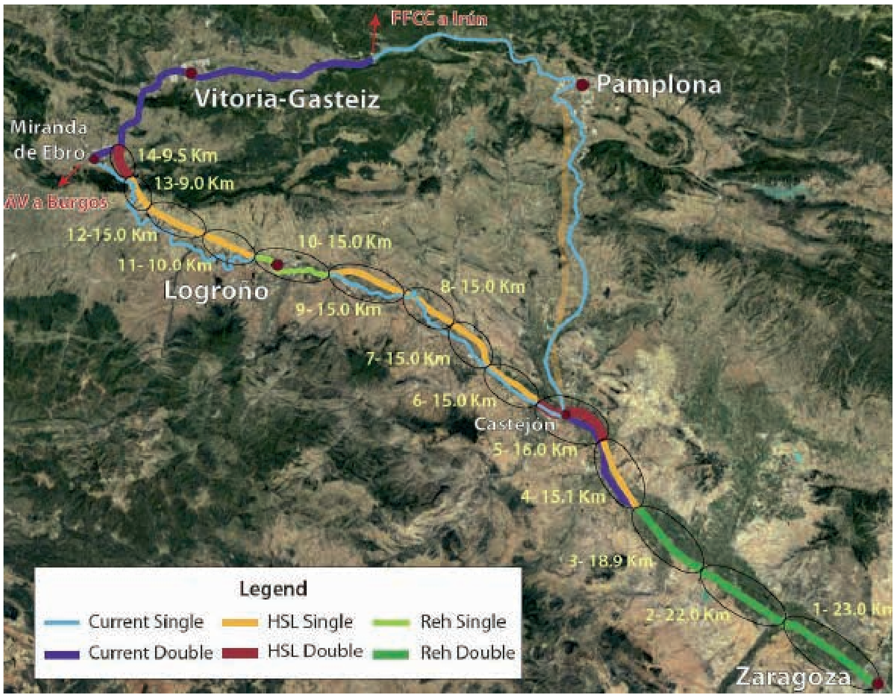


Figure 16. Final layout of the proposed line showing the single- and double-track segments. Vitoria-Zaragoza Case

3. Conclusions

Once the previous examples have been performed, the main conclusions are that the Alternate Double-Single track projects allow to:

- Minimize the construction cost with reduced travel times.
- Design railway lines under current and future demands.
- Define railway alternatives which not impact seriously over the environment.
- Reduce maintenance costs.
- Optimize timetables and improve significantly the current services travel time.
- Model the timetable and the line layout in response to premises of the network.

The case studies have been developed assuming a demand clearly above to de actual one. Despite of that, the ADST alternative provides a solution with far enough rail capacity for all its expected operational life

4. Bibliography

- Amit, I. & Goldfarb, D., 1971. *The timetable problem for railways.* Developments in Operations Research, Volume 2, pp. 379-387.
- Assad, A., 1980. *Models for rail transportation.* Transportation Research Part A, Volume 14, pp. 205-220.
- Burdett, R. L. & Kozan, E., 2010. *A disjunctive graph model and framework for constructing new train schedules.* European Journal of Operational Research, Volume 2010, pp. 85-98.
- Cacchiani, V. & Toth, P., 2012. *Nominal and Robust Train Timetabling Problems.* European Journal of Operational Research, Volume 219, pp. 727-737.
- Caprara, A., Fischetti, M. & Toth, P., 2002. *Modeling and solving the train timetabling problem.* Operations Research, Volume 50, pp. 851-861.
- Carey, M., 1994. *A model and strategy for train pathing with choice of lines, platforms and routes.* Transportation Research Part B, Volume 28, pp. 333-353.
- Carey, M. & Crawford, I., 2007. *Scheduling trains on a network of busy complex stations.* Transportation Research Part B, Volume 41, pp. 159-178.
- Carey, M. & Lockwood, D., 1995. *A model, algorithms and strategy for train pathing.* Journal of the Operational Research Society, Volume 46, pp. 988-1005.
- Castillo, E. et al., 2015. *An alternate double-single track proposal for high-speed peripheral railway lines.* Computer Aided Civil and Infrastructure Engineering, Volume 30, pp. 181-201.
- Castillo, E., Gallego, I., Ureña, J. M. & Coronado, J. M., 2009. *Timetabling optimization of a single railway track line with sensitivity analysis.* TOP, Volume 17, pp. 256-287.
- Castillo, E., Gallego, I., Ureña, J. M. & Coronado, J. M., 2011. *Timetabling optimization of a mixed double- and single-tracked railway network.* Applied Mathematical Modelling, Volume 35, pp. 859-878.
- Castillo, E., Grande, Z., Moraga, P. & Sánchez Vizcano, J., 2016. A *time partitioning*

technique for railway line design and timetable optimization. Computer Aided Civil And Infrastructure Engineering, Volume 31, pp. 599-616.

- Cordeau, J. F., Toth, P. & Vigo, D., 1998. *A survey of optimization models for train routing and scheduling*. Transportation Science, Volume 32, pp. 380-404.
- D'Ariano, A., Pacciarelli, D. & Pranzo, M., 2007. *A branch and bound algorithm for scheduling trains in a railway network.* European Journal of Operational Research, Volume 183, pp. 643-657.
- D'Ariano, A. & Pranzo, M., 2004. *A real time train dispatching system based on blocking time theory.* Delft, DUP Science, pp. 129-152.
- D'Ariano, A., Pranzo, M. & Hansen, I. A., 2007. *Conflict Resolution and Train Speed Coordination for Solving Real-Time Timetable Perturbations.* IEEE Transactions on Intelligent Transportation Systems, Volume 8.
- Haghani, A. E., 1987. *Rail freight transportation: a review of recent optimization models for train routing and empty car distribution.* Journal of Advanced Transportation, Volume 21, pp. 14172.
- Hellström, P., 1998. *Analysis and evaluation of systems and algorithms for computer-aided train dispatching*, Sweden: s.n.
- Higgins, A., Kozan, E. & Ferreira, L., 1996. *Optimal scheduling of trains on a single line track.* Transportation Research Part B, Volume 30, pp. 147-161.
- Jia, L. M. & Zhang, X. D., 1993. *Distributed intelligent railway traffic control based on fuzzy decision making.* Fuzzy Sets and Systems, Volume 62, pp. 255-265.
- Kraay, D. R. & Harker, P. T., 1995. *Real-time scheduling of freight railroads.* Transportation Research Part B, Volume 29, pp. 213-229.
- Lin, D. Y. & Ku, Y. H., 2013. Using Genetic Algorithms to Optimize Stopping Patterns for Passenger Rail Transportation. Computer-Aided Civil and Infrastructure Engineering, pp. n/a--n/a.
- Ouyang, Y. et al., 2009. *Optimal Locations of Railroad Wayside Defect Detection Installations.* Computer-Aided Civil and Infrastructure Engineering, Volume 24, pp. 309-319.
- Pachl, J., 2014. *Railway Timetabling and Operations. Analysis - Modelling - Optimisation - Simulation - Performance Evaluation*. In: Hamburg(): Eurailpress, pp. 23-24.
- Petersen, E. R., Taylor, A. J. & Martland, C. D., 1986. *An introduction to computer aided train dispatching.* Journal of Advanced Transportation, Volume 20, pp. 63-72.
- Sahin, I., 1999. *Railway traffic control and train scheduling based on inter-train conflict management.* Transportation Research Part B, Volume 33, pp. 511-534.
- Yang, Z. & Hayashi, Y., 2002. *GIS-Based Analysis of Railways Origin/Destination Path-Selecting Behavior.* Computer-Aided Civil and Infrastructure Engineering, Volume 17, pp. 221-226.

High Speed Rail Infrastructure

360.revista de alta velocidad
número 5 - junio - 2018. Pág 301 - 311

Importance of vertical rail track stiffness on dynamic overloading: Limitations of the Eisenmann formulation

Balmaseda, Lucía
Gallego, Inmaculada
Sánchez-Cambronero, Santos
Rivas, Ana

University of Castilla-La Mancha[1]

Abstract

Since the development of high-speed rail in Europe in the 1970s, the vertical loads that are transmitted to the rail track by vehicles have been measured. The formulations obtained by Eisenmann and Prud'Homme are the most notable formulations of this period due to their extensive application.

The formula proposed by Eisenmann considered the quality of the rail track and was widely proven for maximum speeds of 200 km/h. With the emergence of high speed trains, Eisenmann proposed a modification to the previously proposed formula to adapt it to the case of high-speed vehicles and lines.

Additionally, the formula of Prud'Homme is important because it introduces new criteria and reveals how the vertical rail track stiffness, the unsprung mass of the vehicle and the quality of the rail track, in addition to the speed of the vehicle, affect the dynamic overloads.

As expressed by this formula, for a given speed and rail track quality, different geotechnical and geometric compositions of infrastructure, which determine the stiffness, cause different dynamic overloads. This fact was not considered in the Eisenmann formulation, exposing its limitations.

The objective of this article is to analyze these limitations; for this analysis, a threedimensional (3D) finite element model of the rail track will be employed to calculate static and dynamic stiffness and obtain the dynamic load values for different types of infrastructure. These results will enable us to analyze the relationship between the strength characteristics of the rail track and the dynamic coefficientC_d, which is understood to be the ratio between the total dynamic loads and the static loads that are transmitted to the rail track.

Keywords: high-speed rail, Eisenmann formula, dynamic coefficient, overload, Prud´Homme formula.

---

[1] *Balmaseda, Lucía. University of Castilla-La Mancha, Spain. Email: Lucia.balmaseda@uclm.es*
Gallego, Inmaculada. University of Castilla-La Mancha, Spain. Email: Inmaculada.Gallego@uclm.es
Sánchez-Cambronero, Santos. University of Castilla-La Mancha, Spain. Email: Santos.Sanchez@uclm.es
Rivas, Ana. University of Castilla-La Mancha, Spain. Email: Ana.Rivas@uclm.es

1. Background

The calculation of the vertical forces on rail tracks is traditionally determined from a static analysis. However, experimental measurements confirmed that the loads that are transmitted to the rail track increased with speed. This finding prompted research in the railway field that proposed the use of a dynamic amplification coefficient C_d.

The expressions of Whinkler and Pihera in 1915, Driessen in 1936 and Schramm in 1955 are the most notable expressions among the first empirical expressions that were proposed to quantify the magnitude of this coefficient.

During an extensive test campaign by Deutsche Bahn (DB) in the 1960s in Germany, the test results indicated that the dispersion of the dynamic loads increased with speed compared with their average. The magnitude of these dispersions was directly related to the quality of the rail track and the vehicle. These results caused Birmann to propose the following formula to the Committee D-71 of the ORE in 1966:

$$C_d = 1 + 0.04 * \left(\frac{v}{100}\right)^3 + a * b * \left(0.1 + 0.01 * \left(\frac{v}{100}\right)^3\right) \qquad (1)$$

In 1969, Eisenmann proved that these dispersions follow a normal distribution, as shown in figure 1.

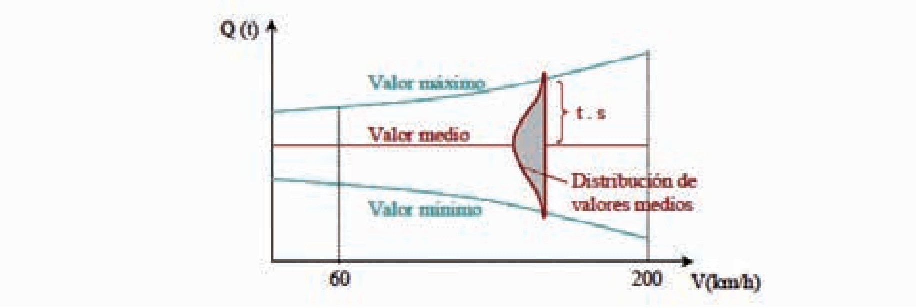


Figure 1: Dynamic oscillation of the load per wheel. Source: Teixeira, 2003

Based on the experimental results, Eisenmann proposed an empirical formula to determine C_d:

$$C_d = 1 + t * s * \varphi \qquad (2)$$

where

t is the factor of probabilistic certainty such that

$t = 1$ 68.3% of the values,

$t = 2$ 95.5% of the values,

$t = 3$ 99.7% of the values;

s is a factor that is dependent on the condition of the rail track,

$s = 0.1$ very good condition,

$s = 0.2$ good condition,

$s = 0.3$ poor condition,

φ is a factor that is dependent on the running speed,

$$\varphi = 1 \qquad \text{for } V \leq 60 \text{ km/h}$$
$$\varphi = 1 + \frac{V-60}{140} \quad \text{for } V > 60 \text{ km/h} \qquad (3)$$

Eisenmann's formula considered the quality of the rail track and the confidence interval. This formula has been widely proven for maximum speeds of 200 km/h.

Since the introduction of the high-speed rail in Europe in the 1980s, rail vehicles have been designed to ensure that their loads, which are transmitted to the rail track, are substantially less than the loads transmitted by conventional vehicles. The quality of the newer rail tracks is significantly higher than the quality of existing rail tracks. Consequently, Eisenmann's formula from 1969, which was developed using data from conventional vehicles on conventional lines, was not valid for determining the stresses caused by high-speed trains, such as France's highspeed trains (*Train à Grande Vitesse,* TGV). In 1993, Eisenmann proposed a modification of his previous formula to adapt it to the case of high-speed vehicles and lines. The following expression was defined for the parameter φ for speeds from 201 km/h to 300 km/h:

$$\varphi = 1 + \frac{V-60}{380} \qquad (4)$$

In the 1970s, the National Society of French railways (SNCF) developed important theoretical and experimental research to analyze the effect of rail tracks and vehicles on dynamic overloads. This research was performed by Prud'Homme, who used a classic model for modeling a rail track-vehicle system and its behavior and analyzed the excitations produced by the irregularities of a rail track. Prud'Homme applied the theory of random vibrations to develop the well-known formula for calculating the dynamic overloads produced by unsprung masses of a vehicle:

$$\sigma_{(\Delta QNS)=0.45*\frac{V}{100}*b*\sqrt{m_{ns}*K*\gamma(\varepsilon)}} \qquad (5)$$

where

$\sigma_{\Delta QNS}$: standard deviation of the dynamic overloads due to unsprung masses,

V: running speed of the vehicle, km/h,

b: variable related to rail track defects and vehicle defects,

m_{NS}: unsprung mass of the vehicle,

K: vertical rail track stiffness, t/mm,

$\gamma(\varepsilon)$: rail track damping

The significance of Prud'Homme's formula is that it introduces new criteria and reveals how the vertical rail track stiffness, the unsprung mass of the vehicle, and the quality of the rail track, in addition to the speed of the vehicle, affect the dynamic overloads.

As expressed by this formula, for a given speed and rail track quality, different geotechnical and geometric compositions of infrastructures, which determines the value of K, produces different dynamic overloads. This fact was not considered by Eisenmann's formula and exposes its limitations. The objective of this article is to analyze these limitations and the relationship between the strength characteristics of the rail track and the dynamic coefficient C_d.

2. Dynamic overloads calculation

2.1 Description of the numerical model

The calculation of the vertical rail track stiffness K was performed using a 3D finite element numerical model of a section of railway rail track using the software ANSYS. One of our objectives is to detect the value of K for the ballasted rail track. The passage of a rail load has been simulated using one model (refer to figure 2). To study the projected ballasted rail track, we have developed a model based on the method proposed in (Gallego, I., 2009) that was used to propose new design criteria in (Gallego, I., 2011) and (Gallego I., 2012 and 2013)

A perfect elastoplastic law was assumed to simulate the behaviors of all materials, with the exception of the rails, elastic pads, sleepers, and the granular material treated with cement and concrete slabs, which are assumed to be governed by an elastic law.

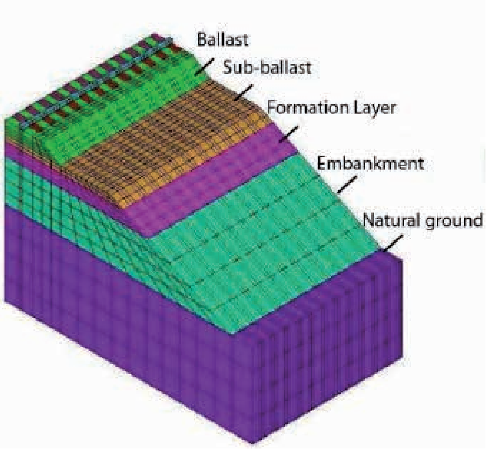


Figure 2. Finite element models for the proposed sections. Source: Gallego et al, 2016.

To obtain the vertical railway rail track stiffness, the following parameters and considerations were used (refer to Gallego, I., (2009) for modeling details):

- The modulus of elasticity EEM of the material that comprises the embankment.
- The height of the embankment hEM.
- Since the natural ground on which the embankment stands along the line consists of rock, with a very high modulus of elasticity, we have assumed that the displacement of this layer is going to be null. Therefore, we can delete it to reduce the computational time to solve the model and achieve similar results.
- The thickness of the ballast and sub-ballast are 35 cm and 30 cm, respectively.
- The well-known rail is UIC 60.
- The elastic pad has k=100 kN/mm.
- The sleeper is the monoblock pre-tensioned AI-99.

To study the influence of these parameters, a series of generic case studies were created with different values (values that are not fixed).

To establish the values that are assigned to the modulus of elasticity of the material that comprises the fill material of the embankment EEM and the other parameters that are needed for the model, the values considered in the soil classification conducted by the UIC are used as a reference. These values are listed in Table 1.

Table 1. Values of the geotechnical parameters considered in the model.

Material	$E(N/m^2)$	ν	$c(N/m^2)$	$\phi(°)$	$\rho(kg/m^3)$
Steel for rails	$2.1\text{x}10^{11}$	0.3	-	-	7500
Base plates	$6.91\text{x}10^{7}$	-	-	-	-
Sleeper E1	$8.01\text{x}10^{10}$	0.25	-	-	2500
Sleeper E2	$5.02\text{x}10^{10}$	0.25	-	-	2500
Sleeper E3	$3.69\text{x}10^{10}$	0.25	-	-	2500
Ballast	$1.3\text{x}10^{8}$	0.2	0	45	1900
Formation layer	$0.8\text{x}10^{8}$	0.3	0	35	2000
Material QS1	$12.5\text{x}10^{6}$	0.3	10000	10	2000
Material QS2	$25\text{x}10^{6}$	0.3	10000	20	2000
Material QS3	$80\text{x}10^{6}$	0.3	0.30	30	2000
TGM	$160\text{x}10^{6}$	0.25	-	-	2300

Source: Gallego, 2016

EEM values of 12.5, 25 and 80 MPa have been employed. The value of 160 MPa, which corresponds to an embankment that consists of cement-treated granular material (TGM), has also been considered, as in the case of an embankment-structure transition. For the height of the embankment hEM, the values 3, 5, 7, 10 and 15 m have been employed. The considered case studies are determined by the possible combinations between the different values of EEM and hEM. In this study, only the results obtained at 7 m are employed since that height is considered to be average in the existing rail tracks.

The load per wheel considered corresponds to the static load of passenger trains on the Madrid-Seville high-speed line (considering all types of trains); it increases to a value of 186.40 kN per axle.

2.2 Dynamic and static vertical rail track stiffness

From the resolution of each case study, the value for the vertical rail track stiffness K was obtained as follows (refer to Esveld (11):

$$K = \frac{Q}{z}, \tag{1}$$

where Q is the vertical static load per wheel and z is the summation of all deflections in the vicinity of the load measured on the head of the rail.

The value of K in the Prud'Homme formula is obtained from the dynamic load. Since this load is unknown, iteration is necessary to obtain a converged value of the stiffness based on a certain static load value.

Different types of infrastructures were selected. For each infrastructure, the static stiffness was calculated by entering the static load in the software and consecutively applying it to four sleepers to simulate the passage of an axle. With the static stiffness value obtained, an initial dynamic overload value was calculated. With the total value of the load, the model was recalculated and an initial dynamic stiffness value was obtained. This process was iterated twice; the dynamic rigidity values obtained are listed in Table 2.

Table 2.

Static and dynamic stiffness values (kN/mm) and their ratios.

TYPE OF TRANSITION	STIFFNESS (KN/mm)	SLEEPER 5	SLEEPER 6	SLEEPER 7	SLEEPER 8
Embankment=QS2 **Natural ground=QS1**	K static	10.800	11.195	10.997	11.263
	K dynamic	12.705	13.023	13.137	13.337
	K dynamic / K static	1.1763	1.1633	1.1976	1.1841
Embankment=QS3 **Natural ground=QS1**	K static	54.186	54.985	55.476	55.808
	K dynamic	55.956	55.671	57.216	57.073
	K dynamic / K static	1.0326	1.0125	1.0313	1.0226
Embankment=QS2 **Natural ground=QS2**	K static	16.380	16.658	16.823	16.915
	K dynamic	18.747	19.106	19.335	19.431
	K dynamic / K static	1.1445	1.1469	1.1493	1.1487
Embankment=QS3 **Natural ground=QS2**	K static	59.936	60.519	60.915	60.519
	K dynamic	61.151	61.735	62.171	62.094
	K dynamic / K static	1.0203	1.0200	1.0206	1.0261

(Source: Gallego, 2012).

From Table 2, we note that

- The dynamic stiffness is always greater than the static stiffness, as expected.
- The difference between the static stiffness and dynamic stiffness decreases as the stiffness value is increased.
- Structures with very elastic infrastructure have 18% greater dynamic stiffness. Rail track structures with a stiffness of approximately 50 and 60 KN/mm have approximately 2% greater dynamic stiffness, which is almost negligible.

Since the minimum appropriate stiffness for high-speed rail infrastructures is 60 KN/mm; the recommended values are 70 and 80 KN/mm. In these cases, the difference between the static stiffness and the dynamic stiffness is very small; use of the simplifying assumption that the static stiffness is equal to the dynamic stiffness seems reasonable.

2.3 Stiffness and dynamic overload results

Due to the nonlinear behavior of the material, the loads need to be applied in several stages. In the first stage, only the weight of the materials is considered until the equilibrium of the stresses is achieved, whereas the load caused by the train is considered in subsequent stages. Because the displacements of interest correspond to the load from the train, they will be obtained by the difference between the totals after applying the train load and the values that correspond to the first stage. According to Committee D-71 of the International Union of Railways Office for Research and Experiments (Office de Recherches et d'Essais de l'Union Internationale des Chemins de fer, ORE, Report No. 28, 1983), the load is distributed on the four sleepers adjacent to the sleeper that is loaded, two on each side. This distribution implies that the real value of the settlement of the rail head caused at a certain point by the load of the wheel that acts on it can only be determined by considering the previous loads that affect this point. The load of at least three consecutive sleepers—the first two sleepers (T5 and T6) previous to the sleeper that is being analyzed (T7), and the latter—must be considered.

Table 3.

Dynamic coefficient values C_d for different types of structure strengths of the rail track.

TYPE OF INFRASTRUCTURE	DYNAMIC STIFFNESS (tn/mm)				DYNAMIC COEFFICIENT			
	K (T5)	K (T6)	K (T7)	K (T8)	SLEEPER 5	SLEEPER 6	SLEEPER 7	SLEEPER 8
Embankment=QS2 Natural Ground=QS1	1.019	0.770	0.674	0.622	1.36	1.33	1.32	1.31
Embankment=QS2 Natural Ground=QS2	1.572	1.371	1.225	1.161	1.41	1.39	1.38	1.37
Embankment=QS2 Natural Ground=QS3	3.323	3.181	3.128	3.061	1.56	1.55	1.54	1.54
Embankment=QS2 Natural Ground=ROCA	3.843	3.781	3.758	3.677	1.59	1.59	1.59	1.58
Embankment=QS3 Natural Ground=QS1	2.890	2.596	2.469	2.393	1.52	1.50	1.49	1.49
Embankment=QS3 Natural Ground=QS2	5.093	4.931	4.905	4.867	1.67	1.66	1.66	1.65
Embankment=QS3 Natural Ground=QS3	7.061	7.061	7.115	7.087	1.77	1.77	1.78	1.78
Embankment=QS3 Natural Ground=ROCA	7.403	7.432	7.471	7.468	1.79	1.79	1.80	1.80
Embankment=MGT Natural Ground=QS1	6.778	6.778	6.853	6.828	1.76	1.76	1.76	1.76
Embankment=MGT Natural Ground=QS2	7.368	7.339	7.397	7.368	1.79	1.79	1.79	1.79
Embankment=MGT Natural Ground=QS3	8.321	8.359	8.434	8.434	1.84	1.84	1.84	1.84
Embankment=MGT Natural Ground=ROCA	8.780	8.776	8.817	8.830	1.86	1.86	1.86	1.86

Source: Prepared by the authors, 2017

From Table 3, we note that the values obtained in sleeper T7 are always similar to the values obtained for the two adjacent sleepers. Three loading conditions did not have to be considered; two conditions would have been sufficient. This finding is explained by the fact that the largest plastic deformation is caused by the loads from the weight of the wheel; it is not caused by the loads from the passage of an axle.

3. Comparison of the numerical model results with the results obtained with the Eisenmann formula.

Figure 3 shows the value of the dynamic coefficient for the two Eisenmann formulations: Values 0.1 and 0.2 were considered for s the infrastructure is in very good condition or good condition. The formulation that is applicable for speeds between 200 and 300 km/h. With regard to the factor of probabilistic certainty t, a value of 3 was employed, which corresponds to the highest statistical reliability.

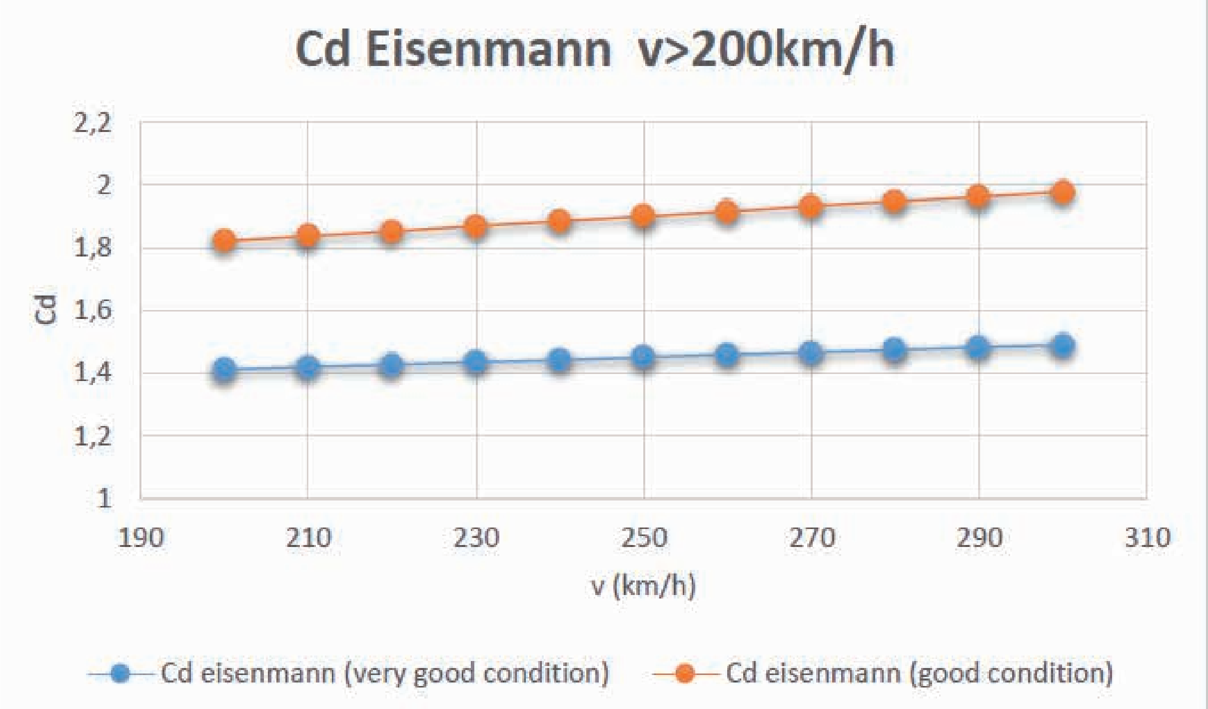


Figure 30. Dynamic coefficient values C_d obtained with the Eisenmann formulations for different conditions.
Source: Prepared by the authors, 2017

The limitations of the Eisenmann formula, which does not consider the structure strength of the rail track, are evident when the results obtained by calculating K and using Prud'Homme's formula for the dynamic coefficient are compared with the results obtained by Eisenmann.

For high-speed rail, the recommended minimum stiffness value is 60 KN/mm, and values between 70 and 80 KN/mm are preferred. These values correspond to the last six structures in Table 2.

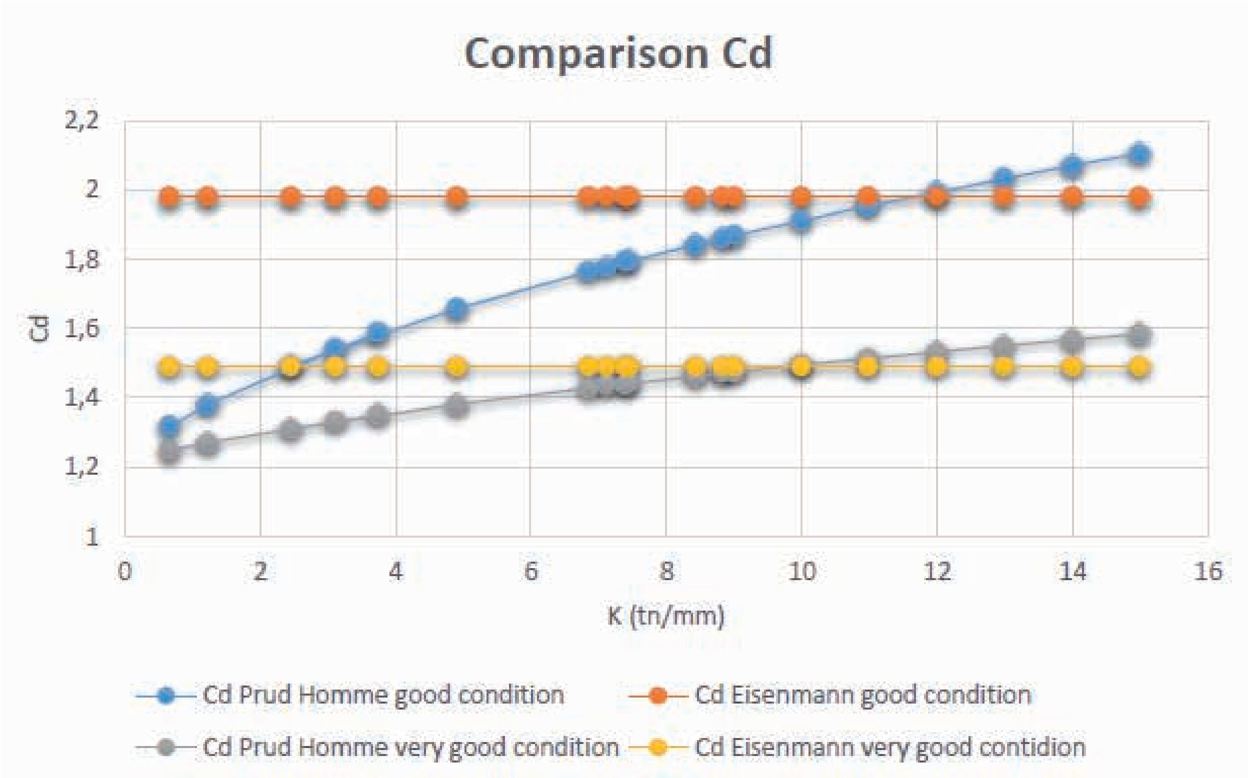


Figure 4. Dynamic coefficient values C_d obtained with the Eisenmann and Prud'Homme formula
Source: Prepared by the authors, 2017

As observed in figure 4 the difference between the coefficient given by Prud'Homme and Eisenmann is smaller as the quality of the high speed track increases.

The difference smaller as it increases the stiffness of the infrastructure, until they converge around the 100 and 120kn/mm where they separate.

Eisenmann offers a superior value to the one of Prud'Homme. In the most flexible regions whereas Prud'Homme occupies the part superior in the most rigid zone.

To determine the proportions on the basis of the criterion of Prud'Homme in the surroundings from the 60 to the 90 kn/mm will help us to rationalize the infrastructure; whereas from the 100 kn/mm when using Eisenmann we would not remain the side of the security.

The values in which Eisenmann approach would correspond to the obtained ones in this paper for the most rigid structures.

However, this infrastructure type is not the one that predominates in the high speed lines, it corresponds to compositions in the zones of transition; reason why to use it is to determine the proportions of the embankment of all the plan would be erroneous.

4. Conclusions

The calculations performed on the 3D numeric model of the rail track proved the following conclusions:

- The dynamic stiffness is always greater than the static stiffness and the difference between them decreases as the static stiffness increases.
- For structures with very elastic infrastructures, the dynamic stiffness is greater than the dynamic stiffness by 18%. However, for rail track structures with rigidities of approximately 50 to 70 KN/mm, the dynamic stiffness is greater by approximately 2%, which is an almost negligible value.
- The stiffness of high-speed infrastructures needs to be greater than 60 KN/mm, and values between 70 and 80 KN/mm are recommended. In these cases, the difference between static stiffness and dynamic stiffness is very small. Therefore, use of the simplifying assumption that the static stiffness is equal to the dynamic stiffness seems reasonable.
- The dynamic amplification coefficient is the result of relating this dynamic loading to the static value of the load. The comparison of the analysis of the coefficient values obtained using Prud'Homme with the analysis obtained using the Eisenmann formula reveals the limitations of the latter formula.
- Is a good approach to use Eisenmann for stiffness superiors to 80 KN/mm, but for more flexible compositions we would be overdesign the infrastructure.
- In addition, the formulas converge in zones of transition where appear the MGT, reason why doesn’t seem suitable to use it to determine the embankment on the rest of the high speed line.

5. References

- Comité D-71, (ORE), (Office de Recherches et d`Essais de l`Union Internationale des Chemins de Fer), (1970), Sollicitation de la voie du ballast et de la plateforme sous l`action des charges roulantes, RP1.
- Comité D-117, (ORE), (Office de Recherches et d`Essais de l`Union Internationale des Chemins de Fer), (1983), Adaptation optimale de la voie classique au trafic de l’avenir, Rapport nº 28 - Abaques de comportement du système voie/fondation-.
- Gallego, I. y López, A. ,2009. Numerical simulation of embankment–structure transition design., Proceedings of the Institution of Mechanical Engineers, PartF: Journal of Rail and Raspid Transit, Vol. 223, pp. 331-343.
- Gallego, I, et al., 2011. Vertical track stiffness as a new parameter involved in designing high speed railway, Journal of Transportation Engineering (ASCE), Vol.137(12), pp. 971-979
- Gallego, I. Heterogeneidad resistente de las vías de Alta Velocidad: Transición terraplénestructura. Editorial Academica Española, 2012

- Gallego, I, et al., 2012. Design of embankment structure transitions for railway infrastructure. Proceedings of the Institution of Civil Engineers Transport, Vol. 165(1), pp. 27-37

- Gallego, I, et al., 2013. Recommendations for numerical rail substructure modeling considering nonlinear elastic behavior, Journal of Transportation Engineering (ASCE), Vol.139(8), pp. 848858

- Gallego, I, et al., 2016. A mixed slab-ballasted track as a means to improve the behaviour of railway infraestructure, Proceedings of the Institution of Mechanical Engineers, PartF: Journal of Rail and Raspid Transit, Vol 230(7), pp. 1659-1672.

- López Pita, A., Infraestructuras Ferroviarias, Ediciones Universidad Politécnica de Cataluña, 2006

- Teixeira, P. F. Contribución a la reducción de los costes de mantenimiento de vías de alta velocidad mediante la optimización de su rigidez vertical. s.l. : Doctoral Thesis, Polytechnic University of Cataluña., 2003.

High Speed Rail Infrastructure

360.revista de alta velocidad
número 5 - junio - 2018. Pág 313 - 321

Development of an alternative maintenance technique for railway ballasted tracks

Sol-Sánchez, Miguel

Moreno-Navarro, Fernando

Rubio-Gámez, Mª Carmen

University of Granada, Spain.[1]

Abstract

Currently, ballast tamping is the most common maintenance intervention in ballasted tracks in order to restore original geometry while recovering damping capacity. However, it is wellknown the limited effectiveness of this technique due to the degradation of ballast particles during tamping, and to the memory phenomenon of ballast layer which quickly restores its pre-maintenance geometry. Then, under the need for developing more durable and effective maintenance techniques, this paper proposes an alternative process based on combining the benefits of using Under-Sleeper Pads and the process known as Stoneblowing, by replacing part of the small stones applied during this last technique with rubber particles from wasted tires that act as flexible aggregates under the sleeper. For this purpose, this paper presents an initial laboratory study focused on analysing the viability and effectiveness of this alternative technique while defining the influence of the quantity of rubber as main designing parameter. Results are based on full-scale testing box tests, and they show that this process allows for reducing short and long-term ballast settlement while reducing ballast degradation and the stress transmitted to sub-layers, in comparison to conventional stoneblowing and to the use of under-sleeper pads. Also, it is possible to optimize track deflection and damping capacity by using different quantities of rubber.

Keywords: Stoneblowing, Under-sleeper pads, rubber particles, full-scale testing box, laboratory study.

---

[1] *Sol-Sánchez, Miguel. University of Granada. Email: msol@ugr.es (corresponding author)*
Moreno-Navarro, Fernando. University of Granada. Email: fmoreno@ugr.es
Rubio-Gámez, Mª Carmen. University of Granada. Email: mcrubio@ugr.es

1. Introduction

Railway over ballasted tracks plays an essential role in most transportation networks, since it presents a range of benefits in comparison with other alternative systems such as concrete slabs. However, changes in the geometric parameters, associated with differential settlement along ballasted tracks, is one of the major problems in this type of infrastructures for railway transportation. Passing trains generates cyclic movements in ballast particles that cause the recompaction of the granular layer and its accumulative vertical deformation, and therefore, to maintain track geometry within adequate safety and ride quality levels, permanent way maintenance for re-levelling and realignment is necessary. Based on this, automatic tamper is the machine most used from the 1960s, which consists of lifting and laterally squeezing the ballast to fill the void space generated beneath the sleeper (Selig and Waters, 1994; Indraratna et al., 2011). However, this maintenance operation is accompanied by some detrimental effects such as loosening of compacted ballast layer, particles deterioration and reduced track stability (Indraratna et al., 2011).

In light of these problems, alternative solutions are being developed for the improvement of the track quality and the effectiveness of the maintenance tasks. In particular, a process known as stoneblowing is worthy of note since it allows for the reduction of ballast breakage during maintenance whilst simultaneously maintaining the vertical strength of the layer (Hellawell, 1997; Claisse et al., 2003). In addition, using this technique avoids rapid ballast recompaction after maintenance. This process, developed in the U.K., consists of adding small stones (around 14-20 mm) to the existing gap between the sleeper and the ballast surface once the former has been raised to the desired level. However, there are some concerns associated with stoneblowing, such as the stiffening of the granular layer and its retarded capacity to damp loads. Another alternative solution to reduce ballast maintenance, applied primarily in European transition sections (where differential settlements are more probable), consists of using elastic Under-Sleeper pads (USPs) in order to reduce the stress on ballast and the settlement of the granular layer, among other benefits (Plica, 2007; Dahlberg, 2010). Nonetheless, its widespread application is limited due to its high costs (which can reach up to 40% of the final price of each sleeper) and to the need for bonding these elements to the bottom of sleepers.

With a view to considering both alternative measures for reducing ballasted track maintenance, this paper proposes a new method that consists of replacing part of the small stones used in the process of stoneblowing with rubber particles obtained from waste tires, acting as flexible aggregates under the sleeper with capacity to damp loads. Thus, this proposed solution, refereed as “Stone-Rubber Blowing”, could replace the incorporation of USPs and would avoid the increase of costs associated with the use of these elements (since the rubber particles are applied at the time of the small stones during the stoneblowing process, without the need of fixing the elastic elements to the bottom of the sleeper) while at the same time an abundant waste material (end-of-life tires) is reused to improve the effectiveness of stoneblowing process. In addition, no bonding agents are needed since the rubber particles are used as aggregates, which also avoids increasing maintenance costs in reference to other solutions (Ho et al., 2013; Fontserè et al., 2016).

For this purpose, this paper analyses the effect of replacing natural small stones with different quantities of rubber particles (used as flexible aggregates), comparing the track section behavior after this process (Stone-Rubber Blowing) to that measured after conventional stoneblowing technique as well as in comparison to the application of different USPs (with various stiffness values) used as reference elastic element in combination with stoneblowing.

2. Methodology

2.1 Materials

In order to analyse the effect of using rubber particles as flexible aggregates during stoneblowing process (mixed with part of natural aggregates used in such conventional technique), these elastic components were obtained by grinding up waste tires primarily composed of rubber (Figure 1). The density of the rubber particles was around 1.15 Mg/m3. The gradation of the elastic granules was as follows: 2% of the particles were between 8 and 12 mm; 61% from 12 to 16 mm; 34% from 16 mm to 22.4 mm; and 2% from 22.4 mm to 25 mm. The rationale for selecting these sizes was to avoid percolation and package (Selig and Waters, 1994; Fair, 2003).

For the same reason, the size of the small natural stones (used to fill the gap between the sleeper bottom and ballast surface after settlement during stoneblowing process) was mainly fixed between 14 mm and 20 mm. These stones were obtained from ophitic rocks, presenting a resistance to fragmentation lower than 8% (EN 1097-2), which is appropriate to limit the fouling of ballast associated with natural aggregates degradation. In Figure 1, the stones are orange painted in order to distinguish them from ballast particles.

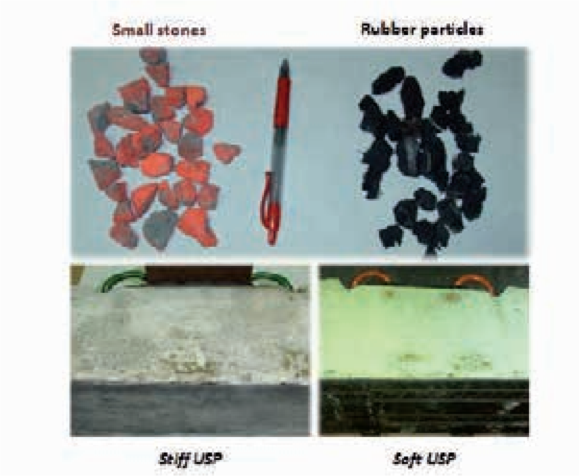


Figure 1. Main materials used for the solutions analysed in this study.

As reference elastic elements to compare the influence of applying rubber particles under sleeper, the USPs used in this study (Figure 1) were manufactured from end-of-life deconstructed tire tread layers (obtained without any need for grinding), which have been shown to be appropriate for application as elastic elements in railway tracks (Sol-Sánchez et al., 2014) These pads were categorized as soft USP (4.5 mm thick) and stiff USP (2.5 mm thick), presenting a static bearing modulus close to 0.20 N/mm3 and 0.40 N/mm3, respectively.

On the other hand, to carry out this study, a conventional track section was reproduced in laboratory by using a testing box that allows for simulating the effect of trains passing, and then, the track settlement due to traffic, which requires maintenance interventions (in this case stoneblowing without and with the combination of the different elastic solutions – rubber particles or USPs) to restore initial position. The box dimensions (1 m x 1 m, and 500 mm in height) allowed for the introduction of a piece of concrete sleeper (250 mm in width and 850 mm in length) with a flexible fastening system and a piece of rail type UIC-54 with a length of 250 mm. The height of the ballast layer was 300 mm under the sleeper, which is a common value found in railway tracks, and allows for maintenance tasks (UIC Code 719-1). The ballast used in this study was composed of ophite, with appropriate size (mainly between 63 mm and 31.5 mm, EN 933-1) and resistance to fragmentation (lower than 8%, EN 1097-2) for its application in railway tracks according to EN 13450, as well as the rest of properties were in consonance with such Standard.

2.2 Test procedure

In order to prove the efficiency of the proposed technique (stone-rubber blowing) to reduce the railway track maintenance, the testing plan included the analysis of the behaviour of a track section reproduced in laboratory when different solutions are carried out as maintenance for re-leveling after ballast settlement. These solutions were: stoneblowing process as conventional maintenance technique; stoneblowing combined with the application of different USPs used as rseference elastic element; and the incorporation of diverse quantities of rubber particles (10%, 25% and 50% over the total volume of the mix of particles blown) used to replace part of the small stones during stoneblowing process (stone-rubber blowing technique). Figure 2 represents the scheme of the different solutions applied.

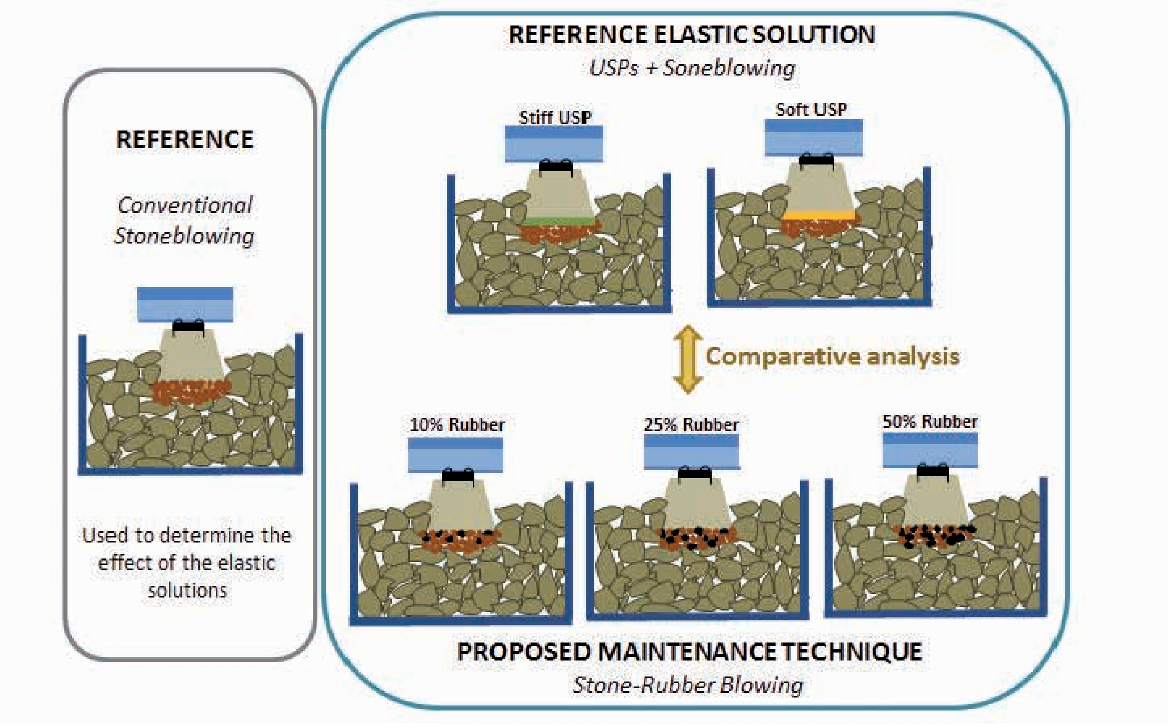


Figure 2. Testing plan.

For this study, a series of dynamic tests were carried out in order to reproduce track section settlement, and then, to simulate maintenance techniques (conventional stoneblowing as well as the combination of this process with the incorporation of elastic elements) for analysing the effect of each solution on track section behavior. The dynamic test consisted of applying 200,000 loading cycles at 4Hz of frequency with a maximum stress over the ballast surface around 250 kPa. These loading conditions are considered appropriate for simulating the effects of passing trains (Indraratna et al., 2006; Bach, 2013; Sol-Sánchez, 2016a). For each maintenance solution, this test began over the conventional section (rail, sleeper, fastenings, rail pad and ballast, without USPs and without blown stones), and it was stopped when the settlement was close to 15 mm and the loading cycles reached approximately 50,000. Following this procedure, a maintenance task was carried out to recover the track geometry, whereupon the test continued until its completion (200,000 cycles). This method was applied twice for all the techniques studied.

The conventional stoneblowing process was conducted after the ballast settlement (up to approximately 15 mm, as indicated previously) by lifting the rail-fastener-sleeper system, and adding the small stones between the sleeper and the ballast surface, which allowed for recovering the original position of the sleeper. In this laboratory work, the process of adding the stones was developed by hand (without compressed air) since the main objective of this initial study was to show the effect of the elastic components, and in particular to examine the feasibility of using rubber particles as flexible stones. The volume of stones and distribution under the sleeper was in agreement with the findings of other authors (Nutbrown and Nicholas, 1999; Tutumluer et al., 2015). Finally, the sleeper piece was returned to its original position (corresponding to its location before ballast settlement), and the dynamic test was continued.

Regarding the combination of stoneblowing with the different elastic elements, in the case of USPs, these components were glued to the bottom of the sleeper by using epoxy resin while the process of stoneblowing was carried out, and then the sleeper with the pad was again placed over the ballast and the stones blown. In reference to the rubber particles (RP) used for stone-rubber blowing process, they were mixed with the natural stones so that both materials were applied at the same time. The quantities of rubber analyzed were 10%, 25% and 50% over the total volume of the mix of elastic particles and natural stones.

3. Analysis of results

Figure 3 analyses the influence of adding different quantities of rubber particles (as replacement of natural small aggregates used during stoneblowing intervention) on the variation in mechanical performance of the global section in comparison to the effect of only apply stoneblowing process (without including innovative elastic solutions). Also, as a referent elastic element to analyse the effect of rubber particles, Figure 3 shows the impact of including USPs in combination with stoneblowing process.

It is possible to understand that the inclusion of USPs and rubber particles during stoneblowing operation, leads to an important reduction in track stiffness while increasing the damping capacity of the section, limiting the negative effect of stiffening of the track when conventional stoneblowing is applied. Also, results show that the use of different percentages of rubber particles allows for gradually varying the track performance, obtaining comparable results to those measured for USPs. Indeed, the replacement of more than 50% of stones with rubber particles presents a comparable reduction in section stiffness to that obtained with soft USPs while a percentage around 10% could be appropriate to use this solution as alternative to stiff USPs. Then, it is seen that the mix of different ratios of rubber particles and stones for track maintenance would allow for the optimization of its vertical stiffness, which is particularly beneficial in transition sections where a gradual change in stiffness is required.

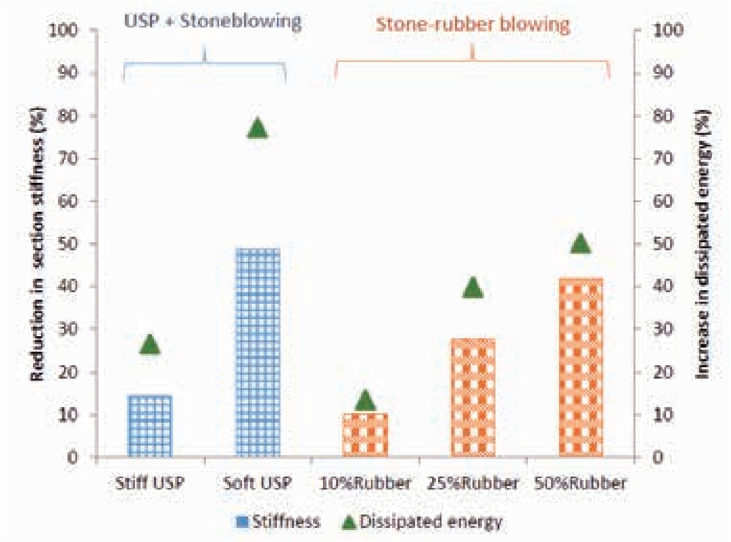


Figure 3. Influence of rubber dosage to modify track performance in reference to conventional stoneblowing, comparing to the effect of USPs.

Regarding the effect of stone-rubber blowing on the settlement trend of the whole track section after maintenance intervention, Figure 4 represents the reduction in total settlement at the end of the test as well as the ratio of settlement per load cycle, in reference to the conventional stoneblowing process. Also, results are compared to those measured when USPs are used as control to analyse the effect of the flexible aggregates.

Results demonstrate that the inclusion of rubber particles during stoneblowing intervention leads to a positive effect on track geometry by reducing cumulative ballast settlement, this fact being more remarkable when increasing the quantity of rubber applied. In comparison to the impact of USPs, it is seen that the quantities of rubber that led to a comparable change in track performance (10% and 50% of rubber in reference to stiff and soft USPs, respectively), allow for a higher reduction in track settlement at the end of the test, and what is more remarkable, a reduction in the long-term trend to settlement. This implies that incorporating rubber particles during stoneblowing seems to be more effective than the application of USPs, which could reduce the need for high-frequency maintenance checks.

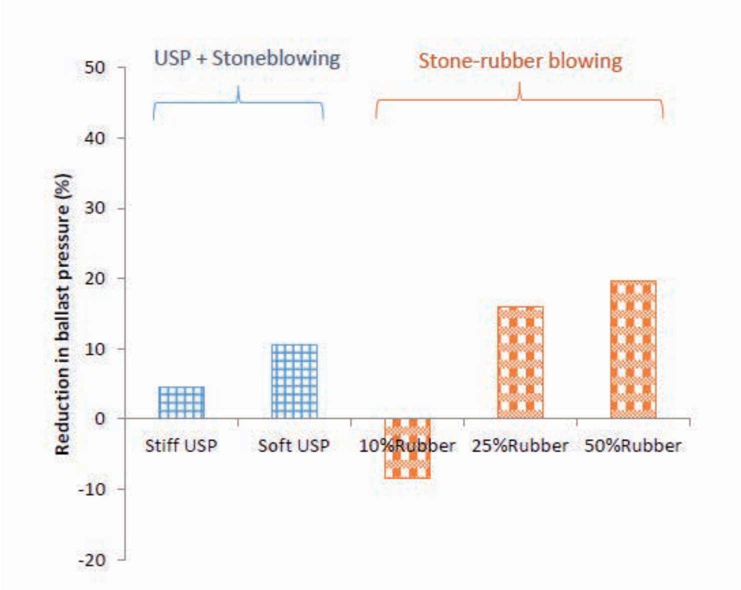


Figure 4. Effect of including elastic solutions to improve the effectiveness of stoneblowing process on reducing track settlement.

On the other hand, Figure 5 compares the effect of rubber particles and USPs on reducing the stress on ballast layer, in reference to the conventional stoneblowing process. Considering that the application of such conventional technique already allows for reducing pressure in the ballast (as a result of the increase of the surfacing contact between sleeper and ballast when smaller stones are added) (Coenraad, 2001; Sol-Sánchez et al., 2016b), results show that the solution with 10% of rubber presented lower reduction than the conventional stoneblowing task, obtaining then a negative decrease in comparison to such reference technique. This is likely to be related to the flexibility of these few particles (having a high capacity to deform under low stress) that cause the concentration of pressure on the natural stones in contact with sleeper.

However, when the quantity of rubber was increased (and there are also rubber particles between the sleeper and the small natural stones), the pressure on the ballast layer was reduced, showing even higher reduction in pressure than that measured for the solutions with USPs. This is due to the strong capacity of the rubber particles to damp loads, associated with their size (up to 20 mm), which is higher than the thickness of the USPs studied.

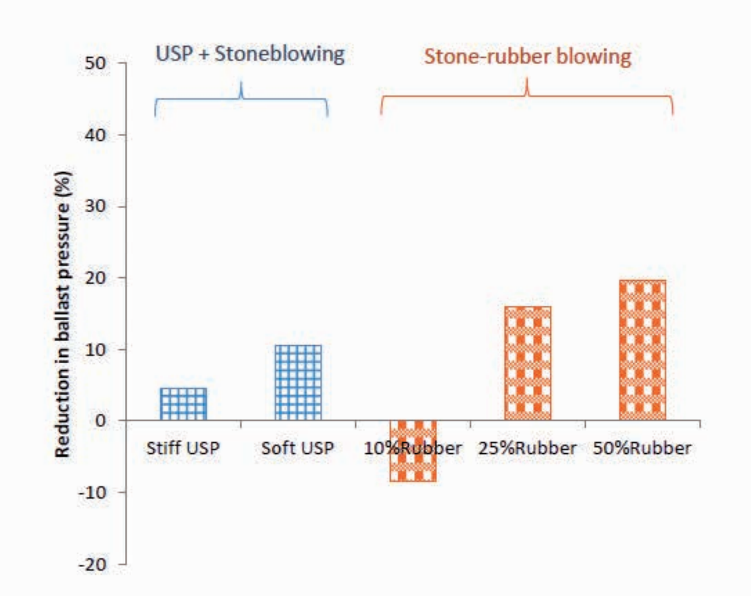


Figure 5. Impact of the elastic solutions on the reduction in ballast pressure, in reference to conventional stoneblowing.

With the aim of analysing ballast degradation after each type of maintenance task based on elastic solutions, Figure 6 displays the effect of combining stoneblowing with USPs and the inclusion of rubber particles as flexible aggregates on two parameters (Index of particle breakage - Bg; Ballast Breakage Index - BBI, %) that have been previously used in other studies to evaluate ballast breakage (Marsal, 1967; Indraratna et al., 2005). As the values recorded for the solutions with the same type of elastic element (USP or RP) were quite similar, Figure 6 shows the mean value recorded for the reduction in degradation when stoneblowing process is combined to both USPs (stiff and soft) and to the different quantities of rubber particles.

The results reflect the fact that both elastic solutions allow for a significant reduction in ballast breakage (regardless of the parameter selected), this fact being more marked when rubber particles were used during the stoneblowing task. This could be related to the better distribution of the stress and lower movement of the ballast stones, which is in accord with the results of ballast settlement and pressure under the granular layer. Thus, it is clear that the use of stone-rubber blowing could lead to an important increase in the durability of ballast, and therefore, in the service life of ballasted tracks.

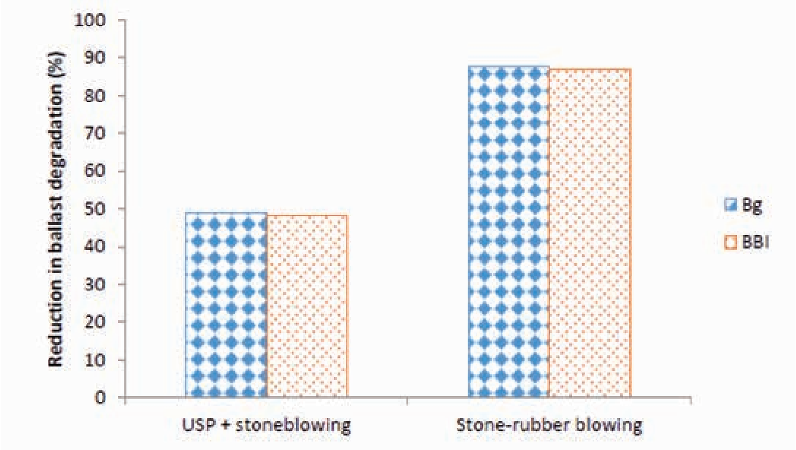


Figure 6. Effectiveness of the elastic solutions to reduce ballast degradation, considering stoneblowing process as a reference.

4. Conclusions

This paper presents a study about the effect of adding rubber particles as flexible stones during the process of stoneblowing, regarded as a new maintenance task (stone-rubber blowing) for railway tracks that combine the benefits of both stoneblowing and USPs (but without the need of bonding such elastic elements) by obtaining an elastic granular layer that fill the gap between the sleepers and the settled ballast layer. With this purpose, this paper shows the influence of adding different quantities of rubber particles as flexible stones at the time that its effect is compared with that recorded when conventional USPs (with different stiffness) are used in combination with the traditional maintenance tasks (tamping and stoneblowing). Based on the results obtained in this study, the following conclusions can be drawn:

- The inclusion of rubber particles during the stoneblowing process improves the effectiveness of this task to reduce ballast settlement and degradation, at the same time that the flexibility and capacity of the granular layer to damp loads is increased, and therefore, longer service life of railway tracks could be obtain.
- The higher the quantity of rubber particles used as flexible stones, the higher the reduction of ballast settlement (at both short and long-term) and vertical stiffness while higher capacity to dissipate the energy transmitted to the ballast is recorded as well as lower stress under such granular layer.
- The use of high quantities of rubber particles (around 50% over the total volume of particles blown) during the process presented as stone-rubber blowing, leads to a comparable (or even better) behaviour of the ballast than that recorded when soft USPs are employed in combination with the conventional stoneblowing process (the most favorable case among the conventional maintenance tasks).
- On the other hand, lower volume of rubber could be comparable to the cases of conventional stoneblowing where stiffer USPs are applied, being able to obtain a gradual behaviour of railway tracks by varying the quantity of rubber employed during the process of blowing. This would allow for smooth stiffness transitions along the track.

Based on the results, it is possible to say that the proposed technique of adding rubber particles to partially replace the stones leads to similar or enhanced benefits in comparison with those obtained when other elastic components such as USPs are applied. Moreover, use of this technique could yield a substantial reduction in maintenance costs, since these rubber particles could be combined with the small stones during the stoneblowing process. Nonetheless, further studies are required to demonstrate the effectiveness of this solution, analysing some issues such as the expected longevity of rubber particles, possible impact on track drainage, and its real-scale application using a stoneblower to assess any possible segregation of the stones and rubber particles.

5. References

- Bach, H. (2013) Evaluation of attrition tests for railway ballast. PhD Dissertation, Graz University of Technology, Austria, 2013.
- Claisse, P. A.; Keedwell, M.; Calla, C. (2003) Tests on a two layered ballast system. Procedings of the ICE - Transport, vol. 156, pp. 93-101.
- Dahlberg, T. (2010) Railway Track Stiffness Variations-Consequences and Countermeasures. International Journal of Civil Engineering, Vol. 8, pp. 1–12.
- Fair, P. (2003) The geotechnical behaviour of ballast materials for railway maintenance. University of Sheffield, PhD Dissertation.
- Fontserè, V.; López Pita, A.; Manzo, N.; Ausilio, A. (2016) NEOBALLAST: New high-performance and long-lasting ballast for sustainable railway infrastructures. Transportation Research Procedia 14, 1847-1854.
- Hellawell, A. (1997) Stoneblower provides ballast maintenance solution. Railway Gazette International(September), pp. 595-596.
- Indraratna, B.; Khabbaz, H.; Salim, W.; Christie, D. (2006) Geotechnical properties of ballast and the role of geosynthetics. Journal of Ground Improvement, Vol. 19, pp. 91-102.
- Indraratna, B.; Lackenby, J.; Christie, D. (2005) Effect of confining pressure on the degradation of ballast under cyclic loading. Geotechnique, Vol. 55, pp. 325–328.
- Indraratna, B.; Salim, W.; Rujikiatkamjorn, C. (2011) Advanced Rail Geotechnology - Ballasted Track.
- Marsal, R.J. (1967) Large scale testing of rockfill materials. Journal of Soil Mechanics and Foundation Engineering. ASCE, Vol. 93, No. SM2, pp. 27-43.
- Nutbrown, D.; Nicholas, R. (1999) Stoneblower experiences. PWI Autumn Technical Seminar.
- Plica, P. (2007) Sleeper Solicitation/Sleeper design. Work Package number 2b of UIC Project Under Sleeper Pads.
- Selig, E.T.; Waters, J.M. (1994) Track Geotechnology and Substructure Management, T. Telford.
- Sol-Sánchez, M.; Moreno-Navarro, F.; Rubio-Gámez, M.C. (2014) Viability of using end-of-life tire pads as under sleeper pads in railway. Construction and Building Materials, Vol. 64, pp. 150-156.
- Sol-Sánchez, M.; Pirozzolo, L.; Moreno-Navarro, F.; Rubio-Gámez, M.C. (2016a) A study into the mechanical performance of different configurations for the railway track section : A laboratory approach, Eng. Struct. 119, pp. 13–23.
- Sol-Sánchez, M.; Moreno-Navarro, F.; Rubio-Gámez, M.C. (2016b) Analysis of ballast tamping and stone-blowing processes on railway track behaviour: the infuence of using UPS's, Geotechnique.
- Tutumluer, E.; Staplin, D.; Trosino, M.; Smak, J.; Tomas, M. (2015) Stone blowing as a remedial measure mitigation differential settlement at bridge approaches. Presented at 94th Annual Meeting of the Transortation Research Board, Washington, D.C.
- UIC (Union International of Railway) (2011) Recommendations for the use of Under Ballast Mats, UIC Code 719-1.

360.revista de alta velocidad
número 5 - junio - 2018. Pág 323 - 331

Control and Maintenance of Railways through Satellites. Is it possible?

Mayoral, Juan
Jiménez, Pablo Lorenzo
Fernández-Sánchez, Gonzalo
Jardí, Ignacio

Universidad Europea de Madrid[1]
Ferrovial

Abstract

The framework of this paper falls within the needs of optimization of the expenses in maintenance in railroads, contemplating the cost of the infrastructure from a life cycle approach.

In the last 25 years, there has been a strong investment in new high-speed railway infrastructures. After putting a new infrastructure into service, it is very important to avoid losses of the value of this asset by means of adequate maintenance actions. Given the economic situation of the last years, it is necessary to optimize the resources allocated to railway maintenance, with criteria of efficiency and austerity.

In this paper, new sources of information have been sought to determine the state of the infrastructures and their maintenance needs. It analyses the possibility of integrating aerospace knowledge to carry out these controls and monitoring tasks through satellites as a substitution of control and visual monitoring or, like the latest innovative proposals, using drones.

Within the capabilities of the satellites, a more detailed evaluation of those that could be useful for railway maintenance, such as satellites with capacity to measure soil moisture, to take high resolution images or to control the surface movements.

Throughout this paper, it is intended to make a comparison between the necessary measurement ranges in railway maintenance and the possibilities that give us the artificial satellites.

Keywords: maintenance; railways; satellites.

---

[1] Mayoral, Juan. Universidad Europea de Madrid. Email: jmayoralramos@gmail.com
Jiménez, Pablo Lorenzo. Universidad Europea de Madrid. Email: pablolorenzo.jimenez@universidadeuropea.es (corresponding author)
Fernández-Sánchez, Gonzalo. Universidad Europea de Madrid. Email: Gonzalo.fernandez@universidadeuropea.es
Jardí, Ignacio. Ferrovial. Email: Ignacio.jardi@ferrovial.es

1. Introduction and Objectives

The applications of the satellites have been developed in different fields, from applications in agriculture, for the control of the crops, through the maritime control of boats, allowing their constant location in zones of high danger, to the control of fires, droughts, flooding or the control of pollutants in the atmosphere.(Duncan, Prados, & Lamsal, 2014)

Within the field of engineering, different applications have been developed, such as the control of surface movements in tunnel covers or slope movements (Guzzetti, Cesare Mondini, & Manunta, 2012), or applications to improve the layout of roads (Copernicus, 2014), but the low resolutions available until relatively recently have prevented a greater contribution to this field.

With the development of satellites, obtaining resolutions lower than the meter, together with the development of image processing and the evolution of radar, this technology becomes an important field of research for the future.

This research includes the following tasks: to study new ways of optimising the conservation and maintenance of railway infrastructure using aerospace technology, more specifically via satellites. The main objective is to find a way of optimising resources for the conservation and maintenance of railways by using artificial satellites.

To this end, intermediate objectives are established that allow us to reach the main one. Firstly, it will be necessary to carry out a study of the current operation of railway maintenance, in order to establish railway's needs. Within the rail system, we focus on those parts that can be measured from satellites.

In addition, we seek a better understanding of the different types of existing satellites and the possible applications offered by each of them. In this way, the aim is to compare the maintenance needs with the possibilities offered by satellites in the control and maintenance of railway tracks to show future practical applications.

2. Methodology

In order to achieve the objectives set, an analysis of conventional railway maintenance will be carried out. Based on this analysis, the elements to be controlled and the measuring ranges used in maintenance operations are determined.

On the other hand, we study the different capacities offered by artificial satellites, establishing also the measuring ranges. Different visits have been made to expert aerospace centres based on the visualisation and control of satellite data in the Madrid sector (AirBus, European Space Agency ESA, academy), as well as interviews with experts and search for technical information in the aeronautical sector.

Finally, a comparison matrix is developed to determine which satellite technologies could be implemented.

3. Results

3.1 Railway maintenance

The standard UNE-EN 13306: 2011 defines maintenance as the combination of all technical, administrative and management actions, during the life cycle of an element, intended to preserve or return it to a state in which it can perform the function required.

In any element, we can distinguish two main types of maintenance, according to the occurrence of the failure: preventive maintenance, when maintenance aims to avoid failure, and corrective maintenance, which is done when the failure of the element has already occurred.

Preventive maintenance operations aim to minimize the likelihood of a failure in the elements, due to the associated operating risks and the higher cost that the corrective maintenance associated with the failure entails. Within preventive maintenance, we can distinguish the predetermined maintenance (which is performed according to established time intervals or operating units), and maintenance based on the condition, where the condition of the element is monitored or inspected to determine the actions to be undertaken.

The railways are composed of a multitude of elements, with different modes of degradation and maintenance operations. To determine how we can facilitate or improve maintenance through the use of satellites, it is necessary to analyse the operations necessary to know the state of the different subsystems. These surveillance operations can be done both on foot and on the road on railway vehicles, whether these are specific to these tasks (or other maintenance), commercial vehicles. All these operations have in common that they are carried out in the own infrastructure. Surveillance tasks that can be remotely controlled (such as knowing the occupation of a track circuit) are not the subject of the present study.

Each of the subsystems that compose the railway system is analysed, in order to ascertain the operations that can be carried out by satellites. In the infrastructure subsystem1, we distinguish the track from the platform. The first includes the following elements: rails, sleepers, fastenings, track foundation (ballast and sub-ballast), or concrete slabs, as well as other devices that are placed on them: diversions, escapes, sleepers, expansion devices, etc. In the second, we include the different types of civil works that we find in a railway infrastructure: natural platform (embankments, trenches, form layer, etc.) and artificial platform (tunnels and viaducts), as well as their drainage systems. The surveillance and inspection operations in this subsystem focus on the following:

In the case of the track, it is interesting to know both the state of the elements and the geometric quality of the elements.

- In the case of the condition of the elements, the main operations are the control of the rails (wear and internal state by ultrasonic auscultation), as well as control of the condition of the materials (ballast, sleepers, joints, fastenings, welds, Track, etc.).
- In relation to the geometric quality, it is determined by five parameters: leveling, longitudinal and transverse, alignment, gauge and warpage (UNE-EN 13.848-1: 2004). Measurement of geometric quality is a way of measuring the state of the set, as it is influenced by the state of the track elements themselves and by the underlying support structures (track foundation and platform).

Regarding the platform, monitoring operations depend on the type of element. In embankments and trenches, movements are sought, paying special attention to the blocks that can be detached and fall on the track in excavation zones.

In relation to the condition of the elements, the following specific aspects of certain elements should be indicated:

- Ballast: it is necessary to verify the possible contamination of the ballast. The contamination is the ballast dust that stems from its crushing, and which is characterized by the whitish colour that appears in the surface of the ballast layer. Other parameters to check are the dimensions of the layer, the slopes, the ballast shoulder, the amount of existing ballast and the presence of vegetation.
- Sleepers: the maintenance operations consist on checking that are no cracks, and that no

voids exist under the sleepers, which could compromise the transmission of the load to the ballast.

- Fastenings: during maintenance operations, it must be checked that they are correctly tightened, that they are correctly positioned and that they have not been removed.
- Bridges and tunnels: check the structural condition, verifying that there are no cracks in walls, boards, etc.
- Drainage systems: cleanliness of pipes, taps, culverts, siphons and gutters and the presence of herbs will be monitored.
- Surroundings: there mustn't be any obstacles that could obstruct access to the road by maintenance personnel, and that the perimeter protections have not undergone any alteration.

The operations to determine the status of this subsystem are as follows:

- Auscultation: measurement of parameters of the track and its elements. We distinguish several types:
  - Geometric auscultation: it measures the parameters of geometric quality previously mentioned. For this purpose, a geometric control car is used, of which there are several examples depending on the network. In addition to the control performed by this type of trains, the geometrical control of the track can be performed by manual or topographic methods or even using the records of the batting machines during their operation. o Dynamic auscultation: it measures and records accelerations in bogies, journal box and vehicle box. Each of the accelerations measured by the auscultation train reflects different path defects. They also determine the comfort of the traveller.
  - Ultrasonic auscultation of lanes: Ultrasonic auscultation allows the detection of internal rail defects. This operation is not the object of this work, since this type of defects is not observable by means of satellites.
- Walking routes/ train rides: routes and rides have a fundamental role in preventive maintenance activities, since they are intended to verify the elements whose parameters are not controllable through the different types of auscultation previously described. They can be on foot (routes) or inside the train (rides): the former are more exhaustive, but much slower; the second are faster, and allow to evaluate the dynamic behaviour of the trains. The periodicity for the realization of both types is marked by the infrastructure manager.

In the energy subsystem, the determination of the state of the mechanical elements (posts, contact wires, supports, etc.) is done by the use of special vehicles. The parameters that are measured are the geometric quality of the catenary, and the wear of the wire. The status of the catenary posts can be determined by means of train rides.

In the control, command and signalling subsystem, most of the elements are remotely controllable (such as signal status). For this reason, most of the maintenance operations are either periodic, or are based on the warnings that the different systems give about their operation. Train rides may be used to determine the condition of cable conducts (open caps or misalignment), or the condition of signal posts.

3.2 Artificial Satellites

There are different classifications of artificial satellites. Depending on their use, these include meteorological satellites, military satellites, telecommunications satellites and so-called earth observation satellites. Within the terrestrial observation satellites, we can distinguish between optical satellites and satellites of image capture with synthetic aperture radar technology.

- Optical satellites are characterized by their high spatial resolution and the range of the electromagnetic spectrum they capture, being influenced by the presence of clouds in the imaging.
- Satellites with synthetic aperture radar technology are not conditioned by the presence of clouds, since their operation is based on the emission of electromagnetic pulses and the capture of the rebound.

Both images of optical satellites and those obtained by radar are subjected to rectification processes such as orthorectification or interferometry.

In addition to classification based on the function they perform, another differentiation in the typology of satellites can be made depending on their orbit. The orbit of a satellite is characterized, firstly, by the centre of rotation (geocentric or heliocentric). In the geocentric satellites, the orbit is classified according to the height, where HEO (High Earth Orbit) exist, above the 35,786 km; MEO, between 20,000 and 35,786 km; and LEO, below 20,000 km (Gaetano Hadad & Blanco Arias, 2009). The latter usually get better resolutions, so they will be those studied in the present work. LEO satellites are divided into orbits synchronized with the Sun, or polar.

The type of orbit and its altitude will characterize the period of passage through the same point, and therefore the frequency of capture of images at the same point that they will have. There is a large amount of satellite data applicable to different fields, given the large amount of information they generate, although there is still some difficulty in interpreting and processing such information. Some of the possible applications are described below.

One of the areas where significant progress has been made is the control of pollutants produced in industrial sectors, allowing the monitoring of columns of contaminants produced in farms and fires. In addition, by observing the concentration of different gases, it is possible to support air quality prediction models, including the generation of new models, and to estimate anthropogenic emissions, thanks to the ability of these satellites to measure the number of molecules of certain gases, like NOx, SO2 in scarce emissions and NOx, CO and CH4 in larger emissions, existing between the Earth's surface and the satellite. These measurements allow the creation of maps of temporal evolutions of emissions, from daily to annual scales, and the monitoring of concentrations (Duncan et al., 2014).

To develop these tasks, agencies like NASA have developed satellites such as Terra, Aura or GOES. However, this use has certain limitations, such as the lack of information regarding the vertical distribution and composition of these columns. However, they do not seem to be applicable in our field, since there are no large concentrations of pollutants in the railway system.

The use of satellites to control soil moisture is mainly focused on improving weather forecasts, drought control and the possibility of associated fires, predicting possible floods and improving water use on farms (NASA, 2004b). This satellite typology uses synthetic aperture radar (SAR) systems for the detection of moisture from the surface layers of the earth. SARs are so-called active sensors, which record high spatial resolution images (Sillerico, Marchamalo, Rejas, & Martínez, 2010). These radars measure the electromagnetic radiation of any body that is at a given temperature (Universidad Politécnica de Cataluña, 2009). This radiation will depend on the composition of the soil, so applied techniques of interferometry can obtain the soil moisture and salinity of the oceans. SAR interferometry is an applicable technique in different areas. It involves the use of a satellite to record two or more images of the same area at different time points. When comparing the different images, it is possible to detect any changes that may have occurred in that period of time (AIRBUS, 2012).The most important satellites in this field are the SMAP of NASA and the SMOS of the ESA. They differ from satellites with radar technology because of their ability to take pictures in different spectra.

Flexibilization of regulations that prevented high resolution images from space has allowed resolutions of up to 0.30 m to be reached in panchromatic mode (Satellite Imaging Corporation, 2015), characterized by an observation made in a single spectral band located in the visible part of the spectrum electromagnetic (Spot Image, 2005).

The applications of the high resolution images in the railway maintenance can vary, from the control of the existing vegetation near the track, until the observation of changes of colour in the infrastructure that can indicate different problems.

Within this field there is a very wide offer, so that only cite some with different resolutions and operated by different companies:

- WorldView-3: resolution of 0.30 m.
- WorldView-2: resolution 0.46 m.
- Pleiades: 0.50 m resolution.
- Spot 6/7: resolution of 1,50 m.
- KazEOSat-1: resolution of 1.00 m.
- Kompsat-3: resolution of 0.70 m.
- RapidEye: resolution of 5m.

Satellites capable of detecting surface movements on slopes are based on the same technologies as soil moisture detection, i.e. SAR satellites, differing from these at the wavelength at which they emit. While satellites whose objective is to measure soil moisture emit in wavelengths around 21 cm and a frequency of 1.4 GHz, band L of the electromagnetic spectrum, the satellites used in the detection of surface movements work in variable wavelengths depending on the type of terrain and in the environment of 3.1 cm and 9.6 GHz (c and x bands of the electromagnetic spectrum).

From the techniques of interferometry, distance data are obtained between the satellite and the terrain surface, calculated by measuring time and time lags. This technique, also known as InSAR, of combination of 2 interferometric passages allows to measure phase differences in two satellite steps in the same zone. These differences can be due to the different position between the two satellite trajectories or because of a displacement of the observed area, so this technique allows the detection of movements and deformations of the Earth's surface. The signal received by the sensor is influenced by parameters such as atmosphere and surface movement (Sillerico, Marchamalo, Rejas, & Martínez, 2010).

Radar images are black and white, although colour images can be produced by combining 3 independent images of different dates, forming a composite image (European Space Agency, 2010).

From the InSAR technology, the differential SAR or D-InSAR interferometry is generated, which allows the generation of maps of the terrain displacement, from “n” steps of the satellite through the same zone. For the processing D-InSAR different techniques already exist. The most used is to obtain displacements from the areas away from the study area, which are assumed to be free of movement and with constant geometry. The DINSAR technique allows to eliminate some of the parameters that influenced INSAR interfeormetry (Sillerico, Marchamalo, Rejas, & Martínez, 2010).

The resolution of this type of images depends on the size of the images taken, varying this in each one of the satellites. Some of the satellites available today are:

1. Terrasar-X: its resolution in the X and Y planes vary between 0.25 m and 3 m, while their accuracy in the measurement of subsidence can reach 3.1 mm.

2. CosmoSkyMed: constellation formed by 4 satellites, located in synchronous orbits to the sun. Its resolution in the x-y plane varies between 1 and 5 m and its accuracy in the z-axis is around cm (Italian Space Agency, 2016).
3. RadarSat-2: Its spatial resolution on the earth's surface varies between 1 and 5 m depending on the image size captured, while the accuracy that can be achieved in the measurement of subsidence is 5.6 mm (Canadian Space Agency , 2015).

3.3 Feasibility of using satellites in railway maintenance

In order to study the usability of satellites in maintenance, a comparison has been made of the measurement ranges of each of the satellite capacities with the measurement ranges used in railway maintenance.

- **High resolution images:**

High-resolution images can be used to observe the variation of distances. For this it is necessary to take into account the highest resolution available in the market today, 25 cm in the horizontal plane, reachable only by the Worldview-3 satellite. This factor immediately discards the possibility of measuring path quality parameters located in millimetric ranges. In addition, it discards the defects of cracks present in sleepers or sleepers, nor the surface defects existing in the rails. Another element that could not be controlled by the use of satellite images are the fastenings between rails and sleepers, since their small size discards this possibility. However, we could control the ballast shoulder, distance between the lane and the beginning of the bank slope, which can vary between 90 and 110 cm.

On the other hand, high resolution images allow us to observe changes in colour between each satellite pass. This variation of colour can give us very useful information for the control of different elements:

- Emergence of vegetation in undesirable areas.
- Detection of areas of surface moisture where it would not be expected, which may reflect poor drainage or obstruction of gutters.
- Control of the existence of fauna or burrows that can deteriorate or weaken the embankments of the infrastructure.
- Contamination of the ballast, characterized by a whitish colour produced by the elevation of the fines of the ballast itself to the surface.
- Existence of obstacles in the access roads.
- Fence perimeter elements and fallen noise attenuation.
- Signalling vertical elements and fallen catenary posts.
- Slopes landslides and landslides.
- Existence of objects in the drainage elements such as ditches or pipes.

- **Humidity:**

The use of satellites such as SMOS or SMAP in the detection of surface soil moisture in rail maintenance has been ruled out due to the spatial resolution of these systems, which exceeds 10 km, and prevents the measurement of problems in linear elements such as Drainage of railway systems. However, it is important to highlight the numerous applications that these satellites could have if they had their spatial resolution improved.

- **Contamination:**

As mentioned in previous chapters, the application of satellites used in pollution control in industrial or urban areas is not applicable to railway systems, since there are no large concentrations.

- **Control of surface movements:**

Satellites studied reach precisions below the centimetre in the vertical axis, making possible the monitoring of superficial movements in earthworks. This control in traditional maintenance is done by sending surveyors to areas where there have already been appreciable movements by the human eye. This way of operating has a high cost and sometimes the control comes too late causing great deterioration in the road.

4. Conclusions

There is a need to find alternative cost-saving alternatives to traditional maintenance, which uses sometimes expensive methods, such as the use of surveyors and pedestrians. They also cannot detect certain deficiencies in the railway infrastructure.

The development of aerospace technology, and specifically artificial satellites, introduces the possibility of an improvement in the control of different elements of the railway platform. These systems allow us to obtain resolutions of up to 0.25 meters, in the case of optical satellites, which we can consider sufficient for the measurement of some important distances in the maintenance like the shoulder of ballast The mentioned resolution together with the possibility of obtaining the colour of each pixel of the image allows the detection of the existence of obstacles or fallen elements in the surroundings of the railway platform, in the elements of superficial drainage or in the own railway infrastructure.

The presence of damp areas that indicate a malfunction of the drainage systems and the accumulation of water in undesirable places can be detected, by means of the control of the colour changes, thus facilitating the improvement in the prevention of possible problems in these systems. In addition, this capacity allows us to perform a better control of the existing vegetation, optimizing the planning of the works of weeding and irrigation of herbicides, and with that of the resources available for these works. Another important application of highresolution colour imaging is the control of the existence of wildlife, such as rabbits, that can deteriorate the state of the embankments, leading to landslides.

On the other hand, satellites based on radar technology offer the ability to measure vertical subcentimetric movements, giving the possibility of control of slopes and embankments in problem areas.

5. References

- ACEX. (2007). Documento base del estado del arte de la conservación de infraestructuras en España.
- AIRBUS. (29 de Febrero de 2012). La interferometría con radar de apertura sintética. Recuperado el 1 de Agosto de 2016, de Airbus. Denfense and Space: http://www.space-airbusds.com/es/noticias-articulos/la-interferometria-con-radarde-apertura-sintetica.html
- Canadian Space Agency . (10 de Julio de 2015). Satellite Characteristics. Recuperado el Agosto de 2016, de Canadian Space Agency: http://www.asccsa.gc.ca/eng/satellites/radarsat/radarsat-tableau.asp

- Copernicus. (2014). Keeping Europe Moving.
- Duncan, B. N., Prados, A. I., & Lamsal, L. N. (21 de Mayo de 2014). Satellite data of atmospheric pollution for U.S. air quality applications: Examples of applications, summary of data end-use.
- Duncan, B. N., Prados, A. I., Lamsal, L. N., Liu, Y., Streets, D. G., Gupta, P., & Hilsenrath, E. (2014). Satellite data of atmospheric pollution for U.S. air quality applications: Examples of applications, summary of data end-user resources, answers to FAQs, and common mistakes to avoid. Atmospheric Environment, 647-662.
- European Space Agency. (2010). Tecnología de radar. Recuperado el 2016, de Eduspace: http://www.esa.int/SPECIALS/Eduspace_ES/SEMGV7E3GXF_0.html
- Gaetano Hadad, O. J., & Blanco Arias, A. (2009). Monografías. Recuperado el Julio de 2016, de Los satélites de comunicaciónes: http://www.monografias.com/trabajos72/satelites-comunicaciones/satelitescomunicaciones2.shtml
- Guzzetti, F., Cesare Mondini, A., & Manunta, M. (30 de Noviembre de 2012). Using GMES to map and monitor landslides and ground subsidence. GMES4REgions.
- Italian Space Agency. (2016). COSMO-SkyMed Mission and Products Description.
- López Pita, A. 2006. Infraestructuras ferroviarias. Colección Temas de Transporte y Territorio. Edicions UPC.
- NASA. (2004b). Soil Moisture Active Passive. Recuperado el 30 de Julio de 2016, de Jet propulsion Laboratory Californian Institute of Technology: http://smap.jpl.nasa.gov/
- Satellite Imaging Corporation. (2015). WorldView-3 Satellite Sensor. Recuperado el 2016, de Satellite Imaging Corporation: http://www.satimagingcorp.com/satellitesensors/worldview-3/
- Sillerico, E., Marchamalo, M., Rejas, J. G., & Martínez, R. (2010). La técnica DInSAR: bases y aplicación a la medición de subsidencias en la construcción . Informes de la Construcción, 47-53.
- Spot Image. (2005). Resolución y modos espectrales. Recuperado el 2016, de Spot Image: http://www.intelligence-airbusds.com/files/pmedia/public/r1996_9_resolucion_modos_espectrales__esp_sept 2010.pdf
- UNE-EN 13.306:2011. Maintenance - Maintenance terminology
- UNE-EN 13.848-1:2004. Railway applications - Track - Track geometry quality - Part 1: Characterisation of track geometry
- Universidad Politécnica de Cataluña. (2009). La UPC controla los sensores del satélite SMOS, en la misión de la ESA para estudiar el cambio climático. Recuperado el 30 de Universidad Politécnica de Cataluña. (2009). La UPC controla los sensores del satélite SMOS, en la misión de la ESA para estudiar el cambio climático. Recuperado el 30 de Julio de 2016, de UPC: http://www.upc.edu/saladepremsa/al-dia/mesnoticies/2009/la-upc-controla-los-sensores-del-satelite-smos-en

6. Acknowledgements

Authors would like to acknowledge the collaboration of Ferrovial and Universidad Europea de Madrid in the development of this initial idea.

High Speed Rail Infrastructure

360.revista de alta velocidad
número 5 - junio - 2018. Pág 333 - 353

Influence of Aerodynamics in Tunnels Design

Sepúlveda Abad, Silvia
Lastra de la Rubia, Emilia

Ineco[1]

Abstract

There are a significant number of factors to consider during tunnels design in order to increase effectivity in construction, reduce schedules, and provide more appropriate cross-sections for each tunnel depending on its location, conditions of exploitation and rolling stock among others.

When a high-speed train enters in a tunnel there is an overpressure and depression in head and tail that propagates along it at the speed of sound, reflecting each time they reach the portals of the tunnel. These variations of pressure, in addition to increasing the aerodynamic drag of the train, can affect the health and comfort of passengers.

Aerodynamically speaking, it is recommended meeting health and comfort criteria. The health criterion is mandatory in accordance with the Infrastructure Technical Specification of Interoperability (TSI). The comfort criteria are variable depending on the country and are usually more restrictive in terms of sizing the free cross-section of the tunnel.

During the last years, designers have developed methodologies and tools to use in calculations for achieving the best tunnel conditions for all the counterparts.

There are two consequences extracted as a result of the development of the methodologies. Firstly, the reduction of the free cross-section of the tunnel to the minimum values as required by the geometrical conditions. Secondly, being able to optimize train speed along the tunnel while complying with the aforementioned criteria, which leads to reductions in cost, improved conditions of exploitation and reduced travel times.

However, aerodynamics research cannot be limited to these studies. The future will require speed improvements and longer tunnels while keeping the same standards of safety and comfort. Currently new factors such as overpressures or depressions in tunnel installations or sonic boom effect are being investigated and included in project requirements.

In relation to this, INECO has developed a software capable of calculating the aerodynamic effects in every specific position within the tunnel, train or the installation. The software most interesting functionality is the capability of modelling and mitigate sonic boom effect.

Keywords: Aerodynamics, free cross-section, tunnels, speed, comfort criteria, health criterion.

---

[1] *Sepúlveda Abad, Silvia. INECO. Email: silvia.sepulveda@ineco.com (corresponding author)*
Lastra de la Rubia, Emilia. INECO. Email: emilia.lastra@ineco.com

1. Introduction

When a high speed train enters in a tunnel there is an overpressure and depression in head and tail that propagates along it at the speed of sound, reflecting each time they reach the portals of tunnel. These variations, among other aspects like the increasing aerodynamic drag of train can affect the health and comfort of passengers.

To avoid the above effects designers, have to comply with several criteria. By one hand, the Health Criterion included in Infrastructure Technical Specification of Interoperability (TSI) that is mandatory. On the other hand, Comfort Criteria included in the different recommendations and national standards which at the end result more restrictive in terms of sizing of the final crosssection.

Throughout this paper, in a perspective vision about what these 25 years have been occurring will be transmitted in order to put in the reader's mind the fast evolution and improvements that the Aerodynamics Studies have introduced in the designing of high speed tunnels.

2. Current spanish regulation

In Spain two main standards are actually employed in relation to sizing the free cross-section due to aerodynamic constraints of high speed tunnels.

2.1 Health Criterion

Infrastructure Technical Specification of Interoperability (TSI), 2014, November 18th, regulates the health criteria. In the article 4.2.10.1 Maximum pressure variations in tunnels, states the following:

- Any tunnel or underground structure intended to be operated at speeds greater than or equal to 200 km/h has to provide that maximum pressure variation, caused by the passage of a train running at the maximum allowed speed in the tunnel, do not exceed 10 kPa during the time taken for the train to pass through the tunnel.
- Above requirement has to be fulfilled along the outside of any train complying with the Locomotives and Passenger TSI.

This is mandatory for all cases, and trains are considered non-sealed trains.

2.2 Comfort Criteria

For the regulation of comfort criteria, in Spain there are 2 main standards.

Draft of the IFI ("Instrucción Ferroviaria para el proyecto y construcción del subsistema de Infraestructura"), now in approval stage, states the following criteria:

TRAIN	TUNNEL	Pressure Difference (kPa)	t(s)
Non sealed trains	Single track	2	4
	Double track	4	4
Sealed trains (τ = 6s)	Single and Double tracks	1	1
		2	10

Table 1. Comfort criteria included in draft of IFI

However, ADIF (Spanish Railway Administrator) includes in the standard NAP 2-3-1.0 (2015), in section 3.2 "Comfort criteria" the following summarized in the next table:

TRAIN	TUNNEL	Pressure Difference (kPa)	t(s)
Non sealed trains	Single track	2	4
	Double track	4	4
Sealed trains (τ = 6s)	Single and Double tracks	1	1
		1.6	3
		2	10
		3kPa during all the time the train is in tunnel	

Table 2. Comfort criteria included in NAP 2-3-1.0 (ADIF)

NAP 2-3-1.0 (2015) also indicates that the UIC Leaflet 779-11 (2005) Code cannot be used for sizing, just for pre-sizing of the cross-section. However, during the first years of development of the high speed in Spain, the UIC Leaflet and the Recommendations from the Ministry of Works were also used.

3. First steps

25 years ago or maybe before, aerodynamic issue was a totally unknown phenomenon in Spain. The main reason was that this issue had never been considered because high speed lines were not a reality in Spain.

The way in which as designers considered this aspect was the following. First of all, the UIC Leaflet was considered for both constraints, health and comfort.

These first steps, which are described below, do not have to be the same as those in other countries, since the evolution in this area has been very fast. On the other hand, the dimensioning that is described is due to aerodynamic effects without taking into account the geometrical condition, always existing, and which was always taken into account as will be seen below.

During the firsts years of evolution in the design of tunnels of High Speed Lines, the tools that later would be available, have not been developed yet. This was the main reason why initially the calculations, or rather estimations, were made in Spain by means of the leaflets provided by UIC Code 779-11 (2005).

These leaflets were first developed in 1995, followed by the European Rail Research Institute (ERRI) Specialist Committee C218 that in 2005, through the use of the SEALTUN and AIRSHAFT programs, these international leaflets were created for users. The main idea of these leaflets was to facilitate the sizing of the high-speed tunnel cross-section. Note that all the recommendations, criteria and considerations that included these leaflets were the same ones that were included in the 2008 TSI's.

The calculation factors that the leaflets from UIC Code considered for sizing cross-sections are the following:

- For single track tunnels it is assumed that only one train is in the tunnel at any given time. Although it is true, the current single track tunnels are much longer than then and could allow several trains in a tunnel at the same time.
- For double track tunnels, it is considered that crossing trains can enter at the same time or at different times in the tunnel.

The type of rolling stock that is contemplated is the following:

Type of train	Rate of speed considered (km/h)
Standard Modern train	180 – 220
Streamlined high-speed train	200 - 350

Table 3 Types of trains considered in UIC Code

However, it is also included the following footnote:

> *"For trains and speeds which are outside the scope of this leaflet, it is recommended that specific calculations be undertaken to determine the design area, using a one-dimensional computer simulation method of the type used in this leaflet for instance"*

- Interpolation is allowed when necessary, but extrapolation is not allowed. Furthermore, in this sense it is indicated that if extrapolation is necessary, specific calculations should be performed using a one-dimensional program.
- The passage of freight trains it is not contemplated, therefore if these are necessary the verification must be done by means of specific calculations.

In short, there were a number of constraints that were gradually limiting the use of the leaflets and directing the aerodynamic calculation towards the use of specific uni-dimensional calculation programs that allowed individual modelling for each case. The most important limitation was the length of the tunnels, which were increasingly longer so in some cases extrapolation was necessary and was not allowed.

It was therefore necessary to start using one-dimensional programs.

4. Thermotun

Throughout 2007, at INECO, an innovation project was developed, which consisted of validating the results obtained through the use of THERMOTUN software were the same as the full scale measurements performed in tunnels. While this software was recognized worldwide, as one of the best aerodynamic calculation software, in Spain was unknown, therefore INECO carried out the verification with the collaboration of ADIF.

During the development of the project, measurements were made in the tunnels of the Madrid - Zaragoza LAV, in particular in Dehesillas, Castejón and Bubierca double track tunnels.

Measurements were also made in the single track Guadarrama tunnel of LAV Madrid - Valladolid.

In addition, to assess the passengers exposure to pressures while crossing a tunnel, measurements were made with the inspection train. Pressure gauges were installed within the train for that purpose.

The characteristics of the instrumented tunnels are the following:

Tunnel	Length (m)	Free Crosssection (m^2)	Average speed (km/h)*	Type of train
Dehesillas	861	81	275	S-102, S-103, S-104 y S-120
Castejón	392	115	278	
Bubierca	2,434	76	278	
Guadarrama	28,400	51.025	300	

Table 4 Characteristic of tunnels analysed by Innovative Project

The instrumentation was set as follow:

- Sensors in the tunnel: Three control sections were installed in each tunnel, one in the centre of the tunnel and the other two near the two portals. If the tunnel contained emergency stop areas or interconnection galleries, the control sections were intended to be located close to these zones.
- Air velocity, atmospheric pressure and temperature were measured. Anemometers were placed on the sidewalks and windlasses were installed at the top of the vault. To measure the piezoelectric pressure, sensors of pressure were used in the sidewalks at a height of a meter and a half. The sections in which the measurements were made were 120m from the inlet and outlet portals as well as the middle on the tunnel. In each section a speed and pressure meters were set up.

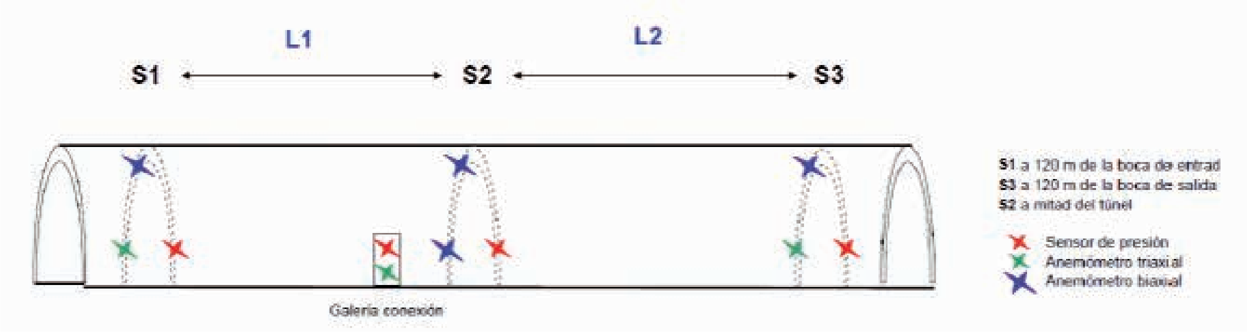


Figure 1 Control sections included in tunnels

- Sensors in the inspection train. The piezoelectric pressure sensors were placed on the roof of the train or installed on a metal plate in the train windows. They were placed in the middle of the driving head, in the central train car and in the other driving head, all of them outside.

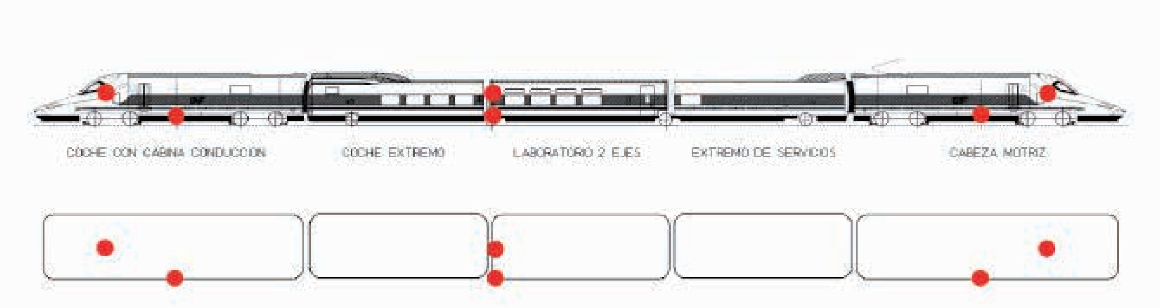


Figure 2 Control sections included in the inspection train

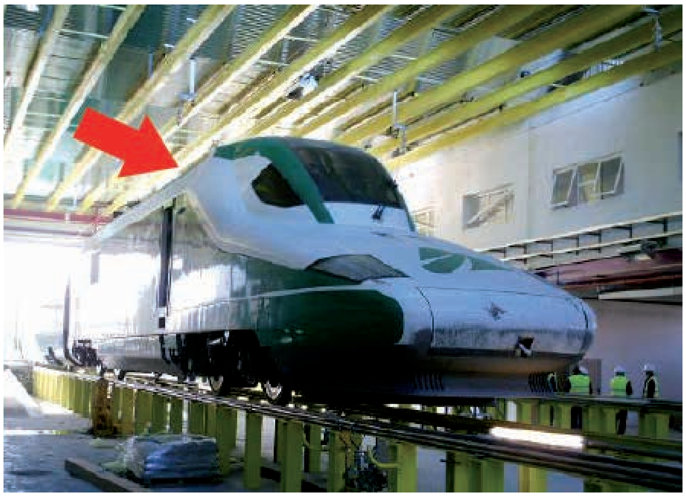


Figure 3. Inspection train from ADIF

The auscultation of the tunnels was carried out from late 2007 to early 2008. The obtained data presented a great similarity with real-scale tests data and simulations carried out with ThermoTun, which meant the use of Thermotun software in Spain as calculation software was authorized by ADIF.

The following is an example of one of the graphs that were obtained by comparison between the auscultation and the full scale tests. It belongs to instrumentation section S1 of the Castejón tunnel at a speed of 278km/h and the S -103.

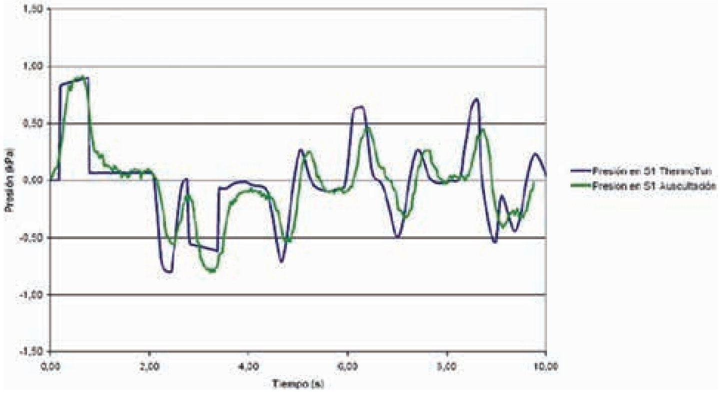


Figure 4. Comparison ThermoTun and full-scale test in Castejón Tunnel.

5. Aerodynamic calculations

Different aerodynamic studies have been carried out at INECO over the years looking for improving the circulation conditions of the infrastructure and the design of tunnels in general. The studies have been carried out working with ADIF and other private entities in a collaborative approach improving the circulation conditions of these infrastructures and the design of tunnels in general.

5.1 Tunnel cross-sections optimization

From 2008 until the end of 2012, a study was developed that enabled the analysis of almost 50 tunnels included in the High Speed Railway Network of Spain. The objectives are summarized in the following aspects:

- Analyse compliance with aerodynamic regulations in terms of health and comfort criteria. At that time the 2008 ETI was in force, which was used for dimensioning of the health criterion. The ADIF IGP-4.1 standard was used for the analysis of the comfort criterion.

- In the case of compliance with the initial data (type of cross- section, longitudinal profile and speed), a section optimization was carried out, resulting in section reductions that were implemented for the following project phases.

- If these criteria were not met, two alternatives were considered: a bigger cross-section was analysed or, if considered appropriate, a possible reduction of speed was contemplated assessing reductions of 5km/h at a time with the aim of not penalizing in excess the final exploitation of the line.

It should not be forgotten that the objective of the development of the High Speed Railway Network was and still is to obtain a High Speed Railway Network that allows acceptable travel times at a relatively inexpensive cost.

With all these premises, the following are the most important results obtained in the project:

- A large majority of the proposed tunnel cross-sections were optimized. There was only one tunnel in which it was not possible, (3rd Tunnel of the Pancorbo - Ameyugo section) in which the section had to be increased by 10m2.

- There were very specific tunnels in which it was only a question of checking compliance with both health and comfort criteria since the starting sections were the minimum stipulated by ADIF in the IGP-4.1. This suggests that the minimums established by ADIF in its regulations are optimised to the aerodynamic needs. These values were 85m2 for the case of double track section and 52m2 for the single track case.

- Given the great versatility of the calculation code used (ThermoTun), it was possible to model the tunnels in different situations. For example, in the case of the Temerosa tunnel, the actual conditions of the analysed tunnel was modelled. In this tunnel three specific narrower sections exist due to constructive needs and were analysed with detail.

It is noteworthy that with the UIC 779-11 Code it is not possible to reflect this situation of the tunnel, which means that the results obtained when using these leaflets are not completely accurate.

The works carried out solved situations during the execution of the tunnels, demonstrating a great adaptability of the calculations to real situations. In this way, it was possible to update the aerodynamic study to the existing conditions of the infrastructure and to introduce changes of speed during the stage of exploitation.

In view of the conclusions drawn in the study, the following points are highlighted:

- The optimization of the free cross-sections of the tunnels has allowed to achieve savings in construction costs derived from the reduction in the concepts of excavation, support system and lining.

- In cases where the section could not be reduced, either because it was the minimum section according to the IGF 4.1 of ADIF, or because it was not possible to reduce it due to other reasons, it was possible to carry out a speed sensitivity study, to determine the maximum speed of circulation of the trains inside the tunnels with the conditions imposed. This improves and optimises the operating conditions.

- ThermoTun's great versatility and flexibility against the UIC 779-11 tokens. This type of calculations allows to implement changes in the layout of tunnels, trains, speeds that the chips do not allow, so it is not possible to quickly adapt them to a construction situation in which changes can occur suddenly.

However, over the years, the following must be taken into account:

- During the last years there have been changes in the characteristics of the rolling stock. This means an improvement of the conditions of the tightness of the trains, as well as the geometrical and aerodynamic characteristics thereof, which make the results obtained from the aerodynamic calculations with the current trains better and better. This makes possible, from the point of view of the operation, to increase speeds while meeting the criteria of health and comfort. Therefore, it is recommended that, if you want to increase travelling speeds, these aerodynamic studies are performed.

- Currently there are many tunnels in Spain that were built without carrying out this type of studies. In many cases, conducting a study of these characteristics could provide with circulation guidelines for trains that improve operating conditions.

5.2 Interoperability certifications. High speed lines to levante and Barcelona

Another relevant piece of work carried out has been the certifications of interoperability, have been carried out for two reasons. On one hand, to demonstrate compliance with the Infrastructure TSI under the supervision of the NoBo (Notify Body). On the other hand, to assess how fast it is possible to circulate complying with such regulations.

With that purpose, studies have been carried out for the Levante High Speed line and the Barcelona - Figueras line.

In the case of the Levante line, S-130 and S-112 trains were considered as well as a generic interoperable train that complied with the Rolling Stock TSI. The calculation speed was 300km/h for a total of 22 tunnels.

In the study of the Barcelona line, the studies were carried out with two types of trains, a S-103 circulating at 350 km/h, and an S-112 circulating at 330 km/h.

In all cases, trains of 200 and 400m long were analyzed. For the case of double track tunnels the following train crossings were considered: 200-200m, 200-400m and 400-400m.

5.3 Construction stage. Aerodynamic studies

On many occasions, one of the problems that a contractor deals during the execution of a tunnel is the "entry into section", or, in other words, the occurrence of convergences and closure of the section. In these cases, it is considered appropriate to carry out an aerodynamic effect study in which introducing the new characteristics of the tunnel, including the location and length of the existing narrowing, demonstrates that aerodynamic health and comfort criteria are still met, or if these are not met, determine the speed value that makes these compliances possible

5.4 New explotaition speeds

Another benefit of aerodynamic studies is the analysis of the possibility of increasing the speed of circulation aiming to achieve improvements in the operating conditions. In these cases, the infrastructure is actually built and takes some time in service, so just the parameter of speed is studied in these cases

As it can be observed, aerodynamic studies are an increasingly important tool for the development of high-speed rail tunnels, since it is a typology of studies that can be carried out over the entire life of the tunnel. It has influence on parameters as diverse as size and type of the section and trains speed.

6. Sonic boom

6.1 Introduction

A major evolutionary step has been that the European high-speed networks are gradually increasing traffic speeds as signaling and security systems allow. In both Asia and Europe, highspeed tunnel networks with speeds in excess of 250km/h have developed considerably in recent decades. Most of these tunnels were designed as a double track monotube since they were more economical and because the tunnel lengths were lower than the current ones.

The construction of tunnels with one type or another of cross-section greatly conditions the aerodynamic effects (overpressures, depressions and sonic boom) that occur in these tunnels. However, there is a tendency for the design and construction of bitube tunnels.

Currently, due to the development that is taking place in the High Speed Railway, tunnels are being built increasingly longer, so there is a generalized tendency to design the tunnels as single track tubes. The main reasons are:

- They are considered safer due to the reduced probability of collision of trains, and due to the better evacuation and rescue conditions.

- Maintenance is performed in better conditions of safety as it is possible to use one tube while the other is closed to traffic.

The main disadvantages are the higher costs of construction and operation.

Switching from dual track sections to single track tunnel sections involves the following aerodynamic considerations.

- Due to the existence of a smaller cross-section in the case of single track tunnels, the pressure fluctuations are greater.

- However, interference with other trains is lower, which implies better compliance with the required comfort criteria.

- In relation to micro - pressure waves (sonic boom), the probabilities of occurrence increase due to a smaller section.

- On the other hand, the presence of slab track instead of a ballasted superstructure, which is very common nowadays given its lower maintenance cost, increases the probability of the sonic boom phenomenon.

Since 2005 - 2006 the existence of the sonic boom phenomenon in Europe has been observed when the lines are put in service or when the speed increases. Until then it had not been observed due to the preference of double-track tunnels versus single-track tunnels.

- In Germany in 2005, this phenomenon was observed on the Nuremberg - Ingolstadt line in the first start - up tests in several bitube tunnels of about 7.2 and 7.7km respectively.

- Here in Spain, in the High Speed line Madrid – Valencia it has been detected in the tunnel of La Cabrera, bitube and about 7km in length. Also in the LAV Madrid - Valladolid / Medina del Campo this effect has been detected in the bitube tunnel of San Pedro of about 8.7km in length.

- In Japan, in 1975, the presence of this phenomenon was detected in the tests of the Okayama - Hakata line of the San - Yo section of the Shinkansen. Until that time, in Japan there was not any legislation limiting the phenomenon.

6.2 Decription of the sonic boom phenomenon

As it was mentioned above, when a high-speed train enters a tunnel, there is an overpressure and depression in head and tail that propagate along the same at the speed of sound, reflecting each time they reach the portals of the tunnel.

At the outlet portal of the tunnel, most of the pressure wave is reflected as an expansion wave propagating in the opposite direction towards the inlet portal. However, part of this wave is emitted in all directions outside the tunnel in the form of impulse-type micro-pressure waves, which generate a detonating sound in the vicinity of the outlet of negative environmental effects.

Thus, the phenomenon of micro - pressure waves has three phases:

- Generation of the compression wave by the entrance of the train in the tunnel. The shape of the compression wave generated depends on different parameters such as: areas of the train and tunnel, geometry of the nose of the train and speed of the train among others.
- Propagation of the wave through the tunnel. The propagation of this wave occurs at the speed of sound, and is affected by the inner structure of the tunnel: tracks, platform ...
- Radiation of part of the micro - pressure wave to the outside (sonic boom). As the velocity of the train increases at the entrance to the tunnel, the amplitude of the pressure wave grows cubically

The effect called "Sonic Boom" is characterized by a sound similar to an explosion in some cases. It can be detected even at distances of 1km, and it can cause vibrations in doors and windows of nearby buildings. In addition, the strong sound generated causes a generally unacceptable noise pollution.

The maximum value of these micro - pressure waves is approximately proportional to the pressure gradient of the wave arriving at the exit portal (generated by the entrance of the train in the tunnel) and inversely proportional to the distance to which the train is located of said portal.

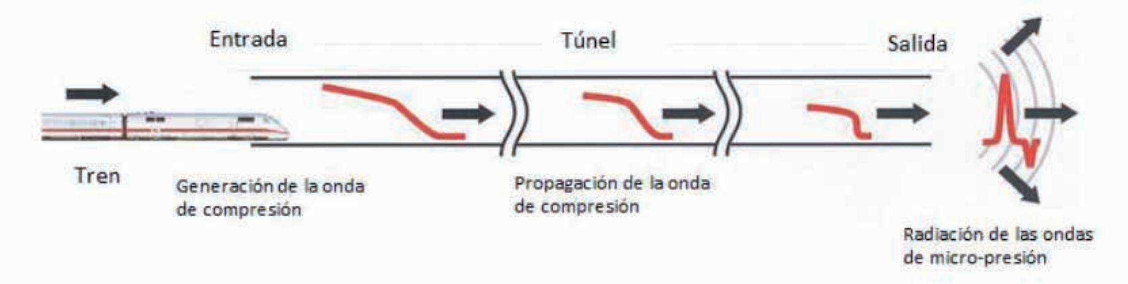


Figure 5 Sonic Boom generation

Currently, from different countries, and mainly at European level, measures are contemplated dissipating this effect once it has been proven to occur. However, there are still few actions that arise at the construction project level, since this effect can not be reliably calculated with the computer tools available today.

The most important measures used to mitigate this effect are detailed below. Some of them can only be included in the drafting phase of the project since these are measures that affect the construction of the tunnel. Other measures may be incorporated later, even if the tunnel is in operation.

As a main measure to mitigate the effect of the "Sonic Boom" is considered the reduction of the pressure gradient of the wave that reaches the exit portal of the tunnel once the train enters the same. This is achieved mainly by reducing the speed at the entrance of the tunnel and optimizing the aerodynamics of the nose of the train. These two measures can be implemented even if the tunnel is in operation.

On the other hand, structural measures can be applied that help reduce the effects of the sonic boom phenomenon. These measures could be:

- Increase the free cross- section of the tunnel which is only possible in the design phase.
- Modification of the design of the packaging: inclusion of "windows" or pores (also called porous or perforated packaging). These "pores" act as pressure sinks. Another measure is to modify the outer geometry of the portals making them flared, as well as to increase the length of the same.
- Use of ballasted track instead of slab track, since ballast, not being a continuous medium, has gaps between the edges that act as "pressure absorbers".
- Use of absorbent materials inside the lining (porous panels). This measure is very innovative and it has detractors because of fire safety and durability.
- Double track tunnel instead of two single track tubes with the aim of presenting a larger cross-section, only applicable in the project phase.
- Inclusion of connecting galleries between tubes, as long as the doors of the same are open and thus act the galleries as fireplaces. It must be taken into account that this aspect should

not go against the safety regulations of evacuation in case of emergency (propagation of smoke in case of fire).

- Inclusion of ventilation wells inside the tunnel near the portal area.

Figure 6. North Porous Portal. Perthus Tunnel.

Due to the great development of the Shinkansen over the years, a regulation has been considered that limits the value of the micro-pressures to 20Pa in the vicinity of tunnels.

Until 2013, this was the only existing regulation regarding the sonic boom effect and was also assumed by much of the world in the civil design of tunnels. In September 2013 a new German standard introduced as a module of the already existing DB Ril 853 was presented at the 15th International Symposium "Aerodynamics, Ventilation and Fire in Tunnels" which limits the value of the decibels produced by the sonic boom. This limitation has two variants:

- In the vicinity of buildings, the decibels in C-weighting (evaluation of high sound level sounds) are limited to the following values:
    - Residential areas: 70dB (C).
    - Parks and Gardens: 85dB (C).
    - Industrial Zones: 95dB (C).
- On the other hand, in the case of non-proximity to the previous elements, the decibels are limited to a distance of 25m from the portal to 115dB (C).

This regulation has been presented after carrying out a comparative study between actual measurements and a numerical model in the Katzenberg tunnel (Karlsruhe - Basel line) of about 9.4km in length, bitube and single track. This comparative study is currently being used in the analysis of the German high-speed rail network.

This limitation of C - weighted "noise" production is based on the fact that most of the energy produced by micro - pressure waves is at low frequencies (<100Hz) and even in the infrasound region (<20Hz). As it will be seen later, the frequency weights "clean" the sound pressure level (SPL), let us call it "raw noise", leaving only the noise produced in the desired frequency ranges so that it can be analyzed better. This cleaning is done by subtracting from the SPL noise a value

(WC) of "decibels" for each frequency. This Wc value is determined by a formula that depends on the frequency at which it is found (f).

Therefore, using the "noise" produced in C-weighting allows the clear analysis of noise produced only by micro-pressure waves.

6.3 Innovative project

In INECO, a tool called PROTAV has been developed, through the Innovation Directorate and thanks to the collaboration of the Ignacio Da Riba Institute of the UPM (Polytechnic University of Madrid) in Madrid, able to predict and attenuate the effect of the sonic boom.

To carry out this project the following steps were developed:

- Instrumentation of a tunnel in which sonic boom was produced
- Modelling of this tunnel and checking in the tool.
- Check using the tool if this effect occurred in tunnels in which the effect of micropressures does not occur.

The following parameters were measured:

- Distribution of pressure and air velocities in the tunnel.
- Static sound produced by the impact of the shock wave generated at the tunnel entrance when Mach 1 speed is reached.

The main objective of these measurements has been the evaluation of the Sonic Boom that produces the pressure wave generated by the trains at their entrance to the tunnel when it reaches the exit portal.

6.4 Cabrera tunnel features

The La Cabrera tunnel belongs to the LAV Madrid - Valencia / Albacete. It is a single-track twintunnel. Its most important characteristics are the following:

- Length: 7229m
- Cross-Section Type: Circular - Bitube of 53m2
- Input Line ch: 351 + 277 at 624m height
- Output Line ch: 358 + 506 at the 461m height
- As for the longitudinal profile according to increasing chainage is characterized by a slope of -30 ‰ during the initial 2.3km, then changes its slope to -8 ‰ for 1.1km, to finish with a slope of -25 ‰ rest of the tunnel.

Figure 7. North Portal of La Cabrera Tunnel.

Figure 8. Inner of North Portal of La Cabrera Tunnel.

6.5 Location of control sections for auscultation

The instrumentation arranged in the tunnel of La Cabrera has been very varied and with very different objectives.

Due to the intention to evaluate the effect of the sonic boom, two sonometers were installed in the vicinity of the portals to record the decibels corresponding to each generated pressure wave, in order to obtain the frequency spectrum of the burst in bands of thirds of octave. The La Cabrera tunnel presents ideal conditions to measure this effect since, as mentioned, it has been verified with a field visit that this effect occurs. Sound level sonometers have been placed 50m from the portal of the Valencia side.

As it has been necessary to measure pressure variations and wind speed inside the tunnel, pressure sensors and triaxial anemometers have been installed.

At the entrance and exit of the tunnel displacement sensors or strain gauges have been installed to record the passage of trains and to determine the speed, composition and type of train that circulates.

Likewise, the pressure, air velocity in the longitudinal component, temperature, humidity and environmental pressure, as well as the passage of the train have been recorded.

The central section of the train has been instrumented with pressure sensors, anemometers and train passage detectors. This section is located towards the middle of the tunnel and has sought the proximity to a connecting gallery in the tunnel of La Cabrera, in order to instrument in and out of the gallery with pressure sensors.

Figure 9. Instrumentation set in La Cabrera Tunnel

To summarize, the scheme of installation of the sensors in the tunnel of La Cabrera was the following one:

- A sensor to measure pressure, temperature and humidity outside the tunnel at ch 351 + 277 (tunnel entrance) and 358 + 450 (section 3).
- Two sound level meters on the outside of the outlet with the corresponding data acquisition equipment and separated by one meter and placed at 50m from the portal (section 4).

- Pressure sensors in ch 354 + 825 (in the middle of the tunnel, section 1) arranged in a tunnel and in the vicinity of the nearest gallery (ch 354 + 885, section 2), as well as inside the tunnel and as far away as possible from the entrance door to the tunnel. These sensors are accompanied by their corresponding data acquisition equipment and also records the temperature and humidity of the air.
- Triaxial anemometers inside the tunnel at ch 354 + 825 (section 1) and section 3 (ch 358 + 450) at the point where the plate is located with strain gauges. This is accompanied by a corresponding data acquisition team.
- Train pass detection sensors have been added inside the tunnel in the central section ch 354 + 825 (instrumentation section 1).

To summarize the instrumentation by sections arranged is as follows:

- Section 1, located at ch 354 + 825: tunnel pressure sensors, triaxial anemometers and strain gauges.
- Section 2, at ch 354 + 885: pressure sensors in a gallery near the central section. Arranged on the outside and inside of it.
- Section 3, in ch 358 + 450, has pressure sensors and triaxial anemometers.
- Section 4, at ch 358 + 568 (outside of the tunnel), in this section two sonometers separate from each other 1m and at a distance of 50m from the outlet portal are arranged.

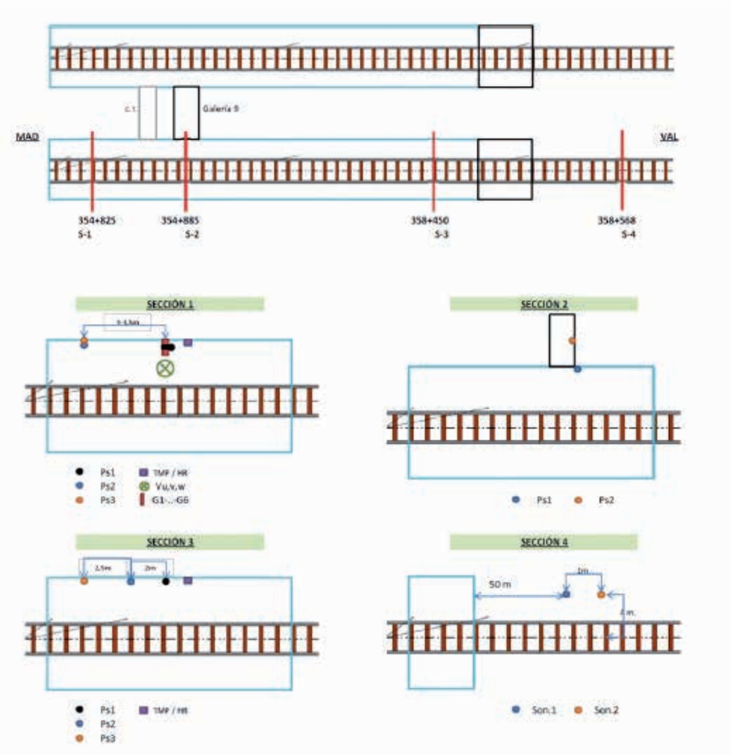


Figure 10. Scheme of the distribution of control sections in tunnel

The detail of the plate on which the gauges are located can be seen in the following diagram. It is possible also to see the position of the pressure sensor on one side of the plate in the next photo:

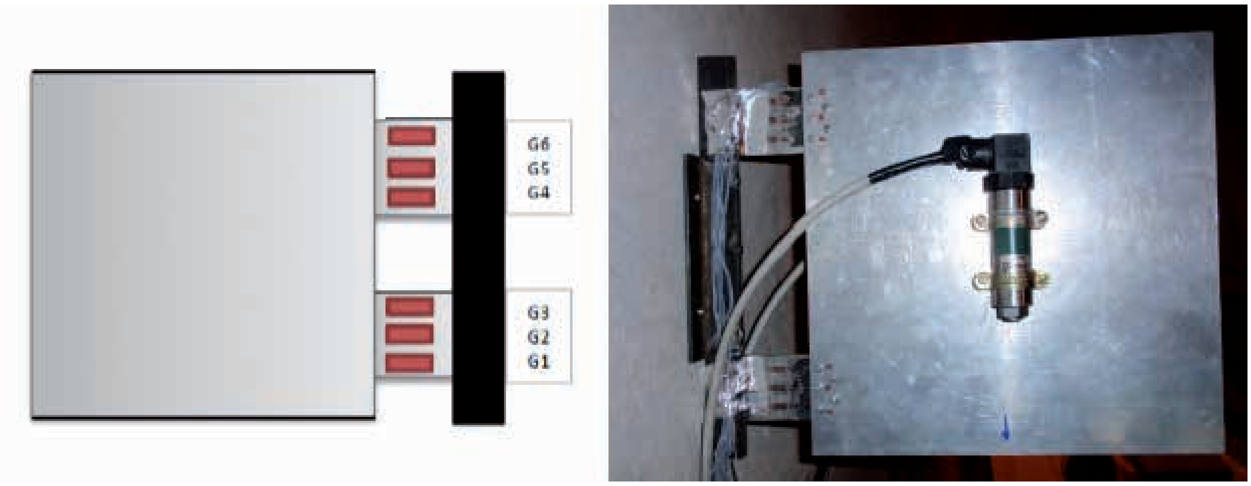


Figure 11. Pressure sensor location

For the installation of the sensors it has been necessary the use of a vehicle in the tunnel of The Cabrera and the existence of energy that feed the equipment to which the instrumentation has been connected.

For the installation of the sensors it has been necessary the use of a vehicle in the The Cabrera tunnel and to provide electricity to feed the equipment to which the instrumentation has been connected.

The instrumentation and start-up of the sensors was performed during the night and the measurements were carried out during the day. Three nights were used for the placement of the instrumentation and two nights to uninstall equipment.

The data of the circulations of these days: schedules, train typology, length, speed ..., were provided by RENFE to be able to compare with the arranged instrumentation.

Once the measurements were made in July 2013, work began on the processing of the same, which materialized in a series of Excel files for each of the circulations analyzed.

Then, a comparative work was started between the field data and the data obtained through the tool finalizing the innovation project with the generation of the PROTAV tool.

7. Protav

PROTAV (PRopagación de Onda en Túnel de Alta Velocidad,) tool has been developed in collaboration with the Institute of Microgravity "Ignacio da Riva", hereinafter IDR of the Universidad Politécnica de Madrid.

PROTAV has been developed as a tool that reflects more or less faithfully the behavior of pressure waves in tunnels, as well as the introduction of dissipating elements of overpressures. The main features of the PROTAV tool are as follows:

- It is an INECO´s proprietary software developed in-house. It allows the analysis and engineering calculations that were not approached with confidence. (Functionality of the sonic boom).

- PROTAV performs an analysis of the generation and propagation of pressure waves inside a tunnel with a higher degree of accuracy than other existing tools and also at a very low computational cost.
- In order to ensure greater accuracy of the studied phenomenon, PROTAV allows to define a multitude of parameters identified as relevant in the calculation of the aerodynamic phenomenon: velocity, length, perimeter and train area, tightness, tunnel geometry, existence of pressure dissipating elements ...
- PROTAV is a tool with a relatively simple and intuitive operation.

In long tunnels with low friction (slab track), the nonlinear effects of wave propagation produce a steepness of the pressure wave front and can form a shock wave. Increasing the pressure gradient of the wavefront increases the wave's ability to radiate energy out of the tunnel by impacting the output.

The method selected by IDR and proposed for the study of radiation is based on the wave separation process proposed by Kikuchi in 2009. This method consists of placing two microphones within the tunnel, whose relative position must be accurately known and placed at some distance from the exit of the tunnel. Another microphone is placed outside the tunnel at such a distance that the flat wave hypothesis can be assured to capture the radiated wave and calculate the transfer function of the tunnel exit.

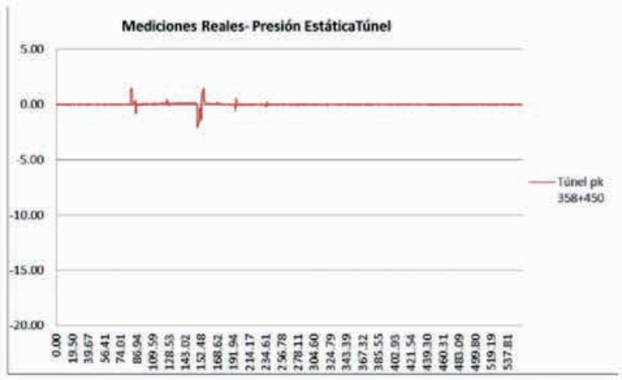
7.1 Protav validation

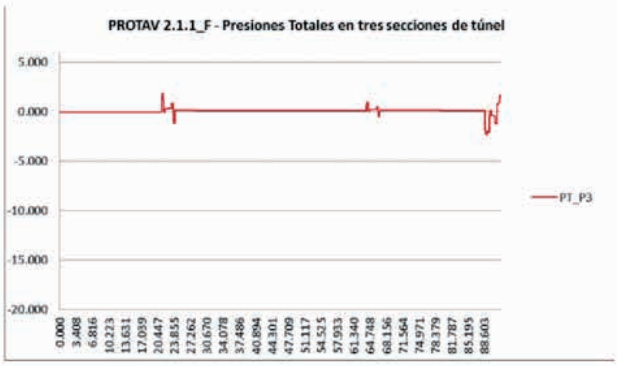
As can be seen in previous sections, four sections were instrumented, three inside the tunnel, and the fourth on the outside with the sonometers.

Sections 1 have been considered as comparison sections located at ch 354 + 825 and section 3 located at ch 358 + 450 direction Valencia. The comparison pass is the T-02, an S-112 train that is detected in the tunnel at 8:44 h at 291 km/h and is detected at the exit section at 8:45 h at 296 km/h. This train has been taken as a comparison train, since in this case the sonometers of section 4 detected a sonic boom.

The comparative graphs of these measurements with PROTAV are shown below. In the graphs:

- For PROTAV, the abscissa represents the time in seconds (s), and the ordinate axis represents pressures in kilopascals (kPa).
- In the Full scale measure graphs, the abscissa axis represents the available data number and the ordinate axis represents pressures in kilopascals (kPa).





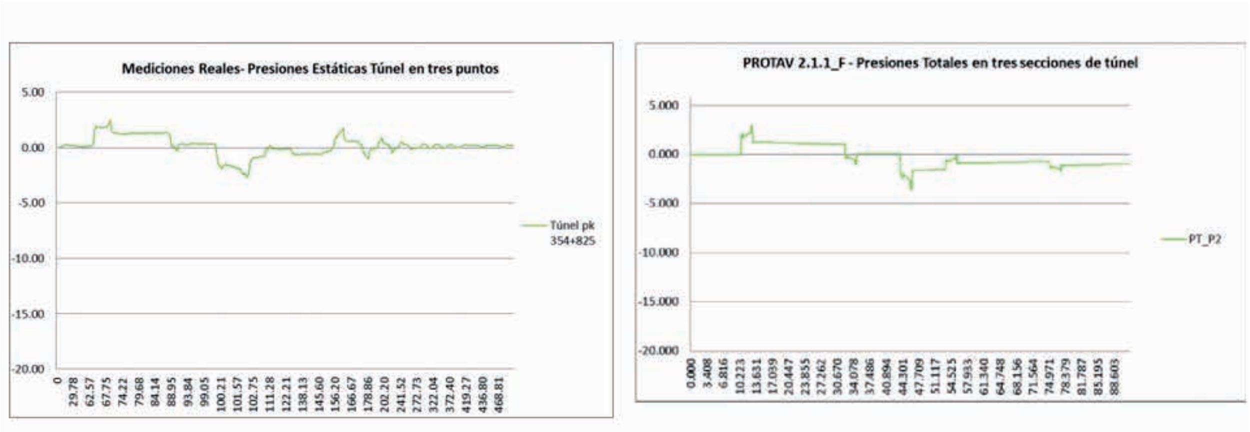


Figure 12. Comparison full-scale test vs PROTAV

As can be seen in the graphs, two aspects to comment are observed:

1. PROTAV does not consider atmospheric pressure values when it measures instantaneous pressures. This results in the results obtained starting at 0 kPa as is done in the case of measuring devices.
2. There is a certain difference between the graphs. This is due to the volume of data. To make the graphs in the case of the actual measurements have been counted with about 550 data, which are provided by the pressure sensors. To realize the graphs of PROTAV have obtained something more than 8800 data which explains the small existing distortions.

There is a certain difference between the graphs. This is due to the volume of data. To make the graphs 550 readings provided by the pressure sensors have been considered. To produce the PROTAV graphs more than 8800 readings have been used, which explains the small existing distortions.

7.2 Sonic boom calculations

To determine the validity of the sonic boom several of the passes of the real measurements of July 2013 have been compared against the simulation of the same conditions in PROTAV in order to evaluate this new functionality of the tool. One of them is shown as illustration.

The data provided by the SON 1 sonometer of instrumentation section 4 include a maximum of 122.6dB Overall at 100Hz. The SON 1 sound level meter is at a distance of 50m from the exit portal to Valencia. This is the graph of the sound level meter that illustrates this data:

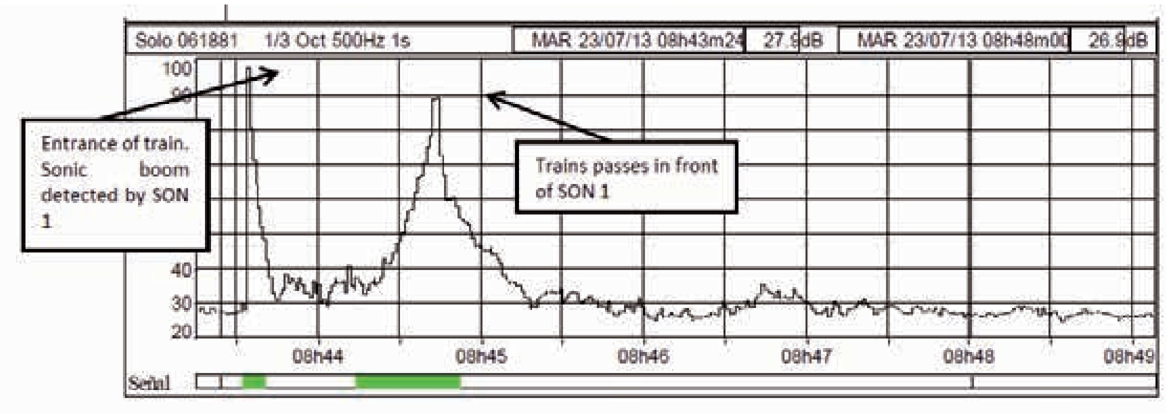


Figure 13. Relation between frequencies and time of the train T-02

As shown in the above chart, SON 1 located in the opposite portal of the entrance detects a very steep rise after 8:43 p.m., time of entrance of the train in the tunnel, so that this peak reflects the sonic boom. Very close to 8:45 a second peak is detected again indicating that the train is passing in front of the sound level meter 1 once it has left the tunnel. Both rises occur in the frequency range of 100Hz or 90Hz (low frequencies) as seen in the chart above.

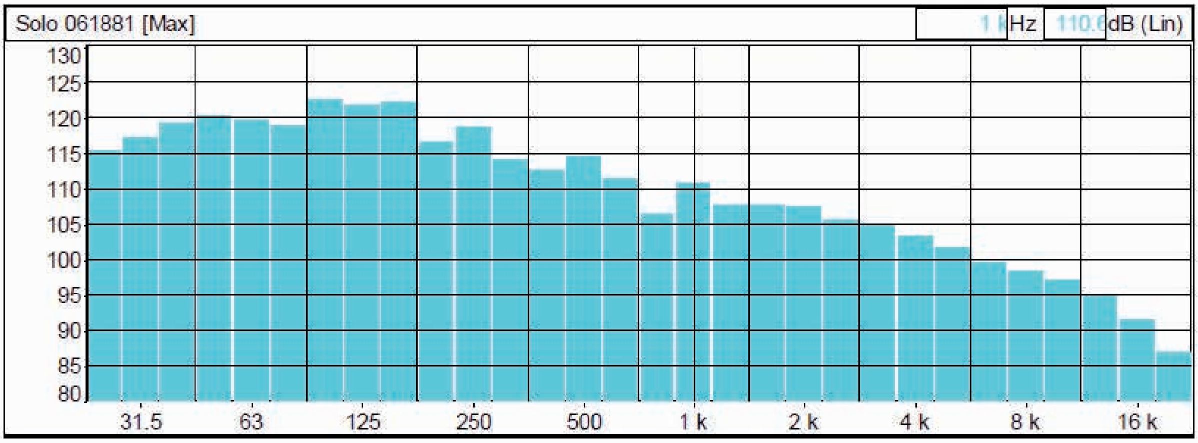
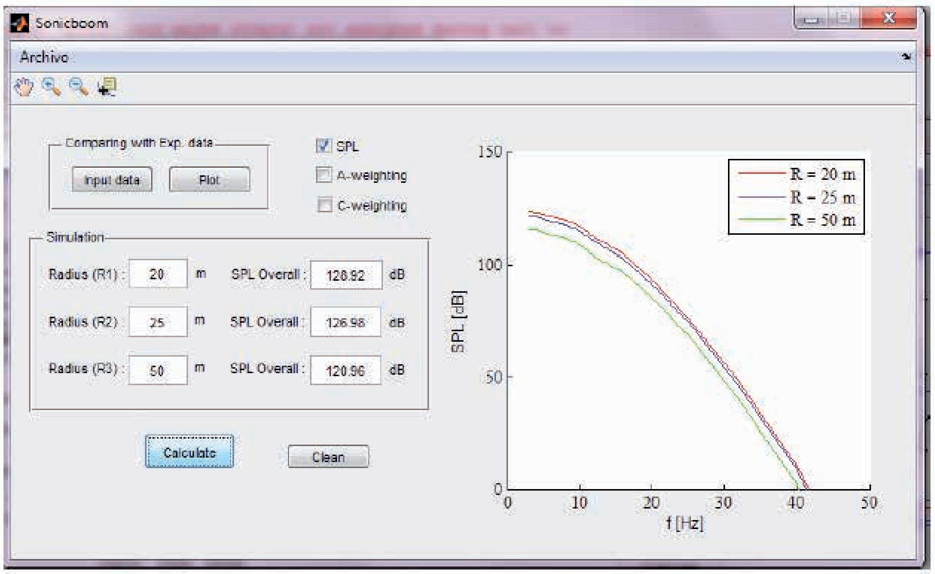


Figure 14. Relation between frequencies and decibels overall of the train T-02

The above image shows the Overall decibels produced with the frequencies at which the sound is produced. In this graph is possible to see that at approximately 100Hz there are more than 120dB, specifically 122.6dB Overall.

With this data, the simulation in PROTAV of the same conditions of this pass was carried out, and it was measured the Overall decibels produced at 50m from the portal (position of SON 1):



☑ SPL
☐ A-weighting
☐ C-weighting

Radius (R1) : 20 m SPL Overall : 128.92 dB
Radius (R2) : 25 m SPL Overall : 126.98 dB
Radius (R3) : 50 m SPL Overall : 120.96 dB

Distance R [m]	20	25	50
SPL Overal [dB]	128.9154544	126.9772541	120.966578

Figure 15. Results obtained with PROTAV in the T-02 train

As shown in the table above, at 50m distance PROTAV detects 120.966dB. This value is less than detected even though it is in the same order magnitude and is only 1.33% lower. Therefore, and in view of the results obtained after several analyzes, it can be concluded the great reliability of the tool in the calculation of this phenomenon

8. Conclusions

Throughout this article a revision of the possibilities of aerodynamic studies in high speed rail tunnels has been performed, including the opportunities offered to the designer and the possibilities of optimization of the free cross- section, as well as the possibility of analysis of variations of speed to improve the operating regimes.

Additionally, it has been presented the PROTAV tool of sonic boom calculation done by INECO in collaboration with the Instituto Ignacio Da Riba (IDR) of the Polytechnic University of Madrid. This new tool represents a new possibility of design and calculation of a growing phenomenon caused by the gradual train speed increments.

9. References

- COMMISSION REGULATION (EU) No 1299/2014 of 18 November 2014 on the technical specifications for interoperability relating to the ‘infrastructure’ subsystem of the rail system in the European Union.
- NAP 2-3-1.0 de Túneles de ADIF, Julio 2015.
- UIC Code 779-11. Determination of railway tunnel cross-sectional areas on the basis of aerodynamic considerations, February 2005.
- DB Netz AG: Infrastrukturseitige Behandlung der Mikrodruckwellen – Thematik, 853.1002ª01. In Richtlinie 853 Eisenbahntunnel planen, bauen und instand halten, 02.2013.
- Micro-pressure wave countermeasures realized in the Katzenberg tunnel and introduction of a new German micro-pressure wave regulation. P. Deeg, M. Hieke, C. Gerbig, T. Tielkes. Proceedings of the 15th International Symposium on Aerodynamics, Ventilation and Fire (2013).
- ThermoTun 7.1 User Manual. Alan E. Vardy. Dundee Tunnel Research. January 2014.
- Kikuchi, K. Ozawa, S. Takami, H. Iida, M. Continuous pressure waves generated by a Train Running in a Tunnel. Journal of low frequency noise, vibration and active control 2009.

High Speed Rail Infrastructure

360.revista de alta velocidad

número 5 - junio - 2018. Pág 355 - 365

A railway culvert maintenance management approach based on risk assessment techniques

Campos Zaldiernas, Javier

Díaz García, Sarai

Galán Alguacil, Álvaro

González Pérez, Javier

University of Castilla-La Mancha[1]

Abstract

Culverts are essential elements of any railway infrastructure. Although there are commonly a large number of these elements in any transport infrastructure, its proper maintenance is often treated as a secondary issue when it comes to the management of ordinary maintenance investments. The aim of this work is to present a methodology intended to create strategies for prioritizing investments in culvert maintenance works. Applying risk assessment techniques, a preliminary diagnosis of the current culvert condition can be obtained, evaluating failure probabilities and quantifying socioeconomic impacts. At the same time, common pathologies and general risk factors are identified in order to determine the optimal solutions to the culvert current issues. Assessment of all culverts in a transport network allows to efficiently managing the investments in ordinary maintenance, optimizing the expenses and minimizing risks.

Keywords: culverts, infrastructure maintenance, risk assessment, investments optimization

---

[1] *Campos Zaldiernas, Javier. University of Castilla-La Mancha. Email: javier.campos@hidralab.com*
Díaz García, Sarai. University of Castilla-La Mancha. Email: Sarai.Diaz@uclm.es
Galán Alguacil, Álvaro. University of Castilla-La Mancha. Email: Alvaro.Galan@uclm.es
González Pérez, Javier. University of Castilla-La Mancha. Email: Javier.Gonzalez@uclm.es

1. Introduction

Culverts are essential elements of any railway infrastructure, however this kind of assets tend to be ignored until a catastrophic failure occurs. Culvert failure may range from structural collapse to insufficient capacity to pass floods and it might affect not only the infrastructure itself but also it could cause serious problems on the railroad and its surroundings, such as damage to the nearest properties, floods and major traffic disruptions.

Although the design of railway culverts is currently subject to national technical regulation: (Ministerio de Fomento, 2015), (Ministerio de Fomento, 2014); maintenance management aspects and evolution of service levels through time are not considered. Experience in other countries (USA, Australia) shows that having up-to-date and detailed information about culvert conditions is crucial when it comes to maintenance management (Federal Highway Administration, 2010) and planning of preventive maintenance activities (Balkham et al., 2010). As many of these structures reach the limit of their service life, management administrations must need to schedule their replacement or repair (Najafi and Bhattachar, 2010), however, currently they do not have and adequate prioritization methodology that help them planning this operations though time, so these operations are often undertaken once the incidence in the service level has occurred (corrective actions).

Figure 1. Traffic disruption events caused by culvert failure: overtopping flows and infrastructure collapse.

The aim of this work is so to develop an objective methodology based on risk assessment techniques for railway and road culverts in order to build investments prioritization models for those infrastructures. The methodology is based on the combination of probabilistic models related to the different culvert failure mechanisms and the repercussion or consequences that the failure event on the global level of service of the infrastructure. Finally, a global assessed risk level derived from the general likelihood and consequence analysis is obtained (Roads and Traffic Authority, 2010). Failure likelihood assessment is based on the analysis of failure modes, or situations where the service levels in the infrastructure are totally cut off due to culvert malfunction, including slope instability, local scour and platform overtopping flows events. Likelihood analysis of failure mechanisms is based on overall material deterioration evolution models (Salem et al., 2012), hydrologic and hydraulic characterization and preliminary analysis of embankment stability and local scour in inlets and outlets (Galán et al., 2016). On the other hand, consequence analysis is not only based on economic costs quantification (Perrin and Jhaveri, 2003), but also in the repercussion assessment for whole society.

Since the main objective of this work is to globally apply the proposed methodology to the railway transport network, being able to massively address the preliminary risk assessment of a large set of assets is a key issue of the process. Some specific applications have been developed in order to overcome this issue. Maintenance investments prioritization will be determined by the profitability of the proposed actions in every analysed culvert based on the nature and entity of the observed pathologies. Thus, it is not only possible to prioritize the kind of proceedings required at a specific time following economic criteria, but also analyse and compare the profitability of a set of preventive actions to be carried out in a certain section of the infrastructure or in the whole transport network.

The present document presents the risk assessment methodology proposed to obtain the assets preliminary diagnosis and its suitability to perform global analysis on railway and road infrastructures.

2. Methodology

The proposed methodology is based on the quantification of the asset risk level. The circumstances that constitute failure of hydraulic structures, including culverts, can be defined differently depending on the viewpoint of the analyst and purpose for the analysis. For the purpose of this analysis, failure events are defined as those situations leading to partial or total railway traffic disruption. Failure may range from structural collapse of any of the elements constituting the culvert (culvert barrel, headwalls, wingwalls and auxiliary structures), to insufficient hydraulic capacity to pass floods. The asset risk level is determined as the combination of the probability of occurrence of each of the events leading to failure (failure mechanisms) and the impact or consequences derived from this situation.

The asset risk level assessment requires a complete characterization of the culvert and its surroundings, thus it is crucial to develop a regular inspection program based on field analysis, including a detailed evaluation of the following items:

- Culvert barrels: including geometry and general condition assessment (related to structural damage).
- Inlet and outlet structures: including general configuration, geometry and condition assessment.
- Embankments: geometry characterization, general soil characterization and current condition (presence of evident signs of slope instability caused by influx or piping) based on visual assessment.
- Waterway: characterization of the waterway downstream the culvert outlet, including shape and approximate geometry of the cross section.
- Adjacent areas: general determination of location and main characteristics of the nearest elements to the culvert inlet or outlet that could be affected by an occasional instability of the railway embankment or a structural collapse of the headwall.

The field assessment must be complemented with further desktop analysis in order to determine all the required parameters to accurately evaluate failure probabilities and potential impacts of the culvert failure. The information required in this case range from average data of railway traffic intensity to the characterization of the waterway drainage basin. The aim of this work is to provide a preliminary diagnosis of the current performance of all culverts inspected, so that it is possible to prioritize the care level of each asset according to an objective parameter: the asset risk level, quantified by the yearly average economic impact of the culvert failure.

Once the assessed risk level has been determined, the preventive actions to be carried out in each case can be prioritized according to risk reduction criteria. According to the preliminary diagnosis, the key pathologies or risk factors that affect the culvert are identified and based on this information, a series of standard actions are proposed to address the deficiencies found. Each action has a specific cost and an economic repercussion on the assessed risk level defined as the reduction of the annual risk of failure caused by such remedial works. Based on these parameters, it is possible to carry out a cost-benefit study of each of the considered actions and lastly, selecting the actions in each case according to profitability criteria in the short, medium and long term.

The repairs or remedial works selected must fit within the budget availability, so that the most cost-effective actions are carried out in the first place

2.1 Failure probability

The annual failure probability of a culvert is determined as the sum of the failure probabilities of each of the failure mechanisms operating in the influence area of the culvert, causing partial or total traffic disruption in the railway. These events are grouped into four main categories called failure modes (Roads and Traffic Authority, 2010):

1. Failure mode 1: Structural collapse. There is only one potential failure mechanism:
    - 1.1 Structural collapse of the culvert barrel.
2. Failure mode 2: Slope instability. There are four potential failure mechanisms:
    - 2.1. Slope instability caused by afflux.
    - 2.2. Slope instability caused by leakage out of barrel.
    - 2.3. Slope instability caused by headwall collapse.
    - 2.4. Slope instability caused by undermining at inlet or outlet.
3. Failure mode 3: Piping. There are three potential failure mechanisms:
    - 3.1. Piping into culvert.
    - 3.2. Piping on outside of culvert due to afflux
    - 3.3. Piping on outside of culvert due to leakage out of culvert.
4. Failure mode 4: Hydraulic flow. There are two potential failure mechanisms:
    - 4.1. Erosion by overtopping flows.
    - 4.2. Cross catchment flooding.

Thus the previous mechanisms are considered independent phenomena and therefore the annual probability of failure of the culvert (global) is assessed as the sum of the probabilities of each failure mechanism. Each mechanism is constituted by a series of events that must occur sequentially so that the failure mechanism is able to take place. These events are divided into two different categories according to its nature:

- Precedent events: these events caused the failure situation, corresponding to an initial situation linked in most cases to a rainfall event or the current condition of the culvert barrel, headwall and wingwalls (deterioration). Adverse circumstances are essentially weather-related precedent events – either exceptional rainfall or extreme groundwater levels. In such cases the probability of failure may be governed by the return period of the necessary precedent event, which would be reflected in the assessment of an annual probability of failure.

- Triggering event: on the other hand, once the precedent event has taken place, the triggering event reveals the likelihood of the infrastructure to be affected by the failure mechanisms. In this case the annual probability of failure may depend on the culvert component geometry and configuration. Depending on parameters such as the height of fill, the soil type and embankment configuration, the global failure probability may vary according to the ability of the current configuration to potentiate or mitigate the adverse effects of the triggering events.

A series of specific studies have to be carried out in order to assess the probability of certain events previously mentioned. Some of which are:

- Hydrologic characterization of the drainage basin: a complete preliminary characterization of the watershed must be undertaken. Hydrologic semi distributed models are applied in order to proceed to the upstream peak flows calculation.
- Hydraulic characterization of the culvert: according to the culvert current configuration, geometry and waterway characteristics, the hydraulic performance of the culvert in the peak flow situations can be obtained.
- Embankment slope instability caused by afflux preliminary studies: considering the hydraulic performance of the culvert previously obtained.
- Embankment slope instability caused by undermining at inlet or outlet preliminary assessment based on experimental studies (Galan et al., 2016).

The integration of these specific studies on the global assessment of the failure probability of each failure mechanism is crucial to identify pathologies linked to potential deficiencies of the hydraulic capacity of the culvert. Additionally, each failure mechanism can be easily decomposed into a series of risk factors, defined as the main pathologies that can lead to the culvert failure. According to its nature, the risk factors are divided into five main categories:

- Partial or total blockage of the culvert barrel.
- Structural deterioration of the culvert barrel.
- Structural deterioration of the culvert headwall or/and wingwalls.
- Evidences of substantial erosion at the culvert outlet or inlet (undermining).
- Deficiencies found in the hydraulic capacity of the culvert.

Figure 2. Risk factors examples: structural deterioration of the culvert barrel, partial blockage of the culvert barrel, evidences of substantial erosion at the culvert outlet.

Identification of the influence level of each risk factor is crucial in order to determine the optimal remedial works to be undertaken in each asset.

Failure consequences

The main purpose of the consequence analysis is to identify the effects of the hazards on the element at risk. The consequences of the culvert failure can be described as the socioeconomic impact produced by subsequent traffic disruption. Those consequences are divided into two main categories:

- Consequence for risk to life: qualitative analysis of the risk to life assessed as the combination of the specific vulnerability (track speed limit, height of fill, security level) and the temporal probability (linked with the average railway traffic intensity).
- Consequence for risk to property and socioeconomic impacts: consequence with respect to property damage and other consequential effects (traffic disruption) of the failure are to be assessed by a quantitative analysis of the economic repercussion caused by each specific failure mechanism. The type of damages and adverse effects are different depending on the nature of the mechanism. In general terms, these consequences can be divided into two main categories:
    - Direct economic costs: there are some consequences that directly affect the railway administration such as rehabilitation and repair costs.
    - Indirect economic costs: on the other hand, some impacts are not directly assumed by the competent administration but they affect society as a whole. These costs include:
        - Traffic disruption costs: linked with the increase in travel times derived from circulation by alternatives routes of the rail network (Perrin and Jhaveri, 2003).
        - Compensation costs derived from landslide impact on near properties due to embankment slope instability: in this case the economic impact depend on many factors such as: height of fill, type of property affected, distance from the element to the failure influence area and general nature of adjacent development.
        - Compensation costs derived from flooding upstream the culvert: in many cases the flooding is caused by the insufficient hydraulic capacity of the culvert so the compensation must be held by an insurance company or the proper administration. Upstream nearest crops and constructions may be affected as long as the flood remains.
        - Traffic disruption costs in upstream adjacent transport infrastructures (other railways or roadways) caused by floods.

Global socioeconomic impact will be assessed for each failure mechanism considering that the consequences of failure may depend on the nature of the failure itself. Where there are multiple consequences, the total should be considered, as it may increase the consequence class above that derived from the individual effects. Note that the consequence classes used for loss of life (combined vulnerability and temporal probability) are not considered equivalent in economic terms to those for the damage to property and consequential effects, reflecting the lower tolerance which exist in society for loss of life compared to pure economic losses.

Culvert failure risk analysis

Once the annual probability of failure linked to each of the ten mechanisms studied is obtained and knowing the level of economic and social impact associated with each culvert, it is possible to determine the preliminary risk level of the asset. The risk level associated with a particular culvert is the direct result of the integration of the annual failure probabilities and the eventual economic and social impact linked to each failure mechanism. The nature of these mechanisms leads to consider that these events are independent so the total risk will be the sum of all the risks evaluated for the same asset.

3. Preliminary results and potential implementation

This section shows the main conclusions drawn from the application of the risk assessment methodology to a set of 757 culverts located in different roadways the province of Ciudad Real, Spain. The set includes different types of culvert barrels, different materials and deterioration conditions. Additionally, the sample covers a considerably wide geographical area so that from the point of view of the topography and the characteristics of the watersheds it is considered sufficiently representative.

An important aspect to consider is that although the results shown below correspond roadways, it has been considered that these results show the potential of the methodology to be applied to different kind of linear infrastructures and its versatility since the changes to be made in the assessment of probabilities and consequences in order to consider the level of risk in railway culverts is minimal.

This analysis has been held in different communication routes (motorways, national highways and autonomic roads) with different traffic intensities in order to fully cover the spectrum of potential socio-economic impact. A number of peculiarities have been recognized to be taken into account during the diagnostic process listed below:

- Evaluation of the hydraulic capacity of groups of culverts connected in series
- Evaluation of the hydraulic capacity of groups of culverts connected in parallel.
- Evaluation and particular recognition of culverts in extensions of roadway, recognizing different typologies and materials in the same culvert barrel in several of them.

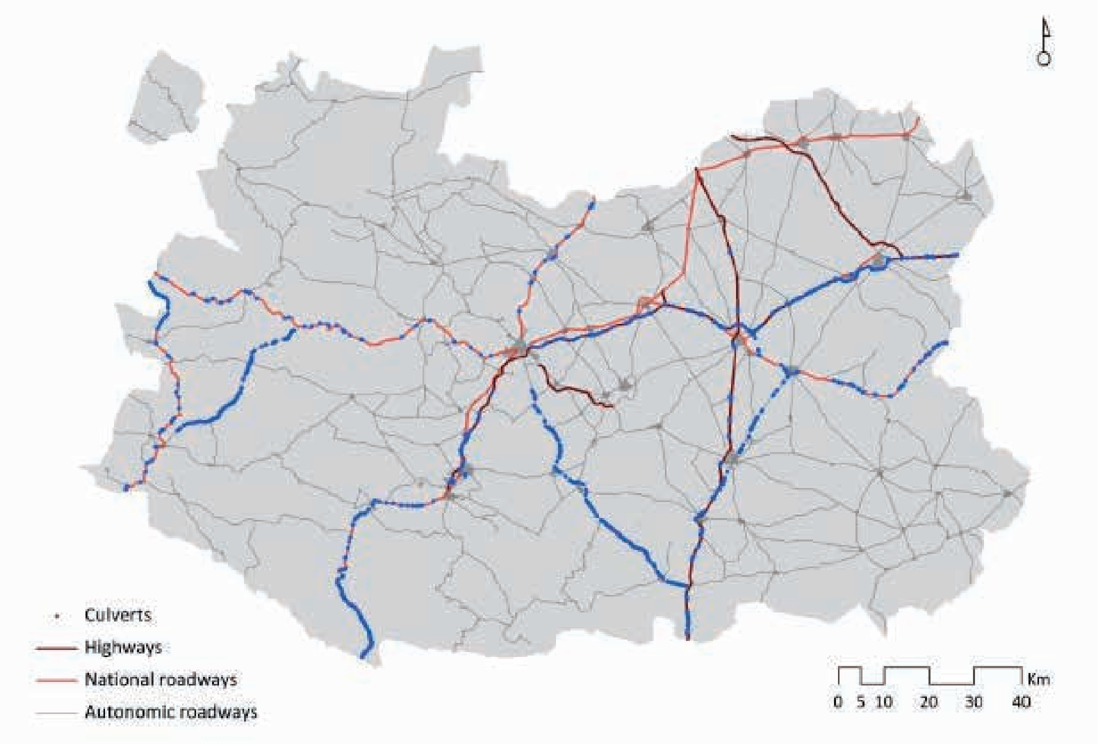


Figure 3. Location of the inspected culverts, Ciudad Real (Spain)

Failure mechanisms	Annual probability of failure	Economic impact	Economic risk	
	(-)	(€)	(€/year)	(%)
Structural collapse of the culvert barrel	0.00842	136829	1151.5	49.3
Slope instability caused by afflux	0.00138	57312	78.8	3.4
Slope instability caused by leakage out of barrel	0.00013	7841 0	10.1	0.4
Slope instability caused by headwall collapse	0.00220	57308	125.9	5.4
Slope instability caused by undermining at inlet our outlet	0.00208	56788	118.3	5.1
Piping into culvert	0.00264	117859	310.6	13.3
Piping on outside of culvert due to afflux	0.00004	112439	5.0	0.2
Piping on outside of culvert due to leakage out of culvert	0.00171	117859	201.6	8.6
Erosion by overtopping flows	0.00472	70461	332.4	14.2
Cross catchment flooding	0.00032	524	0.2	0.1

Table 1. Main results obtained: annual failure probability, economic impacts and risk.

Incidence Risks factor/pathologies	INCIDENCE
	(%)
Partial or total blockage of the culvert barrel	34.50
Structural deterioration of the culvert barrel	7.83
Structural deterioration of the culvert headwall or/and wingwalls	14.34
Evidences of substantial erosion at the culvert outlet or inlet (undermining)	29.60
Deficiencies found in the hydraulic capacity of the culvert	13.73

Table 2. Main results obtained: risk factors and pathologies observed.

Tables 1 and 2 show the statistics of the results obtained in the risk analysis performed in some roadways in the province of Ciudad Real. Table 1 shows the mean values of probability, total economic impact and annual risk of the ten failure mechanisms studied.

The total economic impact is divided into the direct economic impact associated with the repair / rehabilitation costs directly attributable to the competent administration (representing on average 69% of the total economic impact) and the associated indirect economic impact to the costs derived from the traffic disruption on the roadway and various compensation costs, which are held by society (31% of the total). On the other hand, for the purposes of the present study only the indicators of impact and economic risk are shown since the impact and personal risk on the users of the road is assessed only in qualitative terms.

Notice that probabilities and consequences values must be converted to qualitative scales so that the qualitative assessed risk level can be obtained. A scale of 5 risk categories defined by annual probability of failure and economic impact ranges (R1 to R5, where R1 is the highest level of risk) based on recommendations of other authors (Roads and Traffic Authority, 2010) has been established. Applying this rating scale, less than 1% of evaluated assets are in the R1 category, whereas in the majority of the studied cases (41%) the risk level fits in the lowest category (R5).

The failure mechanism presenting the highest risk is the structural collapse of the culvert barrel (directly related to the general deterioration condition observed inside the culvert barrel) and erosion caused by overtopping flows (due to partial or total blockage / insufficient hydraulic capacity during peak flows). Since many of culverts studied are above twenty five years age and the impact caused by the structural collapse of the culvert barrel implies the total cut of service in the roadway and the complete restitution of the culvert, the incidence level on the total risk in this failure mechanism remains high.

On the other hand, the most common pathologies affecting culverts have been identified. Each of these pathologies affects one or more failure mechanisms, depending on their nature. The need to extract these risk factors from the final diagnosis is because they represent the most reliable indicator when planning preventive repair and / or replacement actions based on the preliminary diagnosis results. As shown in Table 2, the risk factor that has the highest percentage of incidence on total risk is that related to the hydraulic capacity of the culvert. This average value includes the cases where the culvert does not present any specific pathology, so the remaining risk (100%) is due to the residual probability associated with failure mechanisms triggered by rainfall events. In order of relevance, the pathologies associated with the structural deterioration of the culvert barrels and headwalls at inlet or outlets, partial or total blockage of the culvert barrel and undermining at inlet/outlet are highlighted.

4. Conclusions and further implementation

In this section the potential of the culvert risk assessment methodology applied in the development of systems of investments prioritization in the maintenance of this type of structures has been discussed.

Once the preliminary diagnosis of the set of assets studied has been obtained, the following level of development of the methodology is presented: the analysis of investments in maintenance and its prioritization based on risk reduction criteria. The selected rehabilitation measurements or remedial works in a particular culvert will be determined in the first place for operational restrictions (the rehabilitation techniques depend, in most cases, on the cross section typology, shape and material of the culvert barrel) and in the second place for the study of the risk factor currently operating on the asset. The cost of remedial works is determined according to the characteristics of the culvert (general field measurements) while the repercussion in terms of risk reduction per year, is evaluated considering the specific impact of every rehabilitation measurement in the global risk assessment considering the current and the forecasted scenario. Since the economic repercussion of the actions to be undertaken can be considered virtual benefits, it is possible to carry out a profitability analysis of the set of potential remedial works previously determined, according to the service life of each action considered.

Once this analysis is performed, it is possible to prioritize the actions according to the profitability indicators provided in the short, medium and long term, not only to evaluate the profitability of different potential actions in the same asset, but to select the most profitable updates linked to the set of culverts studied of a particular railway network, and prioritizing its implantation during a reference period of time.

The type of remedial works to be implemented in the profitability analysis fit the most common pathologies detected in culverts:

- Restitution of the culvert barrels (increasing hydraulic capacity of the culvert during peak flows).
- Erosion protection measures at inlets/outlets.
- Restitution and repair works in headwalls and wingwalls, according to the current condition of the asset.
- Restitution and repair works inside the culvert barrels, leakage prevention measures.
- Removal of blockage elements (debris and vegetation) inside the culvert barrels.

In conclusion, this works outlines the necessity of having a methodology for identifying and prioritizing investments projects in culvert maintenance widely supported by robust applications capable of massively analysing overall culvert risk conditions in transport networks. Global analysis of the results could lead to the establishment of a common management framework including alert and support-decision systems implementation in the planning and execution of investments in preventive maintenance actions of culverts.

5. Acknowledgments

The development of the methodology proposed in this document and its application has been carried out thanks to the development of specific calculation and management tools inside the HidraSmart-CM (Culvert Management) project (Hidralab S.L.)

6. References

- Balkham M., Fosbeary C., Kitchen A., Rickard C., 2010. Culvert design and operation guide. CIRIA Publishing, London, UK.
- Federal Highway Administration, 2010. Culvert assessment and decision-making procedures manual. U.S.Department of Transportation.
- Galán A., Diaz S., Gonzalez J., 2016. Erosión local a la entrada y a la salida de obras de drenaje transversal. XXVII Congreso latinoamericano de Hidráulica, 2016.
- Najafi M., Bhattachar, D., 2010. Development of a culvert inventory and inspection framework for asset management of road structures. Journal of King Saud UniversityScience 32, 243-254.

- Ministerio de Fomento, 2015. Instrucción 5.2-IC Drenaje Superficial.
- Ministerio de Fomento, 2014. Documentos complementarios no contradictorios para la aplicación de los Eurocódigos para el cálculo de puentes de ferrocarril.
- Perrin J., Jhaveri C., 2003. The economic costs of culvert failures.Transportation Research Board, USA.
- Roads and Traffic Authority, 2010. Culvert risk assessment guideline. NSW, Australia.
- Salem O., Salman B., Najafi M., 2012, Culvert asset management practices and deterioration modelling. Transportation Research Record: Journal of the Transportation Research Board 2285, 1-7.

Economic impacts

Fundación de los Ferrocarriles Españoles

360.revista de alta velocidad

Economic impacts

International Congress on High-speed Rail:

Technologies and Long Term Impacts

Ciudad Real (Spain), October, 2017

Escuela de Ingenieros
de Caminos, Canales y Puertos
de Ciudad Real
UNIVERSIDAD DE CASTILLA-LA MANCHA

Economic impacts

360.revista de alta velocidad
número 5 - junio - 2018. Pág 369 - 383

Economic, geographical and time-based exclusion as main factors inhibiting Spanish users from choosing High Speed Rail

Pagliara, Francesca

Menicocci, Fabrizio

Vassallo, José Manuel

Gómez Sánchez, Juan

University of Naples Federico II[1]
Transportation engineer at Mediterranean Shipping Company
Polytechnic University of Madrid

Abstract

Very few contributions in the literature have dealt with the issue of social exclusion related to High Speed Rail systems. The objective of this manuscript is to understand what are the factors excluding users from choosing High Speed Rail services considering the case study of Spain. For this purpose, a Revealed Preference survey was conducted between November-December 2015. A questionnaire was submitted to users of the Spanish transport system travelling for long-distance trips. The research aimed at investigating their perception of the High Speed Rail system and the factors inhibiting passengers or excluding them from its use. Data about their socioeconomic characteristics were collected as well. The analysis of the survey identified a relationship between social exclusion and High Speed Rail in Spain, especially in terms of geographical exclusion.

Keywords: social exclusion, High Speed Rail, geographical exclusion, accessibility.

---

[1] Pagliara, Francesca. University of Naples Federico II. Email: fpagliar@unina.it. (corresponding author)
Menicocci, Fabrizio. Transportation engineer at Mediterranean Shipping Company. Email fabrizio.menicocci91@gmail.com
Vassallo, José Manuel. Polytechnic University of Madrid. Email: jvassallo@caminos.upm.esanchez
Juan Gómez. Polytechnic University of Madrid. Email: juan.gomez.sanchez@upm.es

1. Introduction

Scial exclusion and poverty are two concepts that are often used interchangeably. In the literature, the term "social exclusion" has been explained by several authors. According to Silver (1994) social exclusion is *"a multidimensional process of progressive social rupture, detaching groups and individuals from social relations and institutions and preventing them from full participation in the normal, normatively prescribed activities of the society in which they live."*

Indeed it is generally agreed that exclusion refers to a dynamic process and not necessarily to an end-result (Lucas, 2011; 2012; Jones and Lucas, 2012; Lucas and Musso, 2014), i.e. "who" and "when" someone is excluded can change over time. The concept of social exclusion is based on inclusion into civil society. On the other hand, absolute poverty was defined by the United Nations as *"a condition characterised by severe deprivation of basic human needs, including food, safe drinking water, sanitation facilities, health, shelter, education and information "* (UN, 1996). Therefore, low income categories are not necessarily experiencing social exclusion, since being excluded can be a form of deprivation with innate importance in addition to its causal relations with other issues. Exclusions of the social nature can in turn lead to other deprivations that may significantly decrease the quality of life (Sen, 2000).

Social exclusion is a state in which an individual is not able to take part in activities of civil society, considered normal and expected within society. By social inclusion it is meant, on the other hand, the ability to participate adequately in society. According to Levitas (2007), social exclusion is *"the lack or denial of resources, rights, goods and services, and the inability to participate in the normal relationships and activities, available to the majority of people in a society, whether in economic, social, cultural or political arenas. It affects both the quality of life of individuals and the equity and cohesion of society as a whole."*

Poor transport options and alternatives can be a result of social exclusion and can also reinforce it. Transport could represent a factor of social exclusion since a lack of accessibility prevents people from participating in work, educational activities, community events, etc. (Kenyon, *et al.*, 2003).

Some previous interest can be identified for analyzing the potential relationship between transport systems and social exclusion. This is, for example, the case of UK, since a renewed interest in ameliorating the effects of social exclusion was observed after the election of the Labour government in 1997. A Social Exclusion Unit (SEU) was established to monitor and influence policy across all Whitehall Departments. In 2002 the Unit turned its attention to travel, transport and access, seeing these as processes implicated in the reproduction of social exclusion. In this respect, they pointed out that *"recent years have seen a growing recognition that transport problems can be a significant barrier to social inclusion"* (SEU, 2003). Likewise, in 2004 the FIA Foundation promoted a study to compare the position of the G7 countries in relation to transport and social exclusion at the urban level (FIA Foundation, 2004). In this report, it is worth noticing that no citation to HSR systems was made.

In the academic literature different approaches have been proposed to address the topic of social exclusion related to transport systems. Among other issues, it has been recommended to integrate transport systems planning with urban and social policies. One first step towards the reduction of social exclusion might be that of promoting activities to increase accessibility. At this point, the notable accessibility increases from new High-Speed Rail (HSR) services, which can play a significant role.

In the last decades, an important expansion of the HSR network has been observed in Europe. Indeed several European cities and regions are served today by HSR, and national agendas have planned significant extensions of this type of networks in the next decade. In 2016, the

European HSR network had over 8,100 km in service but it is planned to reach around 22,000 km in 2025. This shows not only the actual relevance of HSR services, but especially the central role that this infrastructure is going to achieve in the European transport policy. HSR is highlighted as a key future transport mode by the EC white paper *“Roadmap to a Single European Transport Area – Towards a competitive and resource efficient transport system”*. Its recent expansion and its planned extension could induce important geographical effects (Givoni, 2006) at different territorial scales (European, national and local ones) even though the role of HSR can differ in European regions due to the fact that networks, services and improvements of accessibility are diverse depending on each country and served city.

Spending public money in the construction of HSR lines has been defended as a socially desirable public investment which produces several types of benefits such as passenger time savings, increase in comfort, generation of new trips, reduction in congestion and delays in roads and airports, reduction in accidents, reduction in environmental externalities, release of needed capacity in airports and conventional rail lines, and wider economic benefits including the development of the less developed regions (De Rus, 2008).

Although these advantages, it is relevant to highlight how expensive is building and operating the new transport system. Indeed, today most of the HSR lines in Europe are subsidized, with the consequence that there is a trade-off between economic exclusion and the economic feasibility of these systems. For instance, for the case of Spain, Betancor, *et al.*(2015) analysed the economic feasibility of the HSR network and did not recognised its economic benefit, therefore other additional social and political factors have motivated the development of the HSR services in the country.

The objective of this research is to analyse whether HSR systems can increase social exclusion for long-distance trips, taking into account that other transport alternatives are available to users. This constitutes a topic of great interest given that future transportation systems investments seem to be focused on these services, mainly in Europe. Specifically, the case study of Spain is considered since it has one of the longest HSR network worldwide.

This paper is organised as follows. Section 2 reviews the current literature addressing the link between HSR and social exclusion. Section 3 briefly introduces the Spanish HSR system. In section 4 a binary logit mode choice model is introduced in order to identify the explanatory variables potentially affecting the choice of HSR and the impact of social exclusion. Section 5 presents and discusses the results. Conclusions and further perspectives are presented in section 6.

2. Social exclusion and HSR systems

Several contributions on the social effects brought by HSR systems have been analyzed in the literature (Vickerman, 1997; Preston and Wall, 2008; Pagliara, *et al.*, 2015). However only few contributions have been conducted on the impact of HSR systems on social exclusion. Among the very few studies present in the literature, the statistical analysis of surveys carried out by Cass, *et al*. (2005) reports interesting results. It indicates that HSR has both positive and negative social impacts. The positive social impact is represented by the increased accessibility and activities for commuting HSR users. The key concept of accessibility highlights the relationship between the system of activities located in a given territory and the transport system serving it.

According to Cascetta (2009), the concept of accessibility may refer alternatively to: a) the need to carry out some activities –shopping, work, education, etc.– by an individual who is in a certain area (active accessibility), or b) the need to be physically reached by potential users –customers, employees, suppliers, etc.– for an activity that is located in a certain area (passive accessibility). The nature of accessibility is influenced by the time-space organization in households, the nature and performances of the transport system, and the nature of time-space organization of the facilities and opportunities individuals are seeking to access. According to Cass, *et al.* (2005), HSR improves trips for working purposes by providing fast rail connections between main cities. On the other hand users who cannot afford HSR or live far from stations can be socially excluded and have problems when searching for better jobs. The introduction of a public transport system plays an important role in the social exclusion or inclusion of "transport poor" populations. HSR might encourage a hyper-mobile society, which can abandon people without access to the fastest modes of transport. This can be avoided only through thoughtful policies.

The study carried out in Spain by Monzón, *et al*. (2010) highlights the role played by the selection of the commercial speed. Indeed, an increase from 220 km/h to 300 km/h, in a given corridor, results in significant negative impacts on spatial equity between locations with and without a HSR service. At this point, it is necessary to point out that HSR cannot reach certain locations due to geographical and/or economic reasons. The same authors propose an assessment methodology for HSR projects following a twofold approach, i.e. addressing issues of both efficiency and equity. The procedure uses spatial impact analysis techniques and is based on the computation of accessibility indicators. Efficiency impacts are evaluated in terms of increased accessibility resulting from HSR projects, with a focus on major urban areas. Likewise, spatial equity implications are derived from changes in the distribution of accessibility values among these urban agglomerations (Monzón, *et al.*, 2013).

For the case of China, Chen and Wei (2013) addressed the case study of the Hangzhou East Rail station. This area is undergoing a rapid industrialization and thus workers' level of income is increasing significantly. However, HSR is still not affordable for the majority of the population. Also in China, Shi and Zhou (2013) aim at analysing transportation equity issues in terms of accessibility change experienced in those cities served by the HSR network. The main research findings, from the equity assessment perspective, reveal that investments in HSR systems do not have a strong impact in fostering social exclusion in terms of being excluded from the use of the new high speed services.

The case study of Italy was treated by Pagliara and Biggiero (2017). They conducted a Revealed Preference survey to identify the main motivations influencing travellers' choice of HSR, and found that users travelling alone choose HSR because of the reduction in travel time. Moreover the cost has an impact on the choice of this service because of the early booking convenient fares, which allow saving money to those travelling within given time periods. On the other hand, for those who have not chosen HSR, the main reason is the geographical exclusion, i.e. the low accessibility to the departure/arrival station. It follows the economic exclusion, i.e. the high cost of the HSR ticket. According to the authors, the fact that both criteria are greatly perceived by low income classes can be interpreted by the residential location of travellers. Furthermore, a quantitative approach was proposed to evaluate mode choice and the perception of social exclusion, considering two main aspects: economic and geographical exclusion. The analysis concluded a significant influence of income and the perceived geographical exclusion on intercity travellers' mode choice.Some of the previous results have been confirmed by a further Revealed Preference survey in UK (Pagliara, *et al.*, 2017). In this case the main motivation for those who have not chosen HSR is the economic exclusion, followed by the low accessibility to the departure/arrival station. In addition, the results of the study suggests that the introduction of a new transport mode, available in few points of the territory, brings social inequality, mainly perceived in terms of economic and geographical exclusion. Without thoughtful policies, HSR systems will encourage a hyper-mobile society that may abandon people without access to the fastest transport modes.

3. The HSR system in Spain

Since January 2016, Spain has the world´s second longest high-speed network, after China, and the longest in Europe (MAEC, 2013), with around 3,100 km of HSR lines in operation (see Figure 1). The service of HSR in Spain –known as AVE, Alta Velocidad Española– is operated by RENFE Operadora, the Spanish national railway company. Since 2005, AVE trains run on a HSR network owned and managed by ADIF, the public company in charge of the management of most of the Spanish railway infrastructure. Although RENFE Operadora is the only company operating the high-speed trains nowadays (CNC, 2013), private companies may be allowed to operate trains in the future, in accordance with the EU legislation. It is envisaged that the Madrid-Valencia corridor will be the first case to introduce competition in the HSR services in the country.

Figure 1. Spain´s HSR network as of May 2016, and annual passengers in the main AVE lines for 2012. Source: El País (2016)

During the last 20 years, the Spanish high-speed network has rapidly developed no matter whether there was sufficient demand to justify the construction of new lines. The expansion of this network has been considered in the last National Transport Plans as an essential element to promote social and territorial cohesion among territories. Indeed, one of the traditional objectives established in the transport agenda by previous governments has been linking the capitals of the 47 provinces in the peninsula by both high capacity roads and high speed rail services.

As a consequence of this policy the system is characterized by a reduced economic feasibility, and the suitability of the investments in the Spanish HSR has been strongly questioned in several occasions (see for instance De Rus, 2012 or Albalate, *et al*., 2011). Recently, Betancor, *et al.* (2015) analysed the economic feasibility of the Spanish HSR and found that only operating costs were covered, so they concluded that the investment was not profitable neither from a financial nor from a social point of view.

The first high-speed line was opened in 1992, connecting the cities of Madrid, Córdoba and

Seville. It was designed according to the technical standards of the French high-speed TGV. In the following years, the network was extended towards the northern part of the country, with the aim to create a connection to France and thus to the European high-speed network. Despite several problems encountered during the construction process, the Madrid-French Border line reached the cities of Zaragoza (2003) and Barcelona (2008). This line connects the two most populated urban areas in Spain -separated by 620 km- in 2h 30 min, which has led to a great success. Later this line was expanded to the city of Figueras, near the French border, and Perpignan (France). In Table 1 the annual passengers in the main AVE lines have been reported for the year 2012.

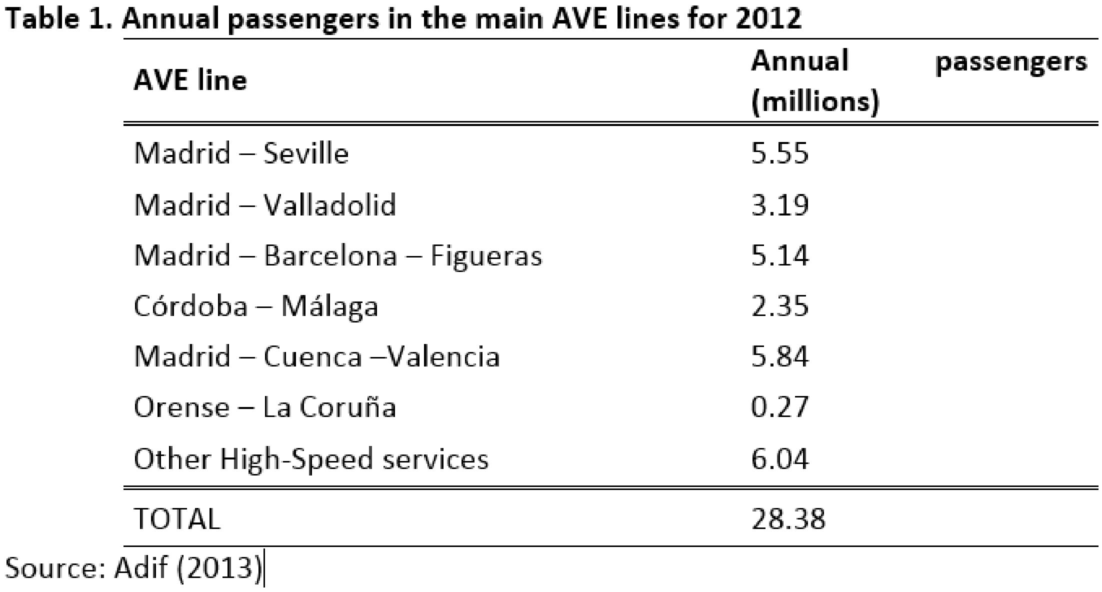


Table 1. Annual passengers in the main AVE lines for 2012

AVE line	Annual passengers (millions)
Madrid – Seville	5.55
Madrid – Valladolid	3.19
Madrid – Barcelona – Figueras	5.14
Córdoba – Málaga	2.35
Madrid – Cuenca –Valencia	5.84
Orense – La Coruña	0.27
Other High-Speed services	6.04
TOTAL	28.38

Source: Adif (2013)

Furthermore, in the last few years the high-speed network has been extended to connect some of the most populated cities in the Spanish Mediterranean coast such as Málaga (2007), Valencia (2010) and Alicante (2013), with great tourist appealing, too. However, due to the shortfall of financial resources, caused by the economic recession in Spain, the government has postponed or cancelled some of the high-speed projects already approved (Pagliara, *et al.*, 2012). The most recent extensions of the network up to Palencia and León (2015) have experienced significant delays, in line with other connections already under construction to reach areas such as the Basque Country, Granada or Extremadura. In this respect, the complex topography of the territory – Spain is the second most mountainous country in Europe – combined with a deterioration in the economic context and therefore in the public budget, have caused that currently some of the Spanish regions are not accessible to HSR services.

Despite the continuous financial losses experienced in previous years, in 2013 RENFE implemented alternative pricing systems in order to make HSR services accessible to a wider range of the population. For instance, discounts up to 70% are currently offered when buying single tickets in advance for certain train services. Reduced prices (up to 60%) can also be found when buying group tickets (4 people). Alternatively, students and elderly people are offered discounts of up to 50% and 60% respectively, over the standard ticket price. This policy, promoted by the Spanish Ministry of Transportation, has resulted in an average reduction of 11% in HSR tickets and has increased rail demand substantially in the last years. In 2015, the HSR services reached a total of 31 million passengers, constituting the peak in the historic trend and almost doubling the passenger traffic in 2008 before the implementation of the general discounts policy. This strategy has also lead to further positive effects, such as a notable decrease in the financial losses experienced by the public company RENFE.

4. The methodology

As pointed out above, this research is aimed at analysing whether HSR systems can increase social exclusion for long-distance trips, taking into account that other transport alternatives are available to users. This contribution is based on the framework of factors that may limit the mobility of socially excluded people, proposed by Church, *et al.* (2000). In this paper, the following categories of exclusion connected to transport, applied for urban trips, are mentioned:

1. Physical exclusion: physical barriers, i.e. lack of disabled facilities or timetable information, limiting accessibility to transport services.
2. Geographical exclusion prevents people from accessing transport services, especially those living in rural or peripheral urban areas.
3. Exclusion from facilities, concerning the low accessibility connected with facilities, like shops, schools, health care or leisure services.
4. Economic exclusion represents the high monetary costs of travel preventing or inhibiting access to facilities or employment and thus having an impact on incomes.
5. Time-based exclusion refers to other demands on time, like combined work, household and child-care duties, reducing the time available for travel.
6. Fear-based exclusion concerns to the fears for personal safety precluding the use of public spaces and/or transport services.
7. Space exclusion is the security or space management preventing given groups having access to public spaces, like first class waiting rooms at stations.

These categories have been adapted to medium-long distance trips to properly address the aim of the paper. Based on this assumption, a Revealed Preference (RP) survey was carried out in Spain between October and December 2015. The questionnaire was created on the Google platform and 414 useful responses were collected. Users were interviewed regarding the last interurban trip they made within Spain, and reported different trip characteristics such as the transport mode chosen, including HSR.

Due to the survey method used, based on the web platform, the sample needed to be weighted. The percentages of gender and age classes, based on the 2011 Spanish Census (INE, 2015), have been considered to adjust the sample. Then, those observations with a trip length lower than 80 km have been removed from the sample since they typically correspond to regional trips, not operated by the AVE services. In this case the authors tried to avoid any bias present in the data set used to make inferences.

Table 2 includes the socioeconomic characteristics of the whole sample, reporting figures for both HSR users and non-HSR users. It is interesting to notice that both groups mainly correspond to full time/part time workers. Particularly, the highest percentage (70%) is observed for HSR users, probably because they can afford this service, while for non-HSR users the percentage is around 55%. Concerning the monthly income, it can be noticed that 70% of the sample has an income between 1,000-3,000 Euro. Among HSR users, less than 70% of them has an income higher than 2,000 Euro, while non-HSR users are less than 53%, this result seems to highlight the influence of income on mode choice.

For the trip purpose, it results that most of the users travel for personal reasons. Among the HSR users around 20% travel for work or study purposes, while the percentage is just 5% for non-HSR users. This can be explained considering that users who do not travel for systematic trips prefer a cheaper transport alternative. This result is in line with previous research on HSR and tourism in Spain, such as Pagliara, *et al*. (2015). For the whole sample, less than 70% of the respondents travel alone or with the partner, therefore no more than two persons. It is interesting to highlight that those who have travelled for work purposes together with their colleagues have chosen AVE probably thanks to the special fares proposed by RENFE.

Moreover, transport mode choices by respondents for interurban trips within Spain revealed that HSR is the most used transport mode (47.1%), followed by car (31.6%). Other transport modes such as conventional rail or plane show a lower share (4.0% and 4.1%, respectively).

Table 2 – Summary statistics of the sample

Characteristics	Categories	HSR Users %	Non- HSR Users %	Sample tot %
Age	18-19	0.0%	9.2%	5.5%
	20-29	17.3%	10.1%	13.0%
	30-39	26.0%	13.6%	18.5%
	40-49	21.6%	16.6%	18.6%
	50-64	19.8%	23.8%	22.2%
	> 65	15.4%	26.9%	22.2%
Gender	M	50.4%	48.0%	49.0%
	F	49.6%	52.0%	51.0%
Nationality	Spanish	89.5%	97.2%	94.1%
	Others	10.5%	2.8%	5.9%
Education	Degree or more	95.2%	73.5%	82.1%
	Others	4.9%	26.5%	17.8%
Occupation	Full time/part time worker	72.1%	56.7%	63.0%
	Student	10.6%	12.5%	11.6%
	Unemployed	0.0%	2.0%	1.2%
	Freelance	1.9%	2.0%	2.5%
	Retired	15.4%	26.8%	22.2%
Monthly income	< €1000	3.4%	2.7%	3.0%
	€1,000 - €2,000	29.6%	44.4%	38.5%
	€2,000 - €3,000	35.2%	30.9%	32.7%
	€3,000 - €4,000	22.2%	11.6%	15.8%
	€4,000 - €5,000	4.0%	7.1%	5.9%
	> €5,000	5.5%	3.3%	4.2%
Trip purpose	Work	15.8%	4.8%	9.2%
	Study	4.3%	0.9%	2.3%
	Holiday	43.9%	33.0%	37.4%
	Personal activities	36.0%	61.2%	51.1%
Travel type	Alone	30.2%	31.7%	31.1%
	Partner	30.4%	39.6%	35.9%
	Colleagues	14.5%	0.0%	5.8%
	Friends	14.8%	15.1%	14.9%
	Relatives	10.1%	13.7%	12.3%

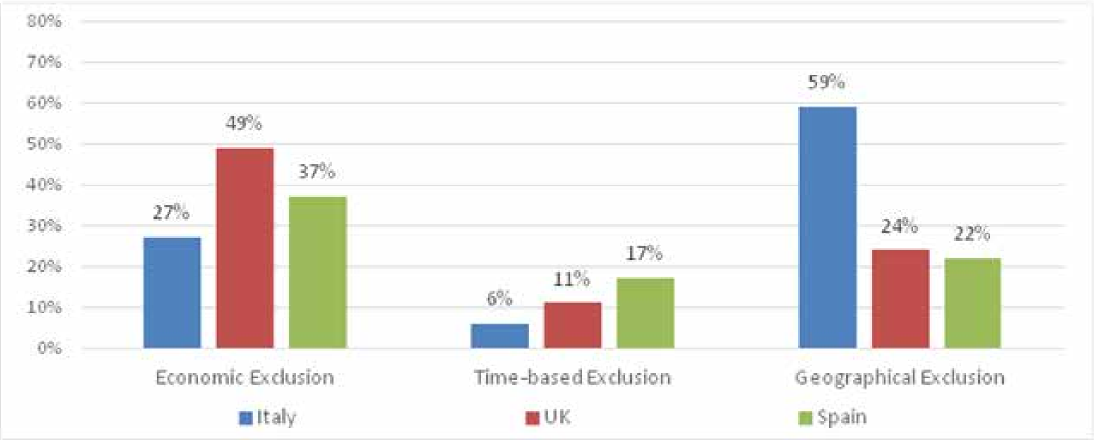
In Table 3, the choice among the seven categories of social exclusion proposed by Church, *et al.* (2000) is analysed. Specifically respondents were asked to rank their perception of each factor of exclusion from 1 (scarcely inhibiting) to 5 (highly inhibiting). As it can be seen, economic and geographical exclusion have turned out to be inhibiting users from choosing HSR, and therefore they have been the fundamental reasons for not using this means of transport by non-HSR users.

Table 3 - Church's categories of social exclusion (non-HSR users)

Economic	**Time-based exclusion**	**Spatial exclusion**	**Fear-based exclusion**	**Geographical**	**Physical**	**Facilities**
85.2%	22.6%	13.4%	2.6%	51.0%	20.8%	18.0%

These values have been compared with two similar case studies previously conducted in Italy and UK (Pagliara, et al., 2015b; 2016). In order to make the comparison properly, the values for Spain have been reported to the unit so that the total sum could be 100%. The results of the comparison (see Figure 2) show that geographical exclusion is a factor of exclusion mainly in Italy (60%) rather than in Spain and UK (less than 25%). This can be explained considering the mode choice and the different extension of the current HSR network in the three countries. In Italy, HSR is the first choice for passengers travelling long distances although the network is not extended and not capillary, in contrast to what happens in Spain, with a quite dense HSR network covering the territory in a quite homogeneous way (see Figure 2). In UK the network is not very extended, and therefore HSR is not considered as a real alternative mode for interurban trips, which could explain why geographical exclusion is less felt than the economic one in this case.

Fig. 2 - Comparison of the main components of social exclusion in HSR services for Italy, Spain and UK



Furthermore, Tables 4 and 5 show the categories of social exclusion and their relation to both trip purpose and household monthly income. As it can be seen, those who travelled for holiday or for personal reasons and have a lower income, feel excluded from HSR due to economic and geographical reasons. At this point, it should be noted that these two categories often coexist, since those individuals who have limited financial resources are also unable to live in areas accessible to HSR services, typically city centres.

Table 4 – Church et al.'s categories of social exclusion versus trip purpose (non-HSR users)

Trip purpose	Categories of Exclusion						
	Economic	Time-based	Spatial	Fear-based	Geographical	Physical	Facilities
Work	4.5%	-	-	-	5.5%	-	-
Study	35.2%	11.7%	24.1%	32.4%	44.5%	23.5%	22.8%
Holiday	1.0%	2.8%	3.0%	-	-	2.6%	3.1%
Personal activities	59.3%	85.5%	72.9%	67.6%	50.0%	73.9%	74.1%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Table 5 – Church et al.'s categories of social exclusion versus household monthly income (non-HSR users)

Monthly income	Categories of Exclusion						
	Economic	Time-based	Spatial	Fear-based	Geographical	Physical	Facilities
Low-medium	77.9%	93.0%	67.4%	57.2%	73.8%	69.4%	55.5%
High	22.1%	7.0%	32.6%	42.8%	26.2%	30.6%	44.5%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

5. HSR or not HSR? That is the question

In order to explore the potential impact of HSR services on social exclusion, an econometric model has been developed based on the data collected from the survey. The model is aimed at identifying the explanatory factors determining the mode choice made by respondents, particularly whether they used the HSR or not HSR for the last interurban trip made. To that end, a binary mode choice logit model (Cascetta, 2009) has been proposed, with the binary variable choosing or not HSR as the dependent variable.

The model follows the traditional binomial logit form, widely referred in the literature. A detailed description of binary choice models is beyond the scope of this paper, so the reader is directed to Ben-Akiva, *et al*. (1999) or Ortúzar, et al. (2011) for further details.

Binary choice models are derived from the utility maximizing theory, according to which decision makers are utility maximizers. Then, the individual choose, among all the options available, the alternative measuring her/his utility, which can be determined by a number of explanatory variables. The utility (V^i_j) gained by individual *i* for choosing alternative *j* can be determined by explanatory variables X^i_{kj}, and written as shown in equation (1):

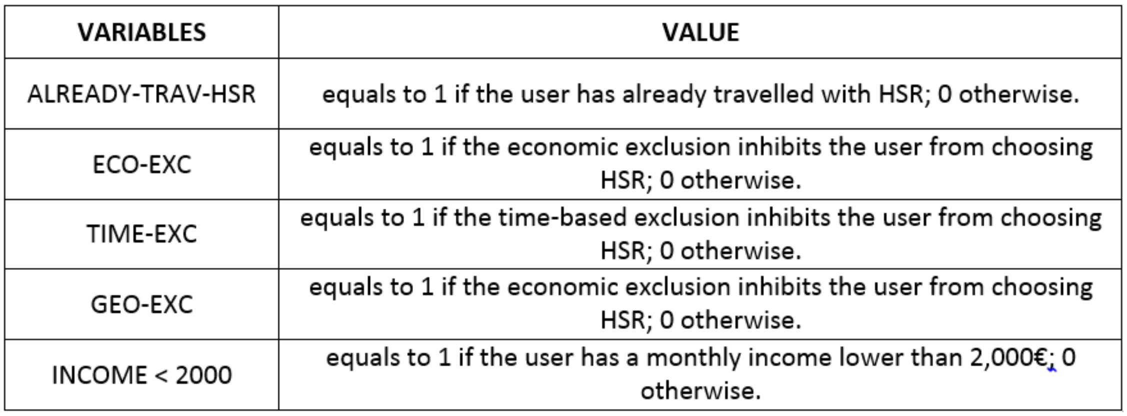
$$V^i_j(X^i_j) = \sum_k \beta_k X^i_{kj} \tag{1}$$

Economic theory assumes that the individual *i* will choose the option with the highest utility. As explained by Ben-Akiva (1985), in the general form of a binary choice model, the probability that user *i* will choose alternative *j* can be expressed as follows:

$$p[m'/oshd] = \frac{exp(V_{m'/oshd})}{\sum_m exp(V_{m/oshd})} \quad (2)$$

In our case, the probability of choosing HSR or other transport mode (not HSR) is here computed, from a given origin o, for a given purpose s, in the time period h and to a given destination *d*. Specifically, the dependent variable equals to 1 if the user travelled by HSR, or 0 if other transport mode was chose. The explanatory variables X_k chosen to model whether users chose HSR for interurban trips are reported in Table 6.

Table 6 – Explanatory variables included in the model specification



VARIABLES	VALUE
ALREADY-TRAV-HSR	equals to 1 if the user has already travelled with HSR; 0 otherwise.
ECO-EXC	equals to 1 if the economic exclusion inhibits the user from choosing HSR; 0 otherwise.
TIME-EXC	equals to 1 if the time-based exclusion inhibits the user from choosing HSR; 0 otherwise.
GEO-EXC	equals to 1 if the economic exclusion inhibits the user from choosing HSR; 0 otherwise.
INCOME < 2000	equals to 1 if the user has a monthly income lower than 2,000€; 0 otherwise.

Then, the binary logit approach predicts the so-called logit of the odds ratio, L_k, given multiple explanatory variables X_k. The model specification finally adopted has the classical form of a binary logit model:

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right) \quad (3)$$

where *p* is the probability of choosing HSR and the argument of the natural logarithm is called odds (Bewick, et al., 2005). The relationship between *p* and explanatory variables Xi can be written as follows:

$$\text{logit}(p) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k \quad (4)$$

where:

β_0 is the intercept, i.e. the expected value of p when all the predictors are 0;

β_i are the regression coefficients, estimated through the calibration process.

As a consequence of this linearization process, the interpretation of the β_k coefficients is different compared to standard linear regression models. The slope coefficient suggests that for a unit increase in a certain explanatory variable X_k, the weighted log of the odds in favor of a certain alternative (Y=1) increases of $e^{\beta k}$. Moreover, for a unit increase of a given explanatory variable X_k, the odds ratio in favor of happening Y=1 increases of $e^{\beta k}$. Furthermore, unlike a simple linear regression, logistic regression parameters are usually estimated with the method of maximum likelihood, an iterative process calculating small corrections until the convergence is reached.

Regarding the goodness of fit of the model, R^2 statistics refer to the entire model and indicates how useful the explanatory variables are in predicting the response variable. The Cox & Snell and the Nagelkerke R^2 are two of the most used statistics. The maximum value for Cox & Snell R^2 is less than 1 while the Nagelkerke R^2 provides a correction of this one and covers the full range from 0 to 1 and therefore is often preferred.

A preliminary analysis has been carried out to check potential collinearity between the explanatory variables ECO-EXC and INCOME < 2000. Specifically a correlation analysis based on the Chi2 test allows determining the level of relationship between both variables. With a level of significance equal to 0.05 and 1 degree of freedom, according to the Chi^2 test the alternative hypothesis is accepted and the variables are independent. Estimation results for the binary choice model specified are reported in Table 7. According to the estimates it can be observed that:

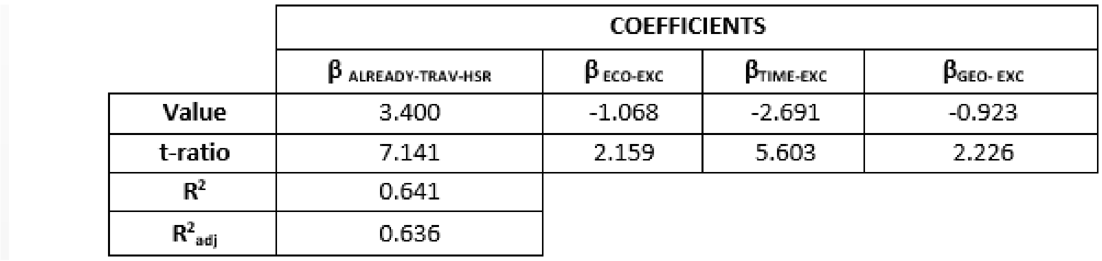
- Those users who have travelled at least once by Spanish HSR (AVE) have a higher probability of choosing it again ($\beta_{ALREADY\text{-}TRAV\text{-}HSR}$ positive and highly significant).
- Those users who feel to be economically excluded (i.e. for whom the cost of the HSR ticket is perceived high) have a lower probability of choosing HSR ($\beta_{ECO\text{-}EXC}$ negative and significant).
- Those users who feel to be time-based excluded (i.e. who feel constrained due to the impossibility of reconciling their commitments with train frequency and timetable) have a low probability of choosing HSR ($\beta_{TIME\text{-}EXC}$ negative and highly significant), keeping the rest of variables constant.
- Those users who feel to be geographically excluded (i.e. who have a lower accessibility to AVE stations) have a low probability of choosing HSR ($\beta_{GEO\text{-}EXC}$ negative and significant).
- Those users who have a monthly income under 2,000 Euro have a lower probability of choosing AVE ($\beta_{INCOME<2000}$ negative). Moreover this variable is not statistically significant (t-ratio = 1.684 < 1.960).

Table 7 - Estimation results (not all the variables significant)

	COEFFICIENTS				
	$\beta_{ALREADY\text{-}TRAV\text{-}HSR}$	$\beta_{ECO\text{-}EXC}$	$\beta_{TIME\text{-}EXC}$	$\beta_{GEO\text{-}EXC}$	$\beta_{INCOME<2000}$
Value	3.671	-1.040	-2.688	-0.844	-0.606
t-ratio	7.099	2.062	5.609	2.009	1.684
R^2	0.649				
R^2_{adj}	0.644				

Regarding the goodness of fit obtained, the Nagelkerke R^2 is quite high (around 0.65) as well as R^2_{adj}. In this respect, the explanatory variables chosen in the model seem to properly reproduce users' behaviour. Moreover, given that the variable INCOME<2000 is not correlated with the ECO-EXC variable and did not resulted statistically significant, it has been removed from the model. Then, the final model estimation is presented in Table 8, including only significant variables.

Table 8 - Estimation results (with all significant variables)



	COEFFICIENTS			
	$\beta_{ALREADY-TRAV-HSR}$	$\beta_{ECO-EXC}$	$\beta_{TIME-EXC}$	$\beta_{GEO-EXC}$
Value	3.400	-1.068	-2.691	-0.923
t-ratio	7.141	2.159	5.603	2.226
R^2	0.641			
R^2_{adj}	0.636			

In this model the values of both R^2 and R_{2adj} are quite high. Moreover this model shows how the choice of AVE is influenced by having already used it (at least once) and by the economic, geographical and time-based exclusion.

6. Conclusions and further perspectives

In this paper the relationship between HSR and social exclusion has been analysed. Following the framework proposed by Church, et al. (2000), the motivations fostering the choice of HSR in Spain have been analysed together with the factors inhibiting from the use of this service.

The results of a Revealed Preference survey have shown that only some criteria are perceived by the users when making the choice. For those who have not chosen HSR, the main reason is the economic exclusion, i.e. the cost of the HSR ticket. It follows the geographical exclusion, i.e. the low accessibility to the departure/arrival station. The fact that both criteria are greatly perceived by low income classes can be interpreted by the residential location of this type of travellers. Regarding the relationship between the perception of economic and geographical exclusion, it results that those individuals with higher incomes live in city centres, generally served by good public transport and taxi services. Indeed a good public transport system can allow an easy access to the departure/arrival stations. Likewise, improving accessibility to HSR stations outside metropolitan contexts can play a key role to reduce the geographical exclusion within the same country, even in the case of an extended HSR network as the Spanish one.

To support the results of the survey, a quantitative approach has been proposed, through the specification and calibration of a mode choice model, which aims at evaluating the perception of social exclusion. Three aspects of social exclusion have been considered (economic, geographical and time-based exclusion). Estimation results show how the choice of HSR services is influenced by having already used it (at least once) and by the economic, geographical and time-based exclusion.

Further perspectives will consider the collection of a larger data set which can support these findings and the specification and calibration of more sophisticated models. Specifically, more complex mode choice model specifications can be adopted to model users´ choices such as nested or mixed logit. These alternatives could be useful to explore additional factors such as the choice among all available alternatives (i.e, not only HSR versus non-HSR) or heterogeneity in preferences among respondents. Furthermore, a structural equation approach could be estimated to identify the specific aspects determining perceptions towards each component of social exclusion (economic, geographical, etc.) in more detail.

7. References

- ALBALATE, D., BEL, G. (2011). *Cuando la economía no importa: auge y esplendor de la Alta Velocidad en España*. Revista de Economía Aplicada 55, 171-190.
- BEN-AKIVA, M. (1985). *Discrete Choice Analysis Theory and Application to Travel Demand, MIT Press.*
- BEN-AKIVA, M., BIERLAIRE, M. (1999). *Handbook of Transportation Science.* Hall, Randolph (Ed.), Springer, Boston.
- BETANCOR, O., LLOBET, G. (2015). *Contabilidad financiera y social de la Alta Velocidad en España. Estudios sobre la Economía Española* – 2015/08, FEDEA.
- BEWICK, V., CHEEK, L., BALL, J. (2005). Statistics review 14: Logistic Regression. *Critical Care 9*, pp. 112-118.
- CASCETTA, E. (2009). *Transportation systems analysis: models and applications*. Springer, New York.
- CASS, N., SHOVE, E., URRY, J. (2005). Social exclusion, mobility and access. *The social review 3*, pp. 539-555.
- CHEN, C-L, WEI, B. (2013). *High-Speed Rail and urban transformation in China: the case of Hangzhou east rail station.* Built Environment 39, pp. 385–398.
- CHURCH, A., FROST, M., SULLIVAN, K. (2000). *Transport and social exclusion in London. Transport Policy 7*, pp. 195-205.
- Comisión Nacional de la Competencia, CNC. (2013). *Informe sobre la competencia en el transporte de mercancías por ferrocarril en España.*
- DE RUS, G. (2012). *Economic evaluation of the high speed rail. Expert Group on Environmental Studies.* Ministry of Finance, Sweden.
- DE RUS, G. (2008). *The Economic effect of High Speed Rail investement, discussion paper no. 2008-16*, University of Las Palmas.
- FIA FOUNDATION. (2004). Transport & Social Exclusion: A survey of the Group of Seven nations, summary report, Editor K. Lucas.
- GIVONI, M. (2006). *Development and impact of the modern High-Speed Train: a review. Transport Reviews 5*, pp. 593-611.
- INE, INSTITUTO NACIONAL DE ESTADÍSTICA. (2016). INE Official website: www.ine.es
- JONES, P., LUCAS, K. (2012). *The social consequences of transport decision-making: clarifying concepts, synthesising knowledge and assessing implications.* Journal of Transport Geography 21, pp. 4-16.
- KENYON, S., LYONS, G., RAFFERTY, J. (2003). *Transport and* social exclusion: investigating the possibility of promoting inclusion through virtual mobility. Journal of Transport Geography 10, pp. 207-219.
- LEVITAS, R., PANTAZIS, C., FAHMY, E., GORDON, D., LLOYD, E., PATSIOS, D. (2007). *The Multi-dimensional Analysis of Social Exclusion.* Department for Communities and Local Government (DCLG), London.
- LUCAS, K. (2011). *Making the connections between transport disadvantage and the social exclusion of low income populations in the Tshwane Region of South Africa.* Journal of Transport Geography 19, pp. 1320-1334.

- LUCAS, K. (2012). *Transport and social exclusion: where are we now? Transport Policy 20,* pp. 105-113.
- LUCAS, K., MUSSO, A. (2014). *Policies for social inclusion in transportation: An introduction to the special issue.* Case Studies on Transport Policy 2, pp. 37-40.
- MINISTERIO DE ASUNTOS EXTERIORES Y COOPERACIÓN, MAEC. (2013). *Spain's positioning. Leadership Key Factors,* Dirección General de Relaciones Económicas Internacionales.
- MONZÓN, A., ORTEGA, E., LÓPEZ, E. (2010). *Social impacts of high speed rail projects: addressing spatial equity effects*. Proceedings of the 12th WCTR, July 11-15, Lisbon, Portugal.
- MONZÓN, A, ORTEGA, E., LÓPEZ, E. (2013). *Efficiency and spatial equity impacts of high-speed rail extensions in urban areas.* Cities 30, pp. 18-30.
- ORTÚZAR, J.D.D., WILLUMSEN, L.G. (2011). *Modelling Transport, 4th ed. John Wiley & Sons, Ltd, Chichester.*
- PAGLIARA, F., LA PIETRA, A., GOMEZ, J., VASSALLO, J. M. (2015), *High Speed Rail and the tourism market: evidence from the Madrid case study.* Transport Policy 37, pp. 187-194.
- PAGLIARA, F., BIGGIERO, L. (2017). *Some evidence on the relationship between social exclusion and high speed rail systems.* HKIE Transactions 1, pp. 17-23.
- PAGLIARA, F., DE POMPEIS, V., PRESTON, J. (2017). *Travel cost: not always the most important element of social exclusion. Paper submitted to The Open Transportation journal, forthcoming.*
- PRESTON, J., WALL, G. (2008). *The Ex-ante and Ex-post Economic and Social Impacts of the Introduction of High-speed Trains in South East England.* Planning Practice & Research 23, pp. 403-422.
- SEU. (2003). *Making the Connections: Final Report on Transport and Social Exclusion, SEU, London.*
- SILVER, H. (1994). *Social exclusion and social solidarity: three paradigms.* International Labour Review 5-6, 531-578.
- SEN, A. (2000). Social Exclusion: concept, application and security, Asian Development Bank.
- SHI, J., ZHOU, N. (2013). How Cities Influenced by High Speed Rail Development: A Case Study in China. Journal of Transportation Technologies 3, pp. 7-16.
- UNITED NATIONS (UN). (1996). *Special Issue on the Social Summit, Copenhagen,* 6-12 March 1995, *Social Policy and Progress,* a review published by the United Nations, Vol 1, No 1, *New York,* United Nations Department of Publications.
- VICKERMAN, R. (1997). H*igh Speed rail in Europe: experiences and issues for future developments.* The Annals of Regional Science 31, pp. 21-38.

Economic impacts

360.revista de alta velocidad
número 5 - junio - 2018. Pág 385 - 407

Measuring The Long-Term Regional Economic Impacts of High-Speed Rail in China Using a Dynamic SCGE Model[1]

Zhenhua Chen

The Ohio State University[2]

Abstract

This paper introduces a comprehensive framework to assess the regional economic impacts of high-speed rail (HSR) in China with a focus on its long-term implications. This research has two major research highlights: First, the regional impacts of HSR are evaluated under a dynamic and spatial (multiregional) general equilibrium modeling framework. Such a framework provides a comprehensive understanding of the impacts with variations in both space and time. Second, the assessment provides a demonstrative example of an ex post evaluation of the impacts based on the actual rail infrastructure investment data for the period of 2002 – 2013 using on a dynamic recursive multi-regional CGE model. The research findings confirm that rail infrastructure development in China has a positive regional economic impacts with a gross output multiplier of 1.01 and a GDP multiplier of 0.1 in the long-run. The aggregate impacts were found to be the highest in the southwest region, whereas the impacts are relatively small in developed eastern regions. The research findings provide implications for future HSR development in both China and other countries.

Keywords: high-speed rail, regional economic impact, investment, dynamic, computable general equilibrium model.

---

[1] *Although the term "High-Speed Rail" is adopted, the discussion de facto focuses more broadly on the Chinese rail infrastructure development in general due to the following concerns: First, the Chinese HSR development policy involves not only the mid- and long-term development for HSR, regular passenger rail and freight rail are also included. Hence, a broader focus on rail would be more appropriate to evaluate the effectiveness of the related infrastructure planning and policy. Second, an assessment with a focus only on the HSR systems is not feasible due to the lack of specific statistical information reflecting the true "HSR" investment strategies and operating performance. Another important consideration is that since many new developed HSR infrastructure facilities, such as stations, rail tracks are also utilized to serve regular passenger rail service, such a broader assessment of rail would be reasonable to achieve a more practical investigation.*

[2] *Zhenhua Chen. The Ohio State University. Email: chen.7172@osu.edu. (corresponding author)*

1. Introduction

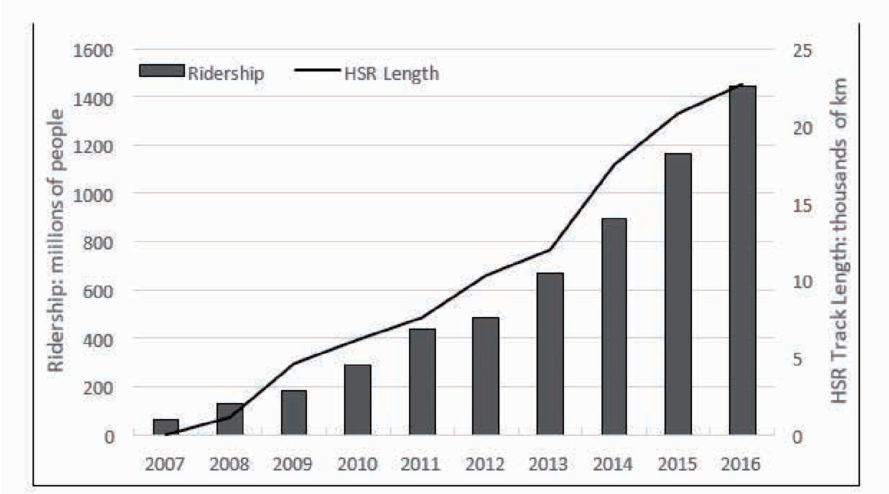
Rail infrastructure system in China has experienced an exponential expansion over the past decade due to a strong support from the central government. As illustrated in Table 1, the National Development and Reform Commission of China has launched three consecutive planning strategies for mid- and long-term rail network development in 2004, 2008 and 2016, respectively, to promote a continuous development of a nationwide rail infrastructure network. These strategies outline both high-level national and regional planning goals and objectives in terms of scale and technological specifications. Hence, they provide a clear guidance to rail industries and local governments for rail infrastructure development. One major highlight of these strategies is the development of an interconnected high-speed rail (HSR) networks to facilitate intercity passenger travel. These HSR systems are expected to be more advanced than conventional passenger rail because most trainsets have a capacity of running at 250 km/h or above given the introduction of Electric Multiple Units (EMUs) and the design of passenger dedicated lines (PDL). In addition to the benefits of a higher speed, the HSR systems also provide better travel experience than conventional rail service in terms of on-time performance, comfort, safety and service frequency (Givoni and Banister, 2012). The systems are also expected to alleviate the conflict between demand and supply for both freight and passenger rail transport (Chen and Haynes, 2015).

Table 1. Mid- and Long-Term Rail Network Planning Strategies in China

Planning Content		
2004 Planning Strategy	2008 Planning Strategy	2016 Planning Strategy
Period		
2003-2020	2008-2020	2016-2025 & 2030
Expected total		
100,000 km	120,000 km	175,000 km
Track length of		
12,000 km	16,000 km	38,000 km
HSR - Highlights of the planning strategy		
Separate passenger and freight traffic for trunk rail lines; Improve the rates of double-track and electrification to 50%; build 4-east-west bound and 4-north-south bound HSR trunk lines; Operating speed of HSR should be 200km/h or above.	Build 4-east-west bound and 4-north-south bound HSR trunk lines with a focus on developed regions with high population density; Build intercity rail systems for major megalopolis; Expand rail networks in the underdeveloped west regions.	Develop 8-east-west bound and 8-north-south bound HSR trunk lines; Operating speed should be 250km/h or above (HSR connecting major cities can be 350km/h, regional HSR connectors can be 250km/h, intercity rail can be 200km/h).

Source: Author's collection.

The deployment of the Chinese HSR is so enormous both in terms of the scale of infrastructure usages and the speed of deployment that none of the systems in other countries could compare with. As shown in Figure 1, the ridership of HSR keeps growing as more HSRs were deployed in operation. The total annual HSR ridership has reached 1.44 billion in 2016, which has expanded by 22 times during the decade as compared to its initial level in 2007. On the supply side, the track length of HSR also experienced a significant increase during the past decade. By the end of 2016, the total rail track in operation in China has reached 124 thousand kilometers, which includes 22 thousand kilometers of HSR connecting more than 400 cities nationwide. As a matter of fact, the pace of development is so rapid that it has exceeded the objectives outlined in the 2004 and the 2008 planning strategies. It is clear that with more people begin to enjoy the benefits of HSR for intercity travel, its regional economic impacts are likely to be even more substantial. In fact, our earlier assessment at the national level already shows that the deployment of HSR in China during 2002-2013 has contributed to a growth in GDP and welfare by 10 percent and 8.5 percent, respectively (Chen et al., 2016).



Source: Author's collection.

Figure 1. Evolution of High-Speed Rail Ridership and Track Length in China

Such a rapid expansion of HSR would not be achieved without a strong commitment and support of the central government in China. As revealed in Chen et al. (2016), public investment in rail sector, particularly in HSR development, grew rapidly, with an annual average rate of 20 percent since the implementation of the initial rail network development strategy in 2003. The latest objective of total rail track length outlined in the 2016 planning strategy has been expanded to 150 thousand kilometers, 30 thousand km of which will be HSR PDL. In particular, the national HSR trunk networks have been expanded from the previous plan with a 4 east-west bound and 4 north-south bound trunk lines to a system that consists of 8 east-west bound and 8 north-south bound HSR trunk lines, most of which will be designed as PDL with a speed of 250 km/h or above. The entire HSR system is expected to be completed by 2025. The ultimate objective is that more than eighty percent of major urban areas in China will be served by HSR, which is likely to significantly reduce the intercity travel time among contiguous provincial capital cities to 1-4 hours and 0.5- 2 hours for a trip that is within a megalopolis.

Despite these facts, skepticism about the effectiveness and economic values of HSR investment was also raised by some scholars. For instance, Button (2017) indicates that although politicians and rail enthusiasts have widely supported HSR infrastructure investment as a catalyst for economic development, their arguments on the anticipated economic growth effects from HSR are generally overly optimistic because most of the conclusions were derived from ex ante assessment in which the actual costs were often underestimated. In the case of China, Wu et al. (2014) suggests that while a limited number of HSR developments in the richest and most densely populated areas are reasonable due to the relative low value of time in China, a massive approach to HSR infrastructure development is problematic as new conventional rail is much more economical than HSR. Hence, they believe that there is no need for a massive development of HSR. The argument was endorsed by Zhao et al. (2015), who further indicates that a large scale HSR construction in China is likely to lead to an increase in market risk and economic loss due to the limited benefits of travel-time savings.

Ansar et al. (2016) also raised concerns on the massive infrastructure investment in China as they argue that such a large-scale investment in projects such as HSR is associated with a high risk due to the builtup of debt, monetary expansion, instability in financial markets and economic fragility. In fact, some scholars, such as Vickerman (2017), pointed out that the effectiveness of HSR investment on regional economic growth can be less transformative, because the contribution from HSR can be redistributive with some regions benefiting and others suffering depending on their abilities to take advantage of new opportunities. Hence, its overall wider economic benefits may not necessary be positive.

Although the intercity travel demand is likely to grow continuously for at least a few years given the strong momentum of regional economic development in China, it remains unclear what the long-term regional economic impacts of rail, in particular, HSR infrastructure development would become. In addition, given that the national rail planning strategies were intended to eliminate disparity across different parts of China so as to achieve a regional coordinated development, it is also essential to understand how do the economic impacts vary among different regions in China as a result of HSR development.

This study addresses these key questions using a dynamic spatial computable general equilibrium (SCGE) model. Our study has three major research highlights as compared to previous studies. First, the regional economic impacts of rail infrastructure investment in China are evaluated using a dynamic SCGE with considerations of capturing both a dynamic temporal evolutions of economic systems as well as the spatial (multiregional) general equilibrium interactions. The model is calibrated and updated with data that reflecting the Chinese economic system and the modeling framework was validated through a comparison with our previous analysis (Chen et al., 2016) that evaluated using a different CGE model at the national level. Hence, the empirical results are expected to be more robust and comprehensive.

Second, a detailed modeling framework based on a dynamic SCGE is developed for the first time, for the assessment of rail infrastructure development. The framework captures both the short-term direct impacts caused by capital investment in the process of rail infrastructure development and the long-term indirect impacts as a results of productivity improvement and technology progress. We believe that such a comprehensive modeling framework provides more meaningful implications to decision-makers and broader applications to practitioners to evaluate regional economic impacts of other types of infrastructures.

Third, an empirical analysis provides a thorough demonstration of the dynamic SCGE modeling process. Specifically, the CGE assessment allows us to capture the evolutions of regional economic impacts in a long-term period as more HSR systems being deployed. The empirical assessment of the long-term regional economic impacts of HSR is critical as it may facilitate future decision-making on infrastructure investment by improving our understanding on the effectiveness of current rail investment policies. In addition, a comprehensive understanding of the regional economic impacts of the Chinese HSR system also provides valuable implications to other countries that are either currently developing HSR or plan to build one in the near future.

The rest of the paper is organized as follows. Section 2 provides a methodological review of economic impact assessments with a focus on rail infrastructure systems. Section 3 introduces the key modeling framework for evaluating the long-term rail infrastructure development. Section 4 introduces the specific modeling structure of the dynamic SCGE model. Section 5 and 6 present data and the simulation results, respectively, whereas Section 7 summarizes and concludes.

2. Literature Review

The traditional approach to economic impact analysis of high-speed rail infrastructure is benefit-cost analysis (BCA). The method has been widely adopted particularly for an ex ante evaluation of HSR (Janic, 2003; De Rus and Nombela, 2007; Brand et al. 2014). The key process of BCA was to justify the value of HSR investment through comparing all the benefits and costs generated from the new developed infrastructure system. For instance, De Rus (2011) considers HSR investment in Spain was a second-best alternative based on a BCA. This is because a positive economic impact is expected given the considerations of levels of modal substitution, traffic volumes and operating costs. However, using BCA to evaluate large-scale infrastructure projects such as HSR, and particularly for a long-term assessment, can be problematic and challenging, as pointed out by Vickerman (2007), due to the uncertainties of project financing in a relative long-term period and the difficulties of selecting an appropriate discount rate to convert future benefits and costs into present terms for a comparison. In addition, BCA also has a limitation in incorporating the wider economic impacts such as agglomeration effects and spatial spillover effects as a result of improved transportation accessibility (Venables 2016; Button, 2017). As a result, the approach was more often applied for a project-level assessment in a short-run rather than a true “social and economic” assessment with a focus on a long-term period.

The second frequently adopted approach to evaluate economic impact of large-scale transportation infrastructure system is econometric analysis, which often follows the tradition of neoclassical growth theory. The key assumption is that transportation infrastructure can be considered as a separate input in addition to capital and labor in a standard production function Y = AF (K, L), where Y often denotes gross domestic product (GDP), A, K, and L represents level of technology, the share of capital and the share of labor, respectively. The output elasticity of transportation infrastructure is then estimated using regression models based on either a time-series or panel dataset. The estimated output elasticities are often found to vary substantially with a range between -0.15 and 0.56, due to the differences in the data and specific modeling forms (Melo et al. 2013). In the case of China, the average output elasticity of Chinese transportation infrastructure was found to be around 0.13 in a meta-analysis by Chen and Haynes (2017). Despite econometric analysis is able to identify the statistical association between infrastructure input and regional economic output from a long-term perspective, the evaluation outcomes using such an approach can still be incomplete due to the implicate assumption of a constant demand as a response to infrastructure change during the investigation period. The indirect impacts on the economic system as a response to demand change cannot be captured due to the lack of a feedback mechanism in regression analysis. In order to fully capture the effects of infrastructure system improvement from both the demand and the supply side, a general equilibrium assessment with a structure of simultaneous equation systems is needed.

The state-of-the-art approach to regional economic impact assessment is computable general equilibrium (CGE) analysis. The model, which is essentially a simultaneous equation system that involves thousands of equations and variables, uses actual economic data in an input-output format to simulate the interactions between the economy and changes in policy, technology or other external factors, the latter of which is often considered as a “shock”. After all parameters were calibrated in the initial simulation, the model then calculates an optimized solution (also known as equilibrium solution) given the introduction of a shock to the economic system. With the improvement of computer technology, CGE has been more frequently adopted for impact assessment of large-scale infrastructure systems. Depending on the regional scale and the consideration of temporal effect, CGE models can be classified into four types (shown in Table 2): a static single-region model, a dynamic single-regional model, a static multiregional model and a dynamic multi-regional model. The first two types of models were generally applied for an impact assessment at the national level or within a single-region. Because these models only include a single-region, the results of assessments on infrastructure investment are often limited due to the ignorance of spatial spillover effects that are manifested as the change of inter-regional commodity and factor input flows.

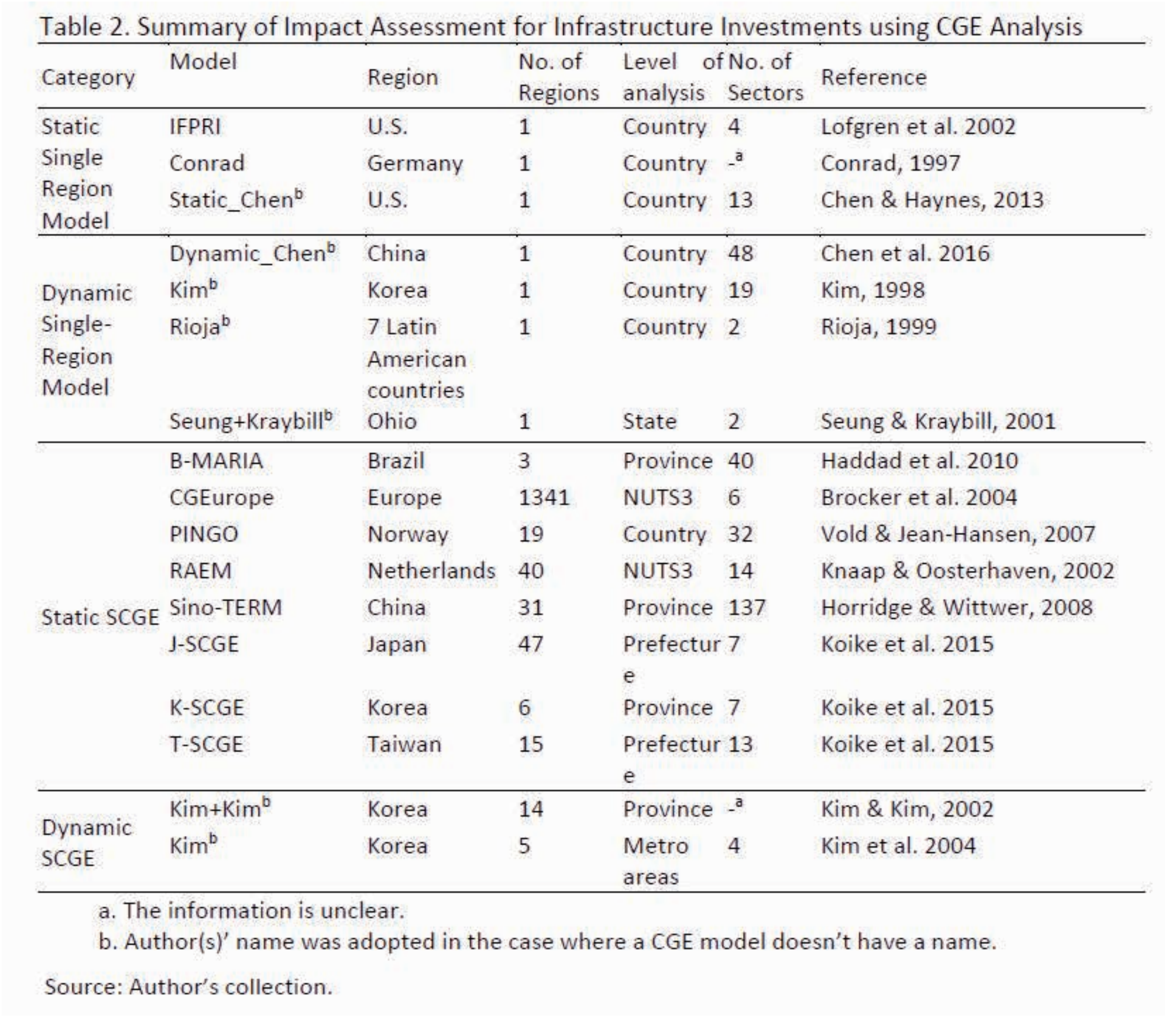


Table 2. Summary of Impact Assessment for Infrastructure Investments using CGE Analysis

Category	Model	Region	No. of Regions	Level of analysis	No. of Sectors	Reference
Static Single Region Model	IFPRI	U.S.	1	Country	4	Lofgren et al. 2002
	Conrad	Germany	1	Country	-[a]	Conrad, 1997
	Static_Chen[b]	U.S.	1	Country	13	Chen & Haynes, 2013
Dynamic Single-Region Model	Dynamic_Chen[b]	China	1	Country	48	Chen et al. 2016
	Kim[b]	Korea	1	Country	19	Kim, 1998
	Rioja[b]	7 Latin American countries	1	Country	2	Rioja, 1999
	Seung+Kraybill[b]	Ohio	1	State	2	Seung & Kraybill, 2001
Static SCGE	B-MARIA	Brazil	3	Province	40	Haddad et al. 2010
	CGEurope	Europe	1341	NUTS3	6	Brocker et al. 2004
	PINGO	Norway	19	Country	32	Vold & Jean-Hansen, 2007
	RAEM	Netherlands	40	NUTS3	14	Knaap & Oosterhaven, 2002
	Sino-TERM	China	31	Province	137	Horridge & Wittwer, 2008
	J-SCGE	Japan	47	Prefecture	7	Koike et al. 2015
	K-SCGE	Korea	6	Province	7	Koike et al. 2015
	T-SCGE	Taiwan	15	Prefecture	13	Koike et al. 2015
Dynamic SCGE	Kim+Kim[b]	Korea	14	Province	-[a]	Kim & Kim, 2002
	Kim[b]	Korea	5	Metro areas	4	Kim et al. 2004

a. The information is unclear.
b. Author(s)' name was adopted in the case where a CGE model doesn't have a name.

Source: Author's collection.

Spatial CGE (SCGE), also known as Multi-regional CGE model, which usually consists of more than two regions as independent economies in the modeling framework, are generally considered more relevant to regional economic impact assessment of infrastructure systems because interregional trade is explicitly taken into account through a bottom-up approach. Hence, the model is able to measure distinct regional impacts and associated regional spillover effects caused by a policy shock. As shown in Table 2, several SCGE models were developed and applied for transportation infrastructure assessment. For instance, Haddad et al. (2010) evaluated the long-run regional impacts of transportation sectors in Brazil using a SCGE model called B-MARIA.

The model was developed based on the MONASH model, which is a multiregional CGE model for the Australian economy originally built by Adams et al. (1994). In order to evaluate the regional economic impacts of the Trans-European Transport Networks, Bröcker (1998) developed a SCGE model consists of 1341 regions at the NUTS 3 level. The impacts were modeled by reducing transport costs along these links and tracing the effects through the economy.

PINGO is another SCGE model developed to predicts regional and interregional freight transport in Norway. Similar to CGEurope, the simulation was implemented through a shock on transport margin but rail transport is combined in the aggregate transport sector (Vold & Jean-Hansen, 2007). RAEM is a static SCGE model designed for the impact evaluation of a potential HSR infrastructure connecting Amsterdam and Groningen in the Netherlands (Knaap & Oosterhaven, 2002). The impact was simulated through reducing transport margins and the results were measured in terms of changes in travel time, numbers of jobs and consumer price index.

SinoTERM is a SCGE model developed by Horridge and Wittwer (2008) for the impact assessment of one railway project connecting Chongqing and Lichuan in Hubei province. The model was modified and updated based on The enormous regional model (TERM) of Australia. One key highlight of this model is that interregional freight transport was represented by interregional trade in the model. Hence, the analysis was able to capture spatial spillover effects on other regions as a response to a policy shock, which in this case, a reduction of freight transport margins (measured as a decline in F.O.B. price).

A series of SCGE models in a similar structure were developed by Koike et al. (2015) for the evaluations of HSRs in Japan, Korea and Taiwan. Different from other aforementioned SCGE models, passenger travels were considered separately for business trip and private (leisure) trip in Koike's models. The specific simulations were implemented through policy shocks on both factor inputs and transport margins. Despite these various approaches, the existing studies on the regional impact assessment of rail infrastructure remains limited, which can be summarized in the following aspects.

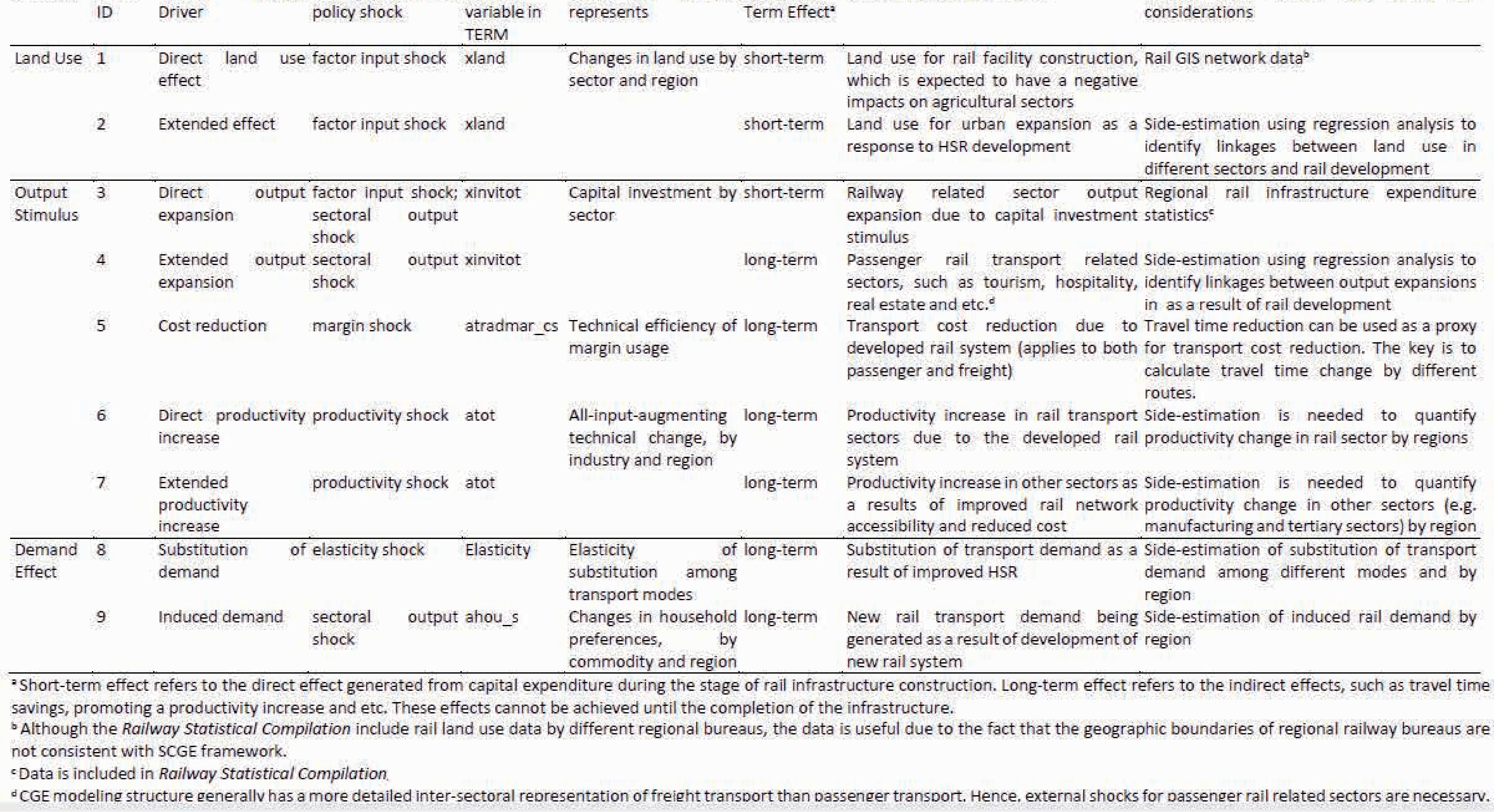
First, most of the SCGE models were essentially evaluated the impacts of transportation infrastructure from an ex ante perspective. Since simulations were generally conducted based on hypothetical scenarios with some arbitrarily specified policy shock values, implications of CGE modeling results can be quite constrained due to the lack of evidence based underpinnings (Chen and Haynes, 2017).

Second, plethora of studies evaluated the economic impacts of rail infrastructure system through a transport margin shock, whereas less attention was paid to other drivers of rail infrastructure development, such as a capital shock and a productivity shock. A related issue is that previous studies generally evaluated the impacts by assuming the infrastructure project is completed and in operation whereas there is a lack of a consideration to differentiate the impacts of a construction period from a post-construction period.

Third, there is a lack of a systematic approach for impact evaluations using a dynamic SCGE model. As a results, the spatial and temporal interactions of impacts as a result of rail infrastructure improvement are often ignored. In fact, only two applications of dynamic SCGE in regional economic impact assessments of infrastructure investment were found in our review, both of which were based on simplistic scenarios with a focus on Korea (Kim and Kim, 2002; Kim, et al., 2004).

Last but not the least, the existing assessments were generally implemented through a deterministic scenario based approach, whereas there is a lack of consideration for modeling uncertainty. In addition, the issue related to model validation is also usually unclear due to the intrinsic complexity of CGE modeling and the lack of reliable data to conduct meaningful validation test.

Table 3. SCGE Modeling Mechanism for Rail Infrastructure Development



Category	Driver ID	Direct Impact Driver	Applicability in CGE policy shock	Related variable in TERM	What the variable represents	Short/long-Term Effect[a]	What the driver represents	Related data source and some key considerations
Land Use	1	Direct land use effect	factor input shock	xland	Changes in land use by sector and region	short-term	Land use for rail facility construction, which is expected to have a negative impacts on agricultural sectors	Rail GIS network data[b]
	2	Extended effect	factor input shock	xland		short-term	Land use for urban expansion as a response to HSR development	Side-estimation using regression analysis to identify linkages between land use in different sectors and rail development
Output Stimulus	3	Direct output expansion	factor input shock; sectoral output shock	xinvitot	Capital investment by sector	short-term	Railway related sector output expansion due to capital investment stimulus	Regional rail infrastructure expenditure statistics[c]
	4	Extended output expansion	sectoral output shock	xinvitot		long-term	Passenger rail transport related sectors, such as tourism, hospitality, real estate and etc.[d]	Side-estimation using regression analysis to identify linkages between output expansions in as a result of rail development
	5	Cost reduction	margin shock	atradmar_cs	Technical efficiency of margin usage	long-term	Transport cost reduction due to developed rail system (applies to both passenger and freight)	Travel time reduction can be used as a proxy for transport cost reduction. The key is to calculate travel time change by different routes.
	6	Direct productivity increase	productivity shock	atot	All-input-augmenting technical change, by industry and region	long-term	Productivity increase in rail transport sectors due to the developed rail system	Side-estimation is needed to quantify productivity change in rail sector by regions
	7	Extended productivity increase	productivity shock	atot		long-term	Productivity increase in other sectors as a results of improved rail network accessibility and reduced cost	Side-estimation is needed to quantify productivity change in other sectors (e.g. manufacturing and tertiary sectors) by region
Demand Effect	8	Substitution of demand	elasticity shock	Elasticity	Elasticity of substitution among transport modes	long-term	Substitution of transport demand as a result of improved HSR	Side-estimation of substitution of transport demand among different modes and by region
	9	Induced demand	sectoral output shock	ahou_s	Changes in household preferences, by commodity and region	long-term	New rail transport demand being generated as a result of development of new rail system	Side-estimation of induced rail demand by region

[a] Short-term effect refers to the direct effect generated from capital expenditure during the stage of rail infrastructure construction. Long-term effect refers to the indirect effects, such as travel time savings, promoting a productivity increase and etc. These effects cannot be achieved until the completion of the infrastructure.

[b] Although the *Railway Statistical Compilation* include rail land use data by different regional bureaus, the data is useful due to the fact that the geographic boundaries of regional railway bureaus are not consistent with SCGE framework.

[c] Data is included in *Railway Statistical Compilation*.

[d] CGE modeling structure generally has a more detailed inter-sectoral representation of freight transport than passenger transport. Hence, external shocks for passenger rail related sectors are necessary.

3. Modeling for Rail Infrastructure Development

Our study fills these gaps by developing a comprehensive modeling framework for the assessment of rail infrastructure development using a dynamic SCGE model. The framework was expanded based on a modeling structure that outlined in Chen et al. (2016). As illustrated in Table 3, regional economic impacts of rail infrastructure development using a SCGE model can be derived from various direct impact drivers from in three categories: including land use effect, output stimulus effect and demand effect.

Specifically, land use effect can be further divided into direct land use effect, which refers to land use and acquisitions for direct infrastructure development purposes, such as converting arable land for the constructions of rail tracks, stations and facility centers. The other type of land use effect refers to extended land use activities as a results of rail development. For instance, the construction of a HSR may lead to a prosperous development of real estate around those new built HSR stations. These development is likely to have a further negative impacts on agricultural sectors due to the fact that reduction of arable land for urban development. From the perspective of CGE applicability, both land use effects can be modeled through a factor input shock and considered as short-term effect, meaning that the effects are achieved immediately once the shock is added to a regional economic system.

The second category of effect is the stimulus to output expansion as a result of an increase in infrastructure investment. Specifically, the stimulus effect includes a direct output expansion, an extended output expansion, a cost reduction and the direct and extended productivity increases. A direct output expansion refers to gross output growth among rail related sectors, such as transport equipment and manufacturing and rail transport, due to the rise in capital input among these sectors. Conversely, an indirect output expansion refers to a growth of output among rail related sectors, primarily those tertiary sectors such as tourism sector. One major difference between these two effects is that the former should be considered as a shortterm effect as the regional economic impacts are generally achieved during infrastructure construction period, whereas the impacts of the latter effect cannot be materialized in a relatively longer-term after the completion of the infrastructure system.

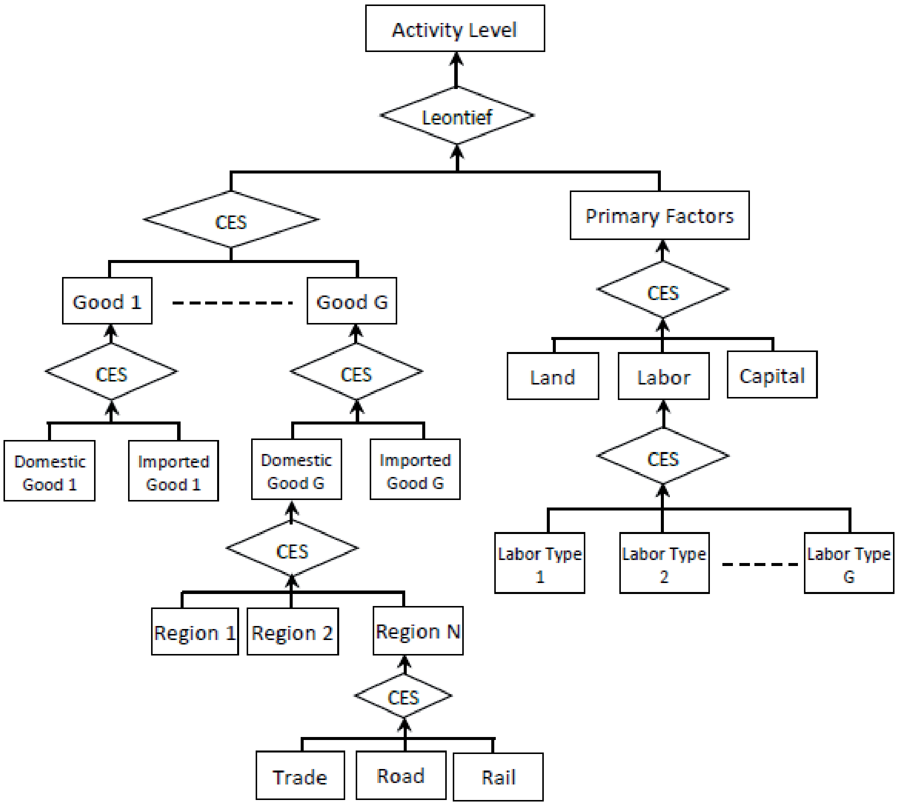
The effect of cost reduction is self-explained. It is generally modeled through a transportation margin shock in CGE analysis and it should only be considered as a long-term effect as the benefits such as transport cost reduction and travel time savings cannot be achieved until the operation of the new developed rail system. Similarly, both productivity effects are considered as indirect and long-term economic benefits of rail infrastructure development because these effects are only enable after a new system being deployed.

The third category of effect is derived from the rail transport demand change as a result of the operation of a new rail infrastructure system. Such an effect can be modeled in two aspects. First, the deployment of a new HSR line, for instance, may lead to a demand change due to substitution among different transportation modes. Although CGE model generally captures inherent substitution through nesting structures of production activities, the adaptive substitution as a response to an introduction of a new transport service has to be modeled explicitly through adjusting relevant elasticity parameters. The second aspect of demand change comes from the induced demand. It is modeled through a sectoral output shock instead as the effect is generated from rail transport sector itself. Understandably, the demand effect is considered to have indirect and long-term impacts on regional economic growth since such an effect is only available after the infrastructure reaches an operating stage.

4. A Dynamic SCGE Model Approach

The dynamic SCGE model adopted in this study is called dynamic TERM, which stands for Dynamic The Enormous Regional Model. The model is an essence of the Centre of Policy Studies (CoPS) at the University of Victoria in Australia and it has been updated by several leading CGE modelers such as Mark Horridge and Glyn Wittwer. The model has several unique features for a large-scale multi-regional CGE assessment. For instance, the model has a capacity to achieve a robust measurement of regional economic impacts given that it is a bottom-up model in which each region is treated as a separate economy. Such a modeling structure is able to provide a high degree of regional details, which can make the model to examine the regional impacts of shocks that may be region-specific. In addition, the model also has a detailed treatment of transport costs, which can help users better simulate the effects of transportation infrastructure improvement.

The original TERM is a comparative static model and it was further developed into various versions for over 13 different countries. The Chinese version of TERM is called SinoTERM, which is a static model covering 31 provinces and municipalities (Horridge and Wittwer, 2008). The model follows the standard CGE structure, which includes equations systems representing the linkages and interactions for four types of economic activities: production, household consumption, government and trade.



Source: Author's update based on Horridge (2013).

Figure 2. Production Nesting Structure of TERM

Specifically, the model assumes that each economic sector produces one commodity each, maximizing profits through a nested production structure with both intermediate goods and primary factors through a Leontief function at the top nest, as illustrated in Figure 2. On the right-hand side, primary goods are produced from three types of factor inputs, including land, capital and labor, the latter one of which is derived from a third-level CES nesting called skill nest, representing a substitution of different types of labors. On the left-hand side, the intermediate goods are produced through various goods under a Constant Elasticity of Substitution (CES) production function, each commodity is further derived from a composite with both domestically produced good and imported good through a third-level CES function (also known as Armington nest). In the fourth-level nest, a domestically produced commodity can be decomposed by different origins of production through a CES function, which essentially reflects inter-regional trade interactions. In the fifth-level, various margin costs are then added to any specific commodity through a Leontief production function. In the sixth-level, the source of each margin is aggregated through a CES function.

TERM assumes household maximize utility, assuming a Klein-Rubin functional form, which is a non-homothetic utility form and subject to budget constraint[1]. The model does not distinguish regional and national government, but government activity functions include government taxes, government income and expenditures. TERM considers two types of taxes, including commodity tax and production tax. Dynamic features of the TERM follow the structures of ORANIG-RD single region model, which includes equations representing rules for capital accumulation, investment and wage adjustments (Horridge, 2012). Specifically, capital accumulation can be represented as:

$$K_{i,r,t+1} = K_{i,r,t}\left(1 - D_{i,t}\right) + I_{i,r,t} \qquad (1)$$

where $K_{i,t}$ denotes the quantity of capital stock available to sector i in region r in year t, $I_{i,t}$ represents the quantity of investment in sector i in region r in year t and $D_{i,t}$ represents the rate of depreciation. The base year quantity of capital stock is provided exogenously, whereas the level of investment is determined by the expected rate of return in sector i in region r in a given time period. Horridge (2012) indicates that the investment mechanism in dynamic TERM involves two basic assumptions: 1) investment/capital ratios are positively related to expected rates of return and 2), expected rates of return converge to actual rates of return via a partial adjustment mechanism. The two assumptions are represented in equations 2 and 3, respectively:

$$G = F(E) \qquad (2)$$

$$G = Q \cdot G_{trend} \cdot \frac{M^{\alpha}}{Q-1+M^{\alpha}} \qquad (3)$$

where G denotes gross rate of capital growth in the next period and E denotes expected gross rate of return in the next period; M represents the ratio between the expected gross rates of return E and normal gross rates of return R_{normal}; Q denotes (max/trend) investment/capital ratio, and G_{trend} is represented as a function of R_{normal}. Implementation of the first equation assumes that each sector has a long-run or normal rate of return and requires an exogenously determined expected gross rate of return, whereas calibration of the second equations requires to specific parameters, such as investment elasticities a, investment/capital ratio G and normal gross rate of return R_{normal}, all of which need to be provided exogenously.

---

[1] *Non-homothetic means that rising income causes budget shares to change even with price ratio fixed.*

Wage adjustment equation in dynamic TERM assumes that wages rise if the actual employment is above the trend (predicted) employment (Wittwer et al. 2005). Since employment is negatively related to real wages, a convergence between the actual employment and the trend employment are expected to occur when the economic system reaches a long-term market clearance. The relationship between wage and employment can be expressed as:

$$\frac{\Delta W_t}{W_0} = \gamma\left[\left(\frac{L_0}{T_0}\right) - 1\right] + \gamma\Delta\left(\frac{L}{T}\right) \tag{4}$$

where W represents real wage, L and T represent actual employment and trend employment, respectively. γ denotes a positive parameter to reflect the speed of labor market adjustment.

5. Data

One of the major challenges for a comprehensive economic impacts assessment of the rail infrastructure development in China is data limitation. This is particularly true if the assessment is conducted at the regional level. As revealed in Chen et al. (2016), the assessment using a CGE analysis requires two types of data: one represents the direct impact drivers, which are used to calculate the magnitude of policy shock for CGE simulations. The other one is often referred to as social accounting matrix, which serves as the benchmark data for CGE calibration. This section discusses the data requirement for a regional economic impacts assessment of rail infrastructure. Our focus is on data that represents direct impact drivers which is often ignored in previous studies[2]. These data reflects land use change, the levels of capital investment, change in transportation cost and productivity, all of which were directly driven by the development of rail infrastructures.

- *A. Land Use Change*

Rail infrastructure development has two types of effect on land use change. One is a direct effect which is due to an immediate land use for the development of rail facilities, such as station, routes and maintenance centers. In addition, the development of rail system may also lead to an extended land use effect due to its stimulus to urbanization, which is manifested by a prosperity of real estate sectors and the development of HSR new towns. All these effects are expected to have a negative effect on agricultural related sectors due to occupation of arable land. Since land use data of rail construction at the regional level is not publicly available, one alternative is to estimate the area size of land use for rail infrastructure systems based on the following equation:

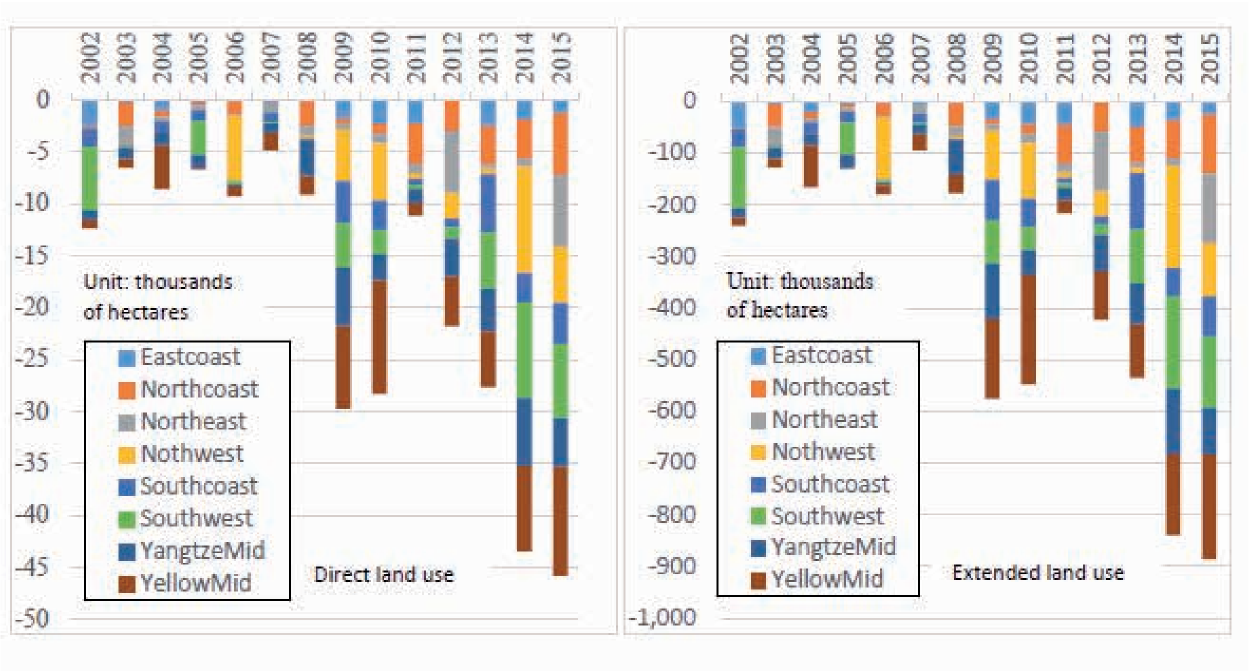
$$Direct\_Land\_Use_{r,t} = \frac{WB_{r,t} - \Delta Track_{r,t} \cdot M}{WB_{r,t}} \times 100\% \tag{5}$$

---

[2] *Plethora of studies using CGE for an impact assessment of transportation infrastructure system was based on hypothetical scenarios, hence the levels of direct impact drivers were generally specified based on arbitrary assumptions, which has led to a lack of underpinnings of policy shocks (Chen and Haynes, 2017).*

where $WB_{r,t}$ and $\Delta Track_{r,t}$ denote the arable land area and new added rail track length in region r in year t, whereas M represents the additional area size required for developing one km of rail track. Following Chen, et al. (2016), the value of 5 hectares of land/km is adopted in this calculation. One should note that the aforementioned calculation is derived upon two assumptions: First, the land use for new rail line construction is solely converted from arable land. Second, the land use efficiency for rail infrastructure development is assumed to be consistent across different regions.

The extended land use effect as a result of rail development can be estimated using the similar approach adopted in Chen et al. (2016). The method assumes that urban land use due to a new rail infrastructure development can be estimated if the linear relationship between them is understood. The estimated results of arable land area change due to HSR Development in China is illustrated in Figure 3, which show two clear patterns: First, the level of arable land reduction caused by the extended land use is more substantial than the direct land use. Second, the level change of the arable land varies significantly both temporally and spatially.

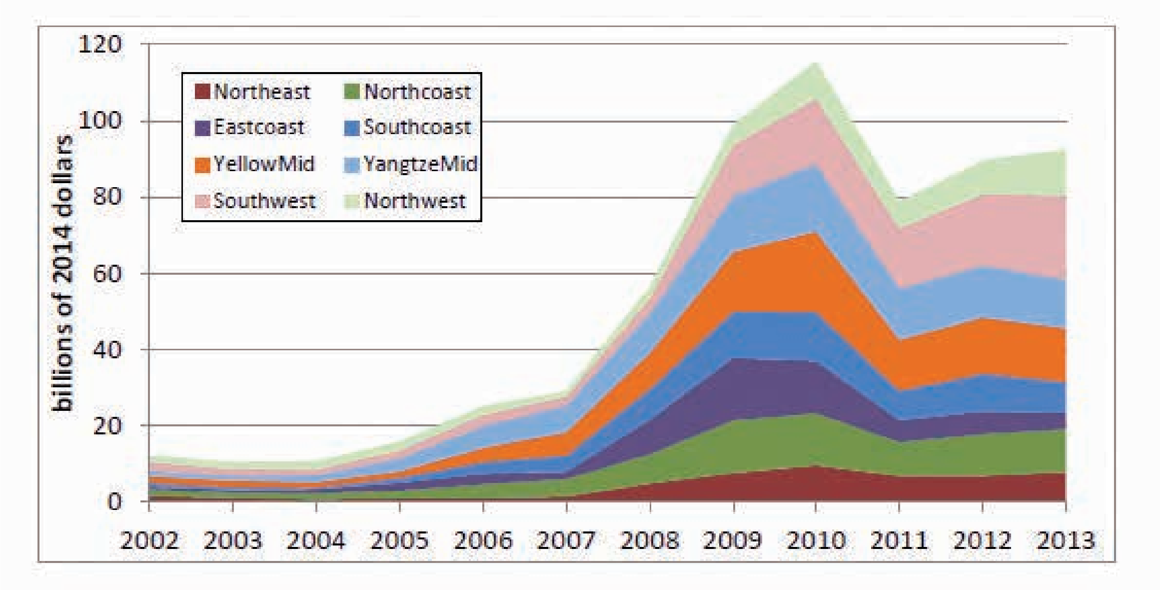


Source: National Statistics Bureau of China.

Figure 3. Estimated Arable Land Area Reduction due to HSR Development in China

- ### *B. Capital Investment*

Capital investment is one of the major drivers for regional economic growth, hence a detailed data source that reflects the regional rail capital investment pattern is essential for a valid regional economic impact assessment. The data of capital investment in rail infrastructure development for the period 2002-2013 is obtained from the Compilation of Railway Statistics. In particular, the data includes capital expenditure in four major fields: rail route construction, facility construction, procurement of rail equipment, such as rolling stock and EMU and the upgrade of existing infrastructures. As illustrated in Figure 4, rail capital investment has experienced a substantial growth since 2007 and reached a peak in 2010, which was then followed by a decline. The investment was dominated in regions such as the Yellow River MidReaches and the Yangtze River Mid-Reaches. The capital investment in rail infrastructure is expected to generate a different regional economic impacts primarily in a short-run through a boost in capital factor input to the economic system. From a modeling perspective, the detailed capital investment data in different fields and regions will be converted into a percent change in K for their corresponding sectors, which will then be used to estimate the indirect economic impacts through CGE simulation.

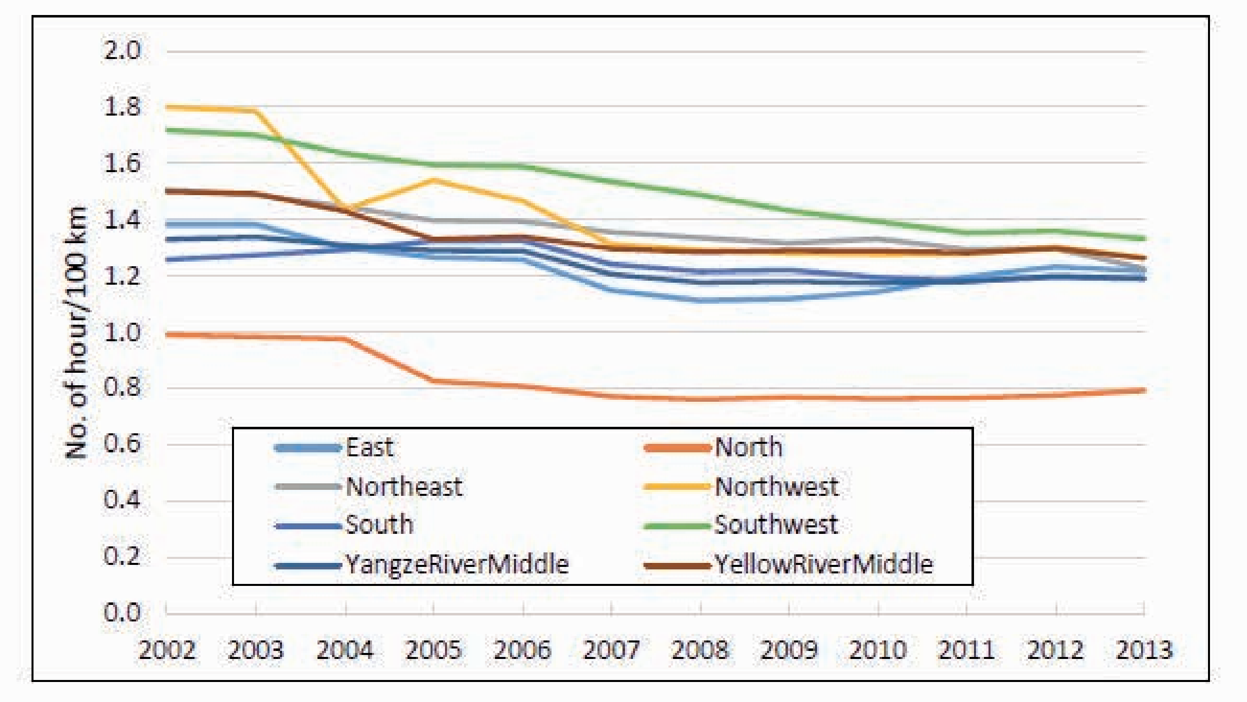


Source: The Compilation of Railway Statistics, 2003-2014

Figure 4. Rail Infrastructure Investment by Regions in China: 2002-2013

- ***C. Rail Transportation Cost***

Transportation cost change is considered as the third key drivers to measure the economic impacts of rail infrastructure improvements. A reduction of transportation cost as a result of infrastructure system development is expected to improve economic efficiency and facilitate the expansion of final demand and supply, which then may lead to a growth of the economy. However, the measurement of the generalized cost can be very challenging as it involves both monetary costs and time costs (Button, 2010). Since most of the data are not available, following Chen et al. (2016), we use the technological speed as a proxy to measure the change of rail transportation cost.



Source: The Compilation of Railway Statistics, 2003-2014

Figure 5. Rail Transportation Cost by Regions in China: 2002-2013

Given the focus of our assessment is HSR, which is essentially a passenger rail system, the average technological speed of passenger rail of different regional rail bureaus was adopted as the proxy to calculate travel time change in different years[3]. As illustrated in Figure 5, the travel time costs measured by number of hours need per 100 kilometers generally decline during 2002 and 2013, which could be considered as the outcome driven by an improvement in rail infrastructure.

One should also note that such a calculation has three limitations: First, given the specific focus of our research objective and the data availability, the speed change of freight rail and monetary travel cost are not considered. Second, we do not differentiate whether the speed is due to a hardware improvement (e.g. infrastructure improvement) or a software adjustment (e.g. a regulatory adjustment due to a concern on safety, the advancement in rail operation and management, and etc.). Nevertheless, all these issues need to be further addressed in the future once such data becomes available.

- ### *D. Productivity Change*

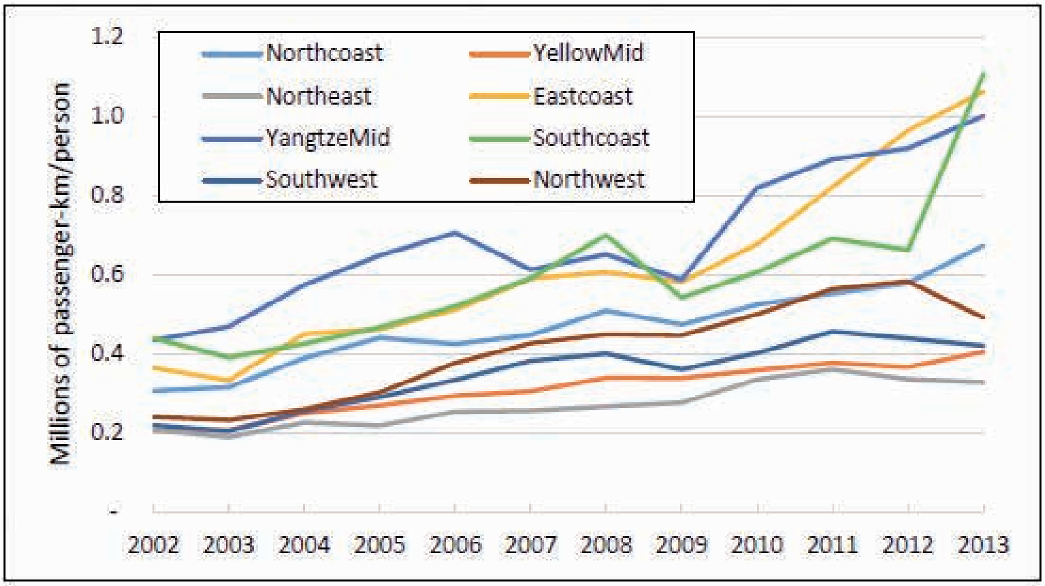
The massive rail investment in China, particularly in developing HSR, is likely to improve the overall productivity of passenger rail system given the adoption of various advanced HSR technologies. Since a productivity shock in CGE model indicates a technology improvement of

---

[3] Although the Chinese railways are owned and managed by the National Railway Corporation (the formal Ministry of Railways), the operation and maintenance are managed by 18 regional railway entities, including Harbin Railway Bureau, Shenyang Railway Bureau, Beijing Railway Bureau, Taiyuan Railway Bureau, Hohhot Railway Bureau, Zhengzhou Railway Bureau, Wuhan Railway Bureau, Xi ' an Railway Bureau, Jinan Railway Bureau, Shanghai Railway Bureau, Nanchang Railway Bureau, Guangzhou Railway (group) Company, Nanning Railway Bureau, Chengdu Railway Bureau, Kunming Railway Bureau, Lanzhou Railway Bureau, and Urumqi Railway Bureau.

production activity, following Chen et al., (2016)'s approach, labor productivity of passenger rail system is adopted as a proxy to measure the productivity change of rail sector. Essentially, as denoted in Equation 6, labor productivity (P) is a ratio which is derived from using the passengerkm (PKM) divided by the number of employees in each region (r):

$$P_{t,r} = \frac{\frac{PKM_{t,r}}{Employee_{t,r}} - 1}{\frac{PKM_{t-1,r}}{Employee_{t-1,r}}} \times 100\% \qquad (6)$$



Source: The Compilation of Railway Statistics, 2003-2014

Figure 6. The Average Labor Output Change of Passenger Rail in China: 2002-2013

Figure 6 illustrates the change of labor output (PKM/Employee) by different regions for the period 2002-2013. The general trend of the average labor output is growing, which suggests that the productivity of passenger rail has been improved since the massive development of HSR. The performance in some regions, such as the Yangtze River Mid-Reaches and the South coast, experienced some fluctuations during 2006-2012. This is primarily due to the expansion of labor force in rail transportation sector due to the opening of several main HSR services, such as the Wuhan-Guangzhou HSR.

- ***E. Data for SCGE Modelling***

The benchmark data used for SCGE modeling is based on the SinoTERM database (Horridge and Wittwer, 2008), which contains the national input-output or use-supply table of China in 2002 as well as regional data used for the estimates of regional distribution of output and final demand. A detailed TERM database structure and development process could be found in Horridge

(2012). One of the advantages of using TERM to evaluate regional impacts of HSR in China is that the model contains a detailed breakdown of trade margin by different transportation modes. As indicated in Horridge et al. (2005), TERM assumes that all users in each region consumes commodities from other regions according to common proportions[4].

Specifically, the value of follows is represented in three respects:

A. Basic values = Output price (for domestically produced goods), or CIF prices (for imports);
B. Delivered values = Basic values + (transport or retail) Margins;
C. Purchasers' values = Delivered values + Tax.

Such a detailed structure enables us to simulate the indirect economic impacts of rail infrastructure development through the shocks on rail transportation cost change. The original database includes 137 sectors and 31 regions. To facilitate CGE simulations, we adopted a condensed version of the SinoTERM database for the assessment, which includes 47 sectors and 8 regions. A detailed bridging table of the sectors and regions are shown in appendix I.

In addition, we made the following updates to improve the accuracy of simulation. First, since the original SinoTERM is a static model, we upgraded the model into a dynamic model by specifying its investment/capital ratio as 25 based on findings of Bai et al. (2006). Second, the elasticity of substitution for factor inputs were updated based on Guo et al. (2014) and Zha and Zhou (2014).

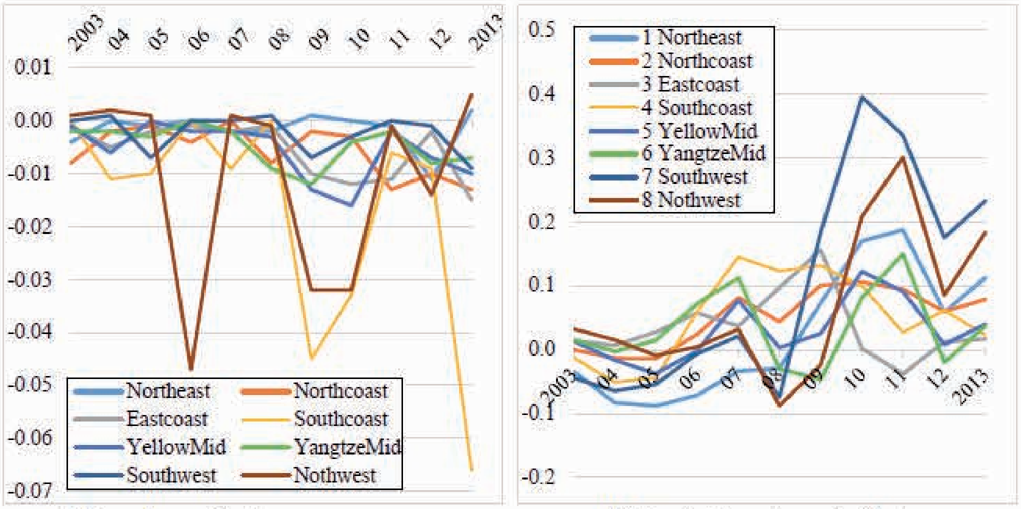
6. Results

The CGE simulations were implemented in five groups in order to measure the regional economic impacts of five scenarios, including land use effect, capital investment effect, transportation cost change effect, productivity change effect and a simultaneous effect. The model was operated using RunDynam, a windows interface developed by the GEMPACK Software team at the University of Victoria in Melbourne. Given that the direct impact drivers capture the period 2002 – 2013, the model was solved recursive-dynamically, in other words, the results are computed one-period-at-a-time. A short-run closure rule was applied for the simulation in order to allows wage to be fixed while employment to be adjustable endogenously. In addition, an additional rule was applied to exogenize the investment variable when the simulation involves a capital investment shock in order to achieve a convergent optimized solution.

The simulation results of regional economic impacts measured by the change of regional gross product (GRP) are illustrated in Figure 7. Specifically, Figure 7(a) illustrates the impacts of land use change as a result of rail infrastructure development on GRP. It is clear that the land use effect has a negative impact on GRP growth due to the constraint of land factor for agricultural related sectors. The magnitude of impacts varies substantially across different regions. For instance, the negative impacts of land use from rail infrastructure development in the northwest and south-coast regions are found to be relatively higher than other regions. The two spikes of negative impacts occurred in 2006 in the northwest region and 2013 in the south-coast region were primarily due to the extended effect of urbanization as a result of rail infrastructure development.

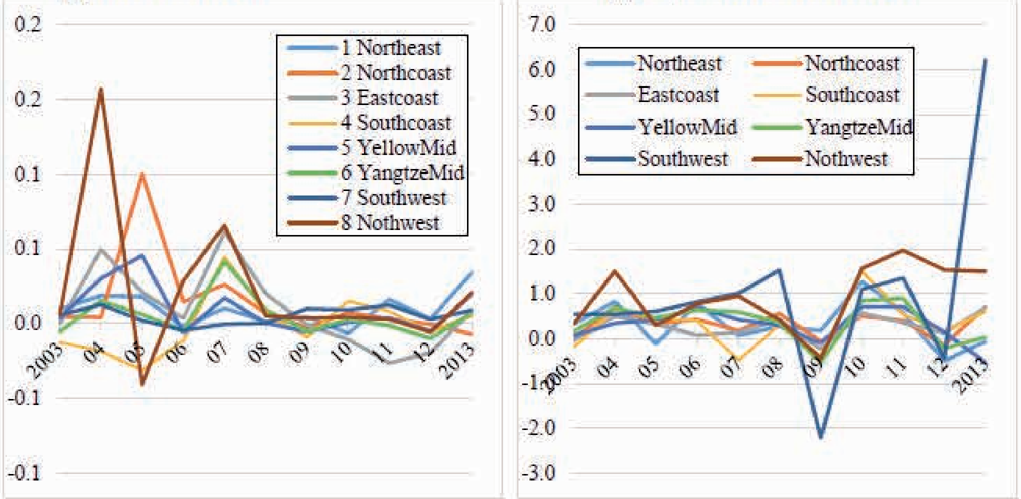
---

[4] Unlike the conventional data structure for a single-region CGE model, inter-regional trade flows are captured in a TRADE matrix, which serves as a part of TERM's data structure. TRADE contains a n X n submatrix, where n represents the number of regions in the model. Each row corresponds to region of origin and each column corresponds to region of use (destination). Locally consumed commodities are denoted as diagonal elements.



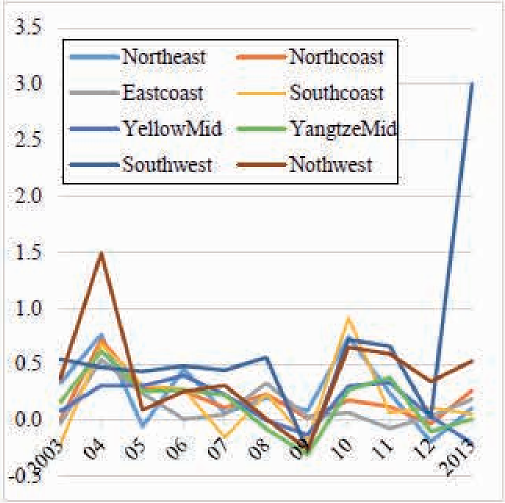
(a) Land use effect

(b) Capital investment effect



(c) Transportation cost change effect

(d) Productivity change effect



(e) Simultaneous effects

Figure 7. Regional Economic Impacts of Rail Development based on Various Effects

Figure 7(b) presents the regional impacts of the capital investment effect as a result of rail infrastructure development in China. The general trend of impact is increasing during 20022013, but there are also two major drops in the middle of the period. The decline in 2008 may be explained by the investment cut in that year due to the economic recession, whereas the fall starting in 2011 is likely due to the disinvestment as a result of the HSR accident in 2011[5]. In terms of the regional differences of impact, regions in the less developed southwest and northwest experienced relatively larger growth during the period given the stimulus from capital investment in rail sector.

The regional impacts of the transportation cost change effect as a result of rail infrastructure development were illustrated in Figure 7(c). Although there were some fluctuations in the initial development period, the impacts on GRP are modest with a minor increasing trend over time. The fluctuation of contribution in during 2003 – 2008 may be caused by the change of the regulatory policies on rail operating speed management, whereas modest increasing trend after 2008 may reflect the fact that transportation costs did not experience a substantial change during this period of time as most rail systems were still under construction.

The regional impacts of productivity change effect from rail infrastructure development are illustrated in Figure 7(d). Generally speaking, a productivity increase in rail transport sector is associated with around a positive impact on GRP during the investigation period with an average magnitude at around 0.7 percent. However, there were also two major declines in 2009 and 2012, which is likely due to the following two reasons: First, since the productivity variable is essentially a labor productivity which reflects a ratio changes in both passenger-km (PKM) and the number of employees, the major decline of impact in 2009 is most likely caused by the drop of passenger rail demand due to the effect of recession, whereas the decline in 2011 is likely to be caused by the increase in the number of rail sector related jobs due to the openings of new HSR services. Last but not the least, the results of the simultaneous simulation that incorporated the four effects are illustrated in Figure 7(e). It is clear that although the economic contributions from rail infrastructure development tend to decline given the influences from economic recession, the overall economic impacts are positive.

Table 4. Aggregate Regional Economic Impacts of Rail Infrastructure: 2002-2013

Region	Nominal Output	Real Output[a]	Real GDP[a]	Agg Employ[b]	Nominal Output	Real Output	Real GDP	Agg Employ	Output multiplier (nominal)	Output multiplier (real)	GDP multiplier (real)
	Level Change				Percent Change						
Northeast	151.36	60.81	6.66	0.11	61.29	23.43	3.94	0.21	3.14	1.26	0.14
Northcoast	172.77	77.45	8.19	0.41	32.55	14.02	4.57	0.38	2.16	0.97	0.10
Eastcoast	295.38	132.06	8.92	-0.04	46.84	19.87	3.07	-0.04	4.67	2.09	0.14
Southcoast	246.28	108.26	11.14	0.11	58.58	24.85	4.39	0.15	3.56	1.57	0.16
YellowMid	251.26	142.03	7.17	0.00	86.27	46.55	3.55	0.00	2.44	1.38	0.07
YangtzeMid	128.75	73.40	3.51	0.14	37.35	20.48	3.32	0.10	1.27	0.72	0.03
Southwest	40.50	20.52	5.61	2.67	13.78	6.85	14.42	1.77	0.40	0.20	0.06
Northwest	22.86	12.42	1.96	0.69	27.45	14.81	9.63	2.28	0.41	0.22	0.04
National	1309.16	626.94	53.16	4.08	46.19	21.19	4.22	0.55	2.10	1.01	0.09

a. Billions of 2014 dollars.

b. Millions of jobs.

c. The results reflect a simultaneous effects of land use, capital investment, changes of transport cost and change of productivity.

---

[5] A HSR accident occurred on July 23 2011 caused 40 deaths and 172 injuries, which was later identified caused by equipment defects under an extreme weather condition. As a result, the pace of the massive rail development was slowed down given the safety concern (Chen and Haynes, 2015).

The aggregate regional impacts of rail infrastructure development in China for the period 2002 – 2013 are summarized in Table 4. The real GDP impacts were found to be the largest in the north-coast region, whereas the smallest one was in the northwest. The impacts on the aggregated employment are similar as most of the jobs were added in the north-coast regions as a result of rail infrastructure development, but southwest has the second highest number of jobs being created due to rail development. The regional economic contributions of rail development during 2002 and 2013 were found to be the highest in the north-coast region if measured in GDP multiplier. The overall multipliers for gross output (real terms) and real GDP are 1.01 and 0.09, which suggests that a one-dollar investment in rail sector is likely to generate one-dollar increase in gross output and 0.09 dollar increase in real GDP.

7. Discussions

China has built the largest HSR system in the world with the strong support from its central government. While more people began enjoying the convenience of intercity travel since the opening of numerous HSR services, the understanding of its regional economic impacts remains unclear. This study introduces for the first time, a comprehensive modeling framework to evaluate the long-term regional economic impacts of rail infrastructure development in China. By applying the state-of-the-art approach to economic impact assessment using a dynamic SCGE model, we developed a detailed modeling procedure to reflect both the short-run effect from rail investment and the long-run effect from the operations of new HSR services. Such a modeling procedure is expected to provide a more reliable estimate than the traditional approach that often evaluated from an ex ante perspective.

After incorporating the four types of effects including land use, capital investment, change of transportation cost and productivity into the modeling framework, the results indicate that rail infrastructure development in China, which is dominated by HSR investment, demonstrates a

positive long-term impacts on regional economic growth with a gross output multiplier of 1.01 and a GDP multiplier of 0.09. The aggregate impacts were found to be much significant in the in the southwest region, whereas the impacts are relatively small in developed eastern regions.

One should note that the aforementioned empirical results are preliminary in the sense that they only reflect the feasibility of the modeling framework for the evaluation of the long-term regional economic impacts of HSR development. Hence, the assessment outcomes should be read with caution. Limitations still need to be clarified so that further endeavors can be made to improve the assessment outcomes. The first limitation is that since the detailed regional level data that reflect the change of inter-regional transport cost and productivity is not available, the existing empirical assessment doesn’t fully capture the regional impacts that caused by other factors, such as a reduction of interregional transportation cost and a productivity increase brought by HSR. Similarly, due to the lack of travel demand statistics at the regional level, the results are also limited as the induced demand effect and effect of the substitution among different transportation modes ignored.

Second, some of the direct impact drivers for CGE simulation need to be further improved. For instance, the productivity change as an outcome of passenger rail system improvement was currently measured in labor productivity. Although such a consideration captures the dynamics of operational efficiency in rail sector, the indicator also has a limitation in that it inevitably included other factors, such as influences from economic performance, regulatory changes and etc. This also explains the negative consequence of a productivity decrease on the regional economy. Hence, in order to reflect the trend of productivity change as a response to the infrastructure and technology improvement, these aforementioned disturbing factors should be removed from the existing indicators or better indicator should be considered.

Third, some key parameters of the SCGE modeling system, such as the elasticity of substitution for factor inputs, the Armington elasticities, remains limited which need to be further validated and updated. For instance, early studies have suggested that the results of CGE can be biased unless key parameters were carefully estimated and chosen based on the specific regional focus of assessment (Partridge and Rickman, 1998; Chen and Haynes, 2017). Hence, in order to achieve a more accurate long-term regional economic impact assessment of the rail infrastructure system, more endeavors are still needed in terms of both data collection and parameter calibration.

Nevertheless, our study still has implications for infrastructure planning and policy, at least in the following two aspects. First, a closer collaboration among different entities, such as government, private sectors, and academic scholars, is essential to achieve a more reliable regional economic impact assessments of large infrastructure system, such as HSR. This is particularly important and relevant in countries like China as information is often limited to certain agencies which as a result, regional impact assessment of HSR can be very challenging. Second, our preliminary results imply that given that the economic impacts of the HSR systems tend to be dissimilar among different regions, future infrastructure development and investment plans need to be more cautiously implemented so as to a maximum benefit to the society and the economy as well as a maximum return to investment.

8. References

- Adams, P. D., Dixon, P. B., McDonald, D., Meagher, G. A., & Parmenter, B. R. (1994). Forecasts for the Australian economy using the MONASH model. International Journal of forecasting, 10(4), 557-571.
- Ansar, A., Flyvbjerg, B., Budzier, A., & Lunn, D. (2016). Does infrastructure investment lead to economic growth or economic fragility? Evidence from China. Oxford Review of Economic Policy, 32(3), 360-390.
- Bai, C., Hsieh, C. & Qian, Y. (2006). The Return to Capital in China, Brookings Papers on Economic Activity, Economic Studies Program, The Brookings Institution, vol. 37, 61-102.
- Brand, D., Kiefer, M., Parody, T.ehndiratta, S. (2001). Application of benefit-cost analysis to the proposed California high-speed rail system. Transportation Research Record: Journal of the Transportation Research Board, (1742), 9-16.
- Bröcker, J. (1998). How would an EU-membership of the Visegrád-countries affect Europe‘s economic geography? EU-membership of the Visegrád-countries. The Annals of Regional Science, 32(1), 91-114.
- Button, K. (2017). High-Speed Railways: Do They Produce Economic Growth?. Mercatus Center Research Paper. George Mason University.
- Button, K. (2010). Transport Economics. Edward Elgar Publishing.
- Chen, Z., & Haynes, K. E. (2017). Transportation Infrastructure and Economic Growth in China: A Meta-analysis. In Socioeconomic Environmental Policies and Evaluations in Regional Science (pp. 339-357). Springer Singapore.
- Chen, Z., & Haynes, K. E. (Eds.). (2015). Chinese Railways in the Era of High-speed. Emerald Group Publishing Limited.
- Chen, Z., & Haynes, K. E. (2013). Transportation capital in the United States: a multimodal general equilibrium analysis. Public Works Management & Policy, 19(2), 97-117.
- Chen, Z., Xue, J., Rose, A. Z., & Haynes, K. E. (2016). The impact of high-speed rail investment on economic and environmental change in China: A dynamic CGE analysis. Transportation Research Part A: Policy and Practice, 92, 232-245.

- Conrad, K. (1997). Traffic, Transportation, Infrastructure and Externalities A Theoretical Framework for a CGE Analysis. The Annals of Regional Science, 31(4), 369-389.
- De Rus, G. (2011). The BCA of HSR: Should the Government Invest in High Speed Rail Infrastructure?. Journal of Benefit-Cost Analysis.
- De Rus, G., & Nombela, G. (2007). Is investment in high speed rail socially profitable?. Journal of Transport Economics and Policy (JTEP), 41(1), 3-23.
- Givoni, M., & Banister, D. (2012). Speed: the less important element of the High-Speed Train. Journal of Transport Geography, 22, 306-307.
- Guo, Z., Zhang, X., Zheng, Y., & Rao, R. (2014). Exploring the impacts of a carbon tax on the Chinese economy using a CGE model with a detailed disaggregation of energy sectors. Energy Economics, 45, 455-462.
- Haddad, E. A., Hewings, G. J., Perobelli, F. S., & Santos, R. C. (2010). Regional Effects of Port Infrastructure: a Spatial CGE Application to Brazil. International Regional Science Review, 33(3), 239-263.
- Horridge, M. (2012). The TERM model and its database. In Economic Modeling of Water (pp. 1335). Springer Netherlands.
- Horridge, M., & Wittwer, G. (2008). SinoTERM, a multi-regional CGE model of China. China Economic Review, 19(4), 628-634.
- Horridge, M., Madden, J., & Wittwer, G. (2005). The impact of the 2002–2003 drought on Australia. Journal of Policy Modeling, 27(3), 285-308.
- Janic, M. (2003). Multicriteria evaluation of high-speed rail, transrapid maglev and air passenger transport in Europe. Transportation Planning and Technology, 26(6), 491-512.
- Kim, E., Hewings, G. J., & Hong, C. (2004). An Application of an Integrated Transport Network– Multiregional CGE Model: a Framework for the Economic Analysis of Highway Projects. Economic Systems Research, 16(3), 235-258.
- Kim, E. & Kim, K. (2002) Impacts of Regional Development Strategies on Growth and Equity of Korea: A Multiregional CGE Model, Annals of Regional Science, 36(1), 165-189.
- Knaap, T., & Oosterhaven, J. (2002). The welfare effects of new infrastructure: An economic geography approach to evaluating new Dutch railway links, Mimeo.
- Koike, A., Ishikura, T., Miyashita, M. & Tsuchiya, K. (2015). Spatial economic analysis for intercity transport policies. In Intercity Transport and Climate Change (pp. 177-213). Springer International Publishing.
- Lofgren, H., Harris, R. L., & Robinson, S. (2002). A Standard Computable General Equilibrium (CGE) Model in GAMS (Vol. 5). International Food Policy Research Institute, Washington, D.C. Melo, P. C., Graham, D. J., & Brage-Ardao, R. (2013). The productivity of transport infrastructure investment: A meta-analysis of empirical evidence. Regional Science and Urban Economics, 43(5), 695-706.
- Partridge, M. D., & Rickman, D. S. (1998). Regional computable general equilibrium modeling: a survey and critical appraisal. International Regional Science Review, 21(3), 205-248.
- Rioja, F. K. (1999). Productiveness and Welfare Implications of Public Infrastructure: a Dynamic Two-Sector General Equilibrium Analysis. Journal of Development Economics, 58(2), 387-404.
- Seung, C. K. & Kraybill, D. S. (2001). The Effects of Infrastructure Investment: a Two-Sector Dynamic Computable General Equilibrium Analysis for Ohio, International Regional Science Review, 24(2), 261-281.

- Venables, A. J. (2016). Incorporating wider economic impacts within cost-benefit appraisal. OECD Discussion paper for International Transport Forum. Paris. 2016-05.
- Vickerman, R. (2007). Cost—Benefit analysis and large-scale infrastructure projects: State of the art and challenges. Environment and Planning B: Planning and Design, 34(4), 598-610.
- Vickerman, R. (2017). Can high-speed rail have a transformative effect on the economy?. Transport Policy.
- Vold, A., & Jean-Hansen, V. (2007). PINGO-A model for prediction of regional and interregional freight transport in Norway. Institute of Transport Economics, TOI report, Vol. 899.
- Wittwer, G., Vere, D. T., Jones, R. E., & Griffith, G. R. (2005). Dynamic general equilibrium analysis of improved weed management in Australia's winter cropping systems. Australian journal of agricultural and resource economics, 49(4), 363-377.
- Wu, J., Nash, C., & Wang, D. (2014). Is high speed rail an appropriate solution to China's rail capacity problems?. Journal of Transport Geography, 40, 100-111.
- Zha, D., & Zhou, D. (2014). The elasticity of substitution and the way of nesting CES production function with emphasis on energy input. Applied Energy, 130, 793-798.
- Zhao, J., Zhao, Y., & Li, Y. (2015). The variation in the value of travel-time savings and the dilemma of high-speed rail in China. Transportation Research Part A: Policy and Practice, 82, 130-140.

Economic impacts

360.revista de alta velocidad
número 5 - junio - 2018. Pág 409 - 433

A first evaluation of the relationship between High Speed Rail (HSR) and the tourism sector in Turkey: The cases of two Turkish cities

Dalkic, Gulcin

Tuydes-Yaman, Hediye

Delaplace, Marie

Middle East Technical University[1]
Université de Paris-Est Marne la Vallée (UPEM)

Abstract

In Turkey, extensive investments have been made to develop HSR services since 2003. As a strategic and growing industry that relies on the increased mobility among cities and countries, tourism can benefit from these HSR services. But there are different types of tourism requiring different travel services. Due to this diversity, it is important to investigate the impact of the HSR lines on the tourism sector in specific case studies. In Turkey, the highest HSR ridership were observed in Ankara-Eskişehir and Ankara-Konya HSR corridors, each having more than 1.5 million HSR passengers annually. Konya is a major place of religious and cultural tourism attracting more than 2 million tourists every year, and Eskisehir is characterized by its identity of "University City", theme parks, cultural and archaeological places that attract mostly domestic tourism. On one hand, recent studies on the HSR impact on Eskişehir and Konya showed that the advantages provided by HSR caused an increase in the number of daily visitors and on the other hand, some studies stated that main trip purposes for HSR trips were tourism and business. However, the net impact of HSR on tourism in these two cities has not been investigated. This is the aim of this paper. First, the change in the HSR ridership for all the existing corridors will be analysed to understand the development of HSR use. Secondly, tourism potential of HSR cities will be evaluated based on available tourism statistics, socio-economic characteristics of the cities, major touristic destinations (in or within the vicinity of the city) and tourism types. Spatial distribution of the touristic destinations in Konya and Eskişehir will be evaluated in more detail with respect to accessibility from HSR station (intermodal connections, public transit availability, etc.) as case studies.

Keywords: high speed rail, tourism, Turkey, Eskişehir, Konya

---

[1] *Dalkic, Gulcin. Middle East Technical University, Civil Engineering Department. Email: dalkic@metu.edu.tr. (corresponding author)*
Tuydes-Yaman, Hediye. Middle East Technical University, Civil Engineering Department. Email: htuydes@metu.edu.tr
Delaplace, Marie. Université de Paris-Est Marne la Vallée (UPEM), Lab'Urba, Ecole d'Urbanisme de Paris (EUP). Email: marie.delaplace@u-pem.fr

1. Introduction

In Turkey, extensive investments have been made to develop the HSR services since the beginning of the 2000s not only to create a more sustainable transportation network, but also to integrate national railway network to Trans-European railway network (Babalik-Sutcliffe, 2007; Dalkic, 2014). Currently there are four HSR lines in operation which connect 7 cities (solid double lines in Figure 1). The existing HSR lines were also combined with the intercity bus and conventional rail services to reach five nearby cities. Three new HSR lines are under construction and 13 HSR projects are in the planning stage, which will serve 47 cities when realized.

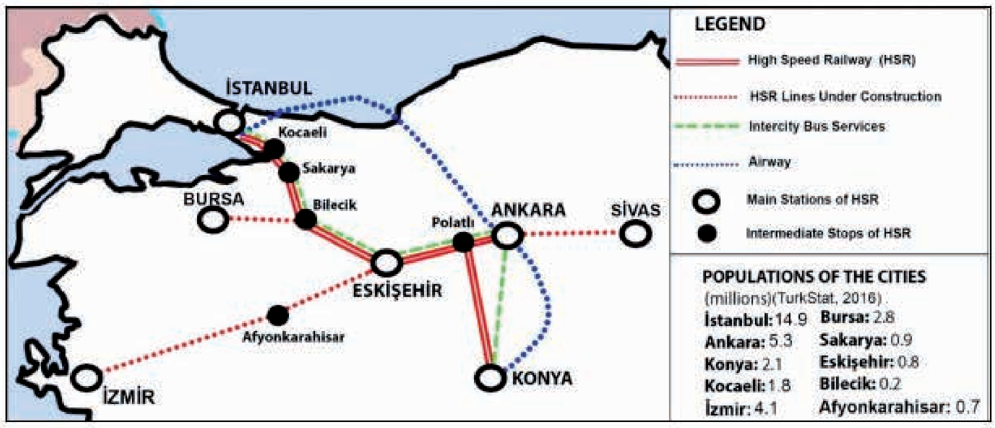


Figure 1. Intercity passenger alternatives along the current HSR network

Tourism (domestic and international) is a strategic and growing sector that can enrich the common culture, and create wealth. According to World Travel Tourism Council--WTTC (2017) Travel & Tourism’s contribution to world GDP was rising to a total of 10.2% of world GDP (US$7.6 trillion). Moreover, 292 million people that is 1 in 10 jobs are employed in this sector. Tourism is also a critical sector continuously included in development plans of Turkey for many decades. With the expansion of HSR network, economic, spatial and social impacts at a national level are expected. As a result of the increase in accessibility, some cities and regions will become more attractive; thus become important from tourism perspective.

As HSR is a newly introduced intercity passenger transportation alternative in Turkey, its impact on tourism has not been investigated much, yet. This study aims to provide a preliminary evaluation of the effects of the HSR on the tourism sector in two Turkish cities served by HSR first: Eskişehir and Konya. After a literature review about the relationship between HSR and tourism in Section 2, the evolution of the HSR ridership in Turkey is presented in Section 3, with some additional background on tourism section. Spatial distribution of the touristic destinations in Konya and Eskişehir will be evaluated in more detail in Section 4 with respect to accessibility from HSR station (intermodal connections, public transit availability, etc.) as case studies. The results of the study are expected to contribute to the existing literature on HSR and tourism and develop recommendations to increase benefits of HSR on tourism. Lastly, Section 5 is devoted to some conclusions.

2. Literature on HSR and tourism

Tourism-related service sectors play an important role in the regional development of the European Union countries. HSR service has a role in increase in tourist travel, and tourism and service sectors are therefore directly affected (Lumsdon and Page, 2004). Among different kinds, urban and business tourism and in-situ tourism (parks, etc.) appear to be the main types of tourism likely to benefit from an HSR service, when the station is located inside the destination. Compared to car journeys, traveling by HSR allows passengers to get to their destination faster while being able to relax and at the same time, avoid road congestion and the increasing difficulties associated with accessing the heart of the city. Compared to air travel, it also saves the time lost in travelling between the airport and final destination. In fact, all forms of forms of tourism in the city could benefit from HSR services; but, tourism forms outside cities (i.e. green tourism or mountain tourism) often benefit less from HSR, unless inter-modality issues are managed to reach their destinations without significant time losses.

Ex-post identified impacts of HSR on tourism can be positive as well as negative or inexistent (see Table 1 and Table 2). Positive impacts (expected or realized) can be listed as widening the tourism markets (increasing the number of tourists, tourism trips, occupancy rates, foreign arrivals, tourism revenues, etc.), improving the accessibility of the destinations (increasing the winter-sport tourism, urban tourism, business tourism). It can also reinforce the competition between tourist destinations (which can be positive or negative). For example, in Spain, Urena et al. (2009) argued that large intermediate cities served by HSR such as Zaragoza and in particular Córdoba were experiencing an increase in urban tourism and business tourism. This was confirmed by Alonso and Bellet (2009) in the case of Zaragoza. Similarly Todorovitch et al. (2011) reported that tourism had grown by 15% annually in Lleida. Guirao and Campa (2015) also reported how HSR was important for tourism in Toledo: over thirty percent of weekday HSR ridership was linked to tourism mobility. But a recent report in Spain suggested that *"the positive impacts of HSR on the number of visitors, the number of nights spent at destination and/or hotel occupancy rates were mostly restricted, at best, to larger cities, but in most cases the impacts are minimal or even negative"* (Albalate *et al.*, 2015).

An increase in tourist movements was mentioned in big cities in Taiwan (Cheng, 2009) as well in Chinese cities such as Wuhan (Wang et al., 2012), Qufu (People's Republic of China, 2014), and Ningbo (Zhao, 2012). Provinces served by HSR in China "were likely to have approximately 20% more foreign arrivals and 25% more in tourism revenue than provinces without such systems" (Chen and Haynes, 2012). Seeking theoretical foundations of the main impacts of HSR on China tourism, Wang et al. (2012) are using the gravitational and the iso-tourist line models with integrating the time–space replacement mechanism. They identified three kinds of impacts: i) an enlargement and a transformation of tourism market space, ii) intensification of the market competition on a larger scale, and iii) a redistribution of tourism centres.

On the other hand, HSR may have negative impacts (expected or realized) such as decrease in overnights stay and tourist revenues. In France there is no growth in small cities and for winter-sports tourism. A reduction in the average length of stay and in the number of nights spent in a city is likely to occur because HSR opens up the possibility of same-day round trips. As an example, in Le Mans in France, the arrival of HSR has contributed to reducing the duration of events and, conversely, to promoting non-residential events, i.e. fewer events lasting several days but more one-day events. In this city, the average length of stay decreased a few years after the arrival of HSR. Furthermore, events had an average duration one day shorter than that generally encountered for national conferences in France in the mid-1990s (Amiard, 1997).

Table 1. A review of the studies showing positive impacts of HSR on tourism

Impacts	Studies (Realized; Expected)
Widening of tourist markets	Sands, 1993 ; Urena et al., 2009; Chen & Haynes, 2015 ; Wang et al., 2012 ; Masson & Petiot, 2009
Increase in the number of tourists	Mannone, 1995 ; Sands, 1993 ; Buttet et al., 2001 ; CRT PACA, 2004, Chen & Haynes, 2014, Bazin et al., 2010, 2011, 2013a, 2014, CSEF, 2005; People's Republic of China, 2014; Okabe, 1980, Kurihara and Wu, 2015, Urena *et al*. 2009,
Increase of tourism trip	Cheng, 2009, People's Republic of China, 2014; Zhao, 2012
Increase of foreign arrivals	Chen & Haynes, 2012, 2014
Increase in occupancy rates	Sands, 1993 ; Mannone, 1995
Increase of the number of hotel rooms	Okabe, 1980, Mizohata, 1995, Vickerman and Ulied, 2006
Growth in the number of nights	Tourisme- Alsace, 2009
Limited impacts to larger cities	Albalate *et al*., 2015
Increase in the number of nights in the short run	INSEE Lorraine 2009
Improvement of the accessibility of the destination	Chen & Haynes, 2014; Coronado et al. , 2013 ; Wang *et al*., 2012; Masson & Petiot, 2009
Increase in the winter-sports tourism	Mizohata, 1995
Impact on hinterland areas	Okabe, 1980
Increase the probability to come again (to be a repeater)	Delaplace *et al*., 2014
Urban tourism development	CRCT PACA, 2003; Bazinetal., 2010, 2011, 2014; Delaplace&Perrin, 2013, Urenaet al., 2009; Delaplace and Benoit-Bazin, 2017; Alonso and Bellet, 2009, Guirao and Campa,2015
Business tourism development	Amiard, 1997; Faye, 1998 ; Tourisme-Alsace, 2009; Bazin et al., 2010, 2011, 2013a; Ville de Marseille, 2011 ; Delaplace & Perrin, 2013 ; AUDRR, 2012, Urenaet al., 2009 ; Alonso and Bellet, 2009 ; Todorovitch *et al*., 2011
Increase in tourism revenues	Chen and Haynes, 2012
New forms of governance	Bazin et Delaplace, 2015, Delaplace and Benoit-Bazin, 2017
Positive impacts on tourist destination choice	Pagliara *et al*., 2014; 2015 ; Delaplace *et al.* 2014, 2016; Saladié *et al*., 2016, Valeri et al., 2012

Source: Adapted and completed from Delaplace et al., 2014.

Similarly, according to an analysis by the French statistics office, INSEE, between 2007 and 2008, the average stay in tourist accommodation declined in almost all areas served by the East European High speed Line, as well as in Reims in 2007 (INSEE Lorraine, 2009). This reduction also reflects the change in the strategies of companies, which are increasingly moving towards one-day conferences to reduce budget. Furthermore, this reduction in overnight stay can lead to an overall reduction in spending by tourists (Levinson, 2012). This trend reveals the contradictory impacts of HSR on tourism (Albalate, Bel, 2010). The analysis of numerous case studies also shows the importance of public policies in served cities (Bazin-Benoit and Delaplace, 2015, Delaplace et Benoit-Bazin, 2017).

Table 2. A review of the studies showing negative or no impacts of HSR on tourism

Impact	Studies (Realized; Expected)
No growth of tourism in small cities	**Bazin** *et al*. **2013a**
Decrease in overnights stay	Bonnafous, 1987; Mannone, 1995 ; Haynes, 1997; Sands, 1993
Fall in the length of stays in the long run	INSEE Lorraine 2009; Bazin *et al.*, 2011; 2014, AUDRR, 2012, Bonnafous, 1987, Mannone, 1995; Amiard; 1997, Okabe, 1980 ; Chen,2013
Expansion of same-day round-trip journeys	Mizohata, 1995
No growth in the winter-sports tourism	**Bonnafous, 1987**
Reinforcement in the competition between tourist destinations	Chen & Haynes, 2012 ; Coronado *et al.* , 2013 ; Wang *et al.*, 2012 Masson & Petiot, 2009
Negative impacts on air transport tourism	Albalate, D., Fageda, X. 2015
Decrease in tourism revenues due to the decrease of the length of stay	Levinson, 2012
No impact on tourism destination choice	**Delaplace et al., 2016; Valeri** *et al.*, **2012**

Source: Adapted and completed from Delaplace et al., 2014

Another important point is that there has been no systematic impact on destination choice. For example, a survey carried out in 2012 in Paris showed that HSR is the third most important element in the choice of destination after historical and cultural heritage, and architecture (Delaplace *et al.*, 2014). Moreover, the modelling results showed that the HSR variable was highly significant in the probability of returning to Paris for tourism purposes. In another study on Disneyland visits, tourists declared that HSR was so crucial in their choice of destination that they would not have come without the presence of an HSR service (Delaplace et al., 2016). Another survey, conducted in Naples, showed that there was an impact with regard to visits to Naples and the intention to visit other cities nearby by HSR (Pagliara, 2014).

The analysis conducted in Madrid showed similar results: no impact on Madrid itself, but an impact on cities linked to Madrid by HSR (Pagliara *et al.*, 2015). Here, the extreme heterogeneity of situations comes to the fore, often with reference to specific cases, underlining the need for contextualization (Delaplace, 2012). The impacts of HSR cannot be understood without considering the socio-economic characteristics of the areas it serves, in tourism as in other fields (Bazin et al., 2013b).

3. Background on HSR and tourism sector in Turkey

3.1 HSR Network in Turkey

The first HSR line, Ankara-Eskişehir (ANK-ESK), started to serve in 2009, followed by Ankara-Konya (ANK-KON) line in 2011. After the completion of the 155 km HSR corridor between Eskişehir and İstanbul, Ankara based HSR service was extended to İstanbul (ANK-IST) in 2014, followed by KON-IST services in the same year (see Figure 1).

Total HSR line length, annual passenger volumes and passenger-km until 2016 are presented in Table 3 based on the most recent statistics by Turkish State Railway (TSR). Total ridership has been increased from about 1 million to 6 million trips in 8 years; ANK-ESK HSR line has been the most demanded service with a constant increase in the first five years of operation. ANK-KON has also been used with an increasing demand in the first four years of operation. But, KON-ESK line has not created as big a demand as the other two lines (see Table 3) and the direct services were cancelled for this line (currently it is served by KON-IST line) in 2015. In a TSR report, it was claimed that HSR shares in ANK-ESK and ANK-KON corridors reached 70% and 66% and induced demand was generated in these two lines as 12% and 18%, respectively (TSR, 2016). As a result, it caused drastic decreases in bus shares (from 55% to 10% in ANK-ESK corridor and from 70% to 17% in ANK-KON corridor), as well as private car shares (from 37% to 18% in ANK-ESK corridor, from 29% to 17% in ANK-KON corridor).

Table 3. Evolution of passenger volume and passenger-km for HSR lines in Turkey (TSR, 2016; TSR 2017a)

Years	2009	2010	2011	2012	2013	2014	2015	2016
Total Line Length (km)	397	888	888	888	888	1213	1213	1213
Total Ridership (x106)	0.94	1.89	2.57	3.35	4.21	5.09	5.69	5.90
Total Passenger-Km (x106)	237	476	665	914	1,186	1,554	1,847	NA
Ridership by HSR lines (10^6)								
ANK-ESK	0.94	1.89	2.15	2.00	2.27	1.92	1.28	2.20
ANK-KON	---	---	0.41	1.39	1.75	1.89	1.80	0.68
KON-ESK	---	---	---	---	0.20	0.25	---	---
ANK-IST	---	---	---	---	---	0.99	1.96	2.20
KON-IST	---	---	---	---	---	0.31	0.66	0.68

Note: NA: Not available data.

In a customer satisfaction survey conducted in 2016 by TSR reaching 1455 passengers travelling along HSR corridors, it was observed that 35% of the participants travelled 2-5 times, 18% of them travelled 6-10 times, 12% travelled 11-20 times and %17 of them travelled more than 21 times in one year (TSR, 2017b).

On ANK-ESK HSR line, 29% of the passengers travelled more than 21 times in a year while 14% travelled 11-20 times. Similarly, for ANK-KON HSR line, 22% of the passengers travelled more than 21 times and 14% travelled 11-20 times in a year. This suggests frequent use of HSR for regular activities.

When the purpose of the realized HSR trips were examined, it was revealed that 57% of the participants had tourism purposes (vacation, family/friend visit), while 20% of them had business purposes. Education trips (for students or lecturers, etc.) consisted only 7% among other trip purposes. For ANK-ESK and ANK-KON lines, almost 50% of the trips had tourism purposes (vacation, family/friend visit).

The most important factor behind the HSR preference was stated by the participants as the “short travel time”, followed by “being fast”, safety and punctuality. However, about half of the participants stated that appropriate schedule of trains and comfort was also important in their choice.

The general satisfaction of participants from the HSR services was determined as 87% (TSR, 2017b). In another study on travel characteristics of HSR users (Celikkol-Kocak et al., 2017) conducted as HSR main stations, a significant amount of first-time HSR users (23.3% of the 412 participants) were interpreted as HSR is still in a growing stage and its current pricing level encourages people to experience it. Also, it was stated that more than 40% of the participants had business or tourism trip purposes.

3.2 Service characteristics of the transportation modes along HSR corridors

As the modal attributes are necessary to develop mode choice models, travel time and cost information of the HSR, bus and air alternatives were compiled from National Transportation Portal (NTP) for each HSR corridors as shown in Table 4 (NTP, 2016).

The weekly service frequencies for all three modes were gathered to give an idea about the size of the intercity passenger sector for these corridors. (Service is defined as the number of round-trips for a given corridor). ANK-ESK was the most frequently served HSR line with 77 services in a week.

Second highest service frequency was observed on ESK-IST line with 56 services a week. KON-ESK and KON-IST are the least served HSR lines with only 14 services a week (corresponding to 2 round- trips a day). These service frequencies are highly correlated with the annual ridership values of HSR lines, but it is not possible to comment on the causality of the relation, as whether low service frequencies cause low ridership, or viceversa.

Table 4 . Service attributes of HSR and alternative modes

	Road Distance (km)	Weekly Service Frequency (One-way Trips)			Average Ticket Fare ($)*			Travel Time (minutes)**		
HSR Lines (in operation)		HSR	Bus	Air	HSR	Bus	Air	HSR	Bus	Air
ANK-KON	262	42	442	30	8.7	8.3 (0.3)	68.6 (9.2)	115	222 (22)	288 (82)
ANK-ESK	235	77	511	---	8.7	7.5 (0.8)	---	95 (2)	191 (20)	---
ANK-IST	450	42	854	350	20.3	17.1 (3.4)	44.8 (21.4)	249 (5)	386 (30)	72 (9)
KON-ESK	340	14	84	---	11.2	12.5 (0.7)	---	100	304 (9)	---
ESK-IST	310	56	359	---	13.1	9.8 (1.7)	---	152 (4)	331 (23)	---
KON-IST	712	14	259	64	24.7	20.7 (1.4)	37.6 (13.1)	260	620 (48)	82 (7)
HSR Lines (under construction)										
ANK-BUR	387	---	492	5	---	13.9 (2.2)	28.5	135	300 (19)	50
ANK-SIV	440	---	279	---	---	12.9 (1.1)	---	120	394 (20)	---
ANK-IZM	585	---	358	237	---	17.7 (3.3)	62.4 (27.1)	210	501 (24)	209 (109)

(*)Ticket costs are converted to US Dollars with a rate of 3.44TL/$ for October, 2016. (**)Travel time values were rounded up to the whole numbers.

While bus transportation is an available alternative for all corridors, its service frequency varies greatly among the corridors: on ANK-IST corridor which connects the two most populated cities in Turkey, İstanbul (14.4 million) and Ankara (5.1 million), there are 854 services a week. It is followed by ANK-ESK corridor (511 services a week), which is mostly a sub-network of the ANK-IST corridor. Air transportation is available for three HSR corridors (as ANK-KON, ANK-IST and KON-IST) and there is no air alternative for travellers in Eskişehir due to the proximity of the city to the other major HSR cities (Ankara, Konya, and İstanbul). Both the legacy and low-cost airline services in ANK-KON corridor include a must-transfer in İstanbul airport which increases cost and travel time significantly ($68.6 and 288 min., respectively) and this makes unattractive air alternative on this corridor. However, there are direct air services for ANK-IST and KON-IST corridors.

Note: While calculating the average travel time and cost values one week service information was gathered (28.11.2016-04.12.2016) and the arithmetic average was taken. For HSR ticket cost, economy class adult ticket; for intercity bus standard ticket (constant price for one-seat) was taken. Additionally, standard deviation values were given in the parenthesis.

3.3 Tourism sector inTurkey

Turkey is characterized by different tourism types such as coastal tourism, cultural tourism, thermal tourism, religious tourism. The share of tourism sector in Gross National Product (GNP) of Turkey has increased from 0.8% to 4% between 1983 and 2013 and it employs more than 12 million people. Turkish tourism sector succeeded to become in the top ten countries that attracts more international tourists with 29.9 million visitors in 2012. In 2014 as in in 2015, Turkey is at the sixth position in the world with respectively 39.8 and 39.5 millions of foreign arrivals (United Nations World Tourism Organization-, 2016). Ministry of Culture and Tourism Statistics showed 25 million foreign tourist arrivals in 2016, with top three countries of nationalities as Germany, Georgia and United Kingdom. The statistics showed that majority of the foreign tourists arrived Turkey by airway (70.1%) followed by road transportation with 23.8%. Railway had a very small share with 0.06% and maritime transportation had about 6% share (Ministry of Culture and Tourism, 2017).

According to the 10th Development Plan (prepared for 2014-2018 years), tourism sector is expected to constitute 45 million dollar share of 1.3 billion dollar Gross Domestic Product (GDP) of Turkey. Also the number of international tourist is expected to reach 48.3 million by 2018 (Ministry of Development, 2014). Moreover, in the 2023 Tourism Strategy of Turkey, the aim is to make Turkey in the top five countries in the world concerning the tourists' number and the tourist' revenues and to develop domestic tourism. Additionally, 8 thematic zones have been proposed including some HSR cities (For instance Phryg cultural and thermal tourism zone concerns Eskisehir and Afyon).

4. Evaluation of the link between HSR and tourism in Turkey

4.1 Methodology

The study aimed to present a preliminary evaluation of the link between HSR and the tourism sector in two Turkish cities as Eskişehir and Konya. In order to give general information on city- based characteristics, a general evaluation was presented for all HSR cities including population, employment rate, satisfaction with urban transportation services, life index ranking and human development index. Secondly, tourism potential of HSR cities was be evaluated based on available tourism statistics, socio-economic characteristics of the cities, major touristic destinations (in or within the vicinity of the city) and tourism types. Spatial distribution of the touristic destinations in Konya and Eskişehir was evaluated in more detail with the prepared maps and the accessibility of touristic destinations from HSR station (intermodal connections, public transit availability, etc.) was investigated.

4.2 General presentation of HSR cities

A background information is presented in Table 5 for the seven cities currently served and six cities soon to be served by HSR, showing population, population change rate for the period of 2016 and 2023, satisfaction with in-city transportation service, life index ranking and human development index. İstanbul and Ankara are the two biggest cities as mentioned above. Bursa and İzmir also have high populations of 3 and 4 million, respectively, that will have HSR services soon. While Kocaeli and Konya have populations around 2 million by 2023; but, Eskişehir, Sakarya and Manisa will be relatively smaller cities with populations around 1 million people. On the other hand, Bilecik, Sivas, Afyon and Uşak are much smaller cities on the HSR corridors, which are also not expected to have high population change rate. While these smaller cities seem to have slightly higher employment rate than others, the employment rates of the cities seem to be stuck within the range of 40%-55%. Big cities like Ankara and İstanbul seem to have problems in transportation services, while people in smaller cities have much higher satisfaction with transportation services.

Table 5. General statistics about HSR cities

	Population			Emp. Rate (2013)* %	Satisfaction w/ transportation service(2013)* %	Life Index Ranking (2015)*	Human Dev. Index (2013) **	HSR Passenger Arrivals (2016)*** $x10^3$
	2016	2023	Change Rate					
Current HSR Cities								
Ankara	5,325	5,927	16	45	71	17	2	2,354
Eskişehir	835	908	13	43	73	15	10	1,375
Bilecik	211	223	8	48	86	13	48	61
Sakarya	949	1,026	12	50	89	2	20	46
Kocaeli	1,765	1,984	18	49	73	23	6	172
İstanbul	14,864	16,569	16	46	69	5	1	779
Konya	2,104	2,175	5	46	88	18	8	1,104
Upcoming HSR Cities								
Bursa	2,839	3,073	12	48	83	19	5	---
Sivas	612	582	-6	45	82	40	36	---
Afyon	708	707	0	51	94	10	42	---
Manisa	1,362	1,371	2	53	91	31	21	---
Uşak	348	353	3	51	94	6	39	---
İzmir	4,138	4,405	9	47	77	21	3	---

**The data were compiled from TurkStat (2016)*

***The source of the data is The Economic Policy Research Foundation of Turkey-TEPAV (2016).*

**** The data source is Turkish State Railways (2017).*

Life index rating which shows the overall rating for the cities considering residents, employment, income, health, education, environment, security, accessibility to infrastructure was obtained for current and upcoming cities (TurkStat, 2016). Among 81 Turkish cities, Sakarya was ranked as 2nd city and İstanbul was ranked as 5th city. The other index was Human Development Index which evaluates development in terms of health, education and income was also given in Table

According to 2013 data, İstanbul and Ankara became the first two cities in terms of human development. Among the upcoming HSR cities, İzmir became the 3rd city followed by Bursa as 5th city.

4.3 Eskişehir as a HSR city

4.3.1 Tourism in Eskişehir

Eskisehir is an attractive commercial, industrial and educational destination in the Central Anatolian Region of Turkey (Figure 2). It has become one of the most popular cities in Turkey and it is setting an example for other municipalities (Samkar & Alpu, 2013). The local administration is one of the most important actor in terms of shaping city image and improving the existing cultural and touristic attractions in the city (Tokay-Argan, 2016). In 2013, the city became the Turkish World Capital of Culture, a title given to the cities having a certain level of cultural heritage (Timur *et al.*, 2014; Tokay-Argan, 2016).

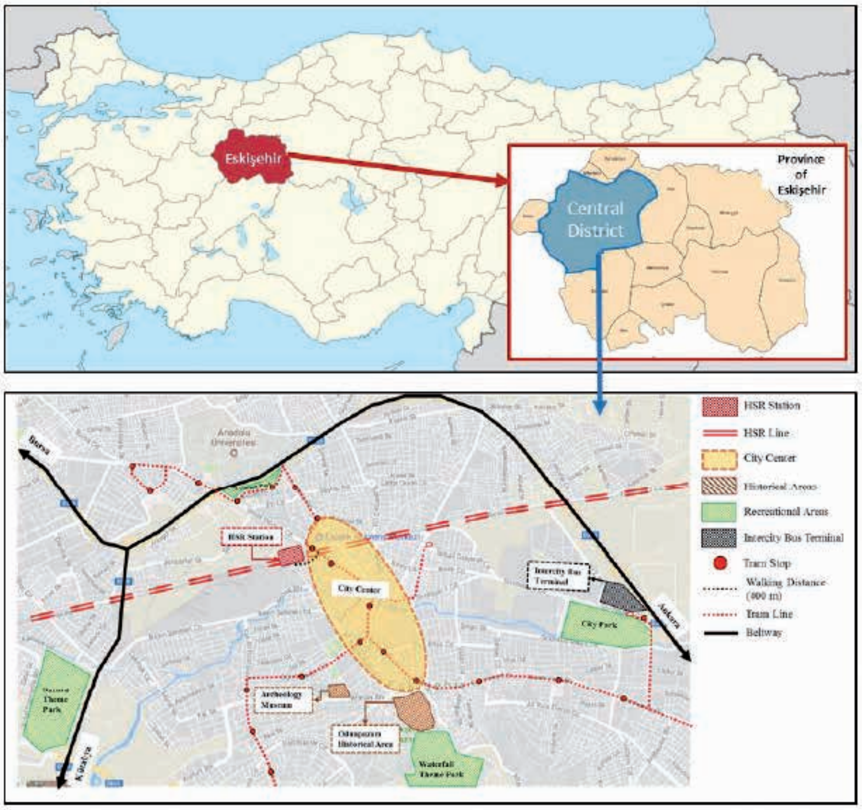


Figure 2. Location of tourist attractive places in Eskişehir

Eskişehir, with its historical and cultural fabric, economy, entertainment and nightlife, infrastructure (roads, communications, transport, etc.), shopping centres, recreational activities, housing, security, sports facilities, parks, services, traffic, medical facilities, restaurants, and green spaces, has the specialty of offering big city comforts in a small town. In Figure 2, a conceptual map showing the locations of attractive places in Eskişehir is given. One of the most attractive place in Eskişehir is Odunpazarı Historic Area which contains buildings with unique architecture constructed in the late period of the Ottoman Empire. The area was restored a few years ago and transformed into a living area with restaurants, boutique hotels and shopping places that attract tourists (Okumuş, 2013; Tokay-Argan, 2016). The city centre was also renewed and became attractive place in the recent years. In addition, there are a few big theme parks (The City Park, Science, Art, and Culture Park, Waterfall Park, etc.) constructed and opened for leisure activities, as well (see Image 1).

When considering tourism sector statistics, the province has much lower inputs compared to the existing potential (Sahin, 2012). However, according to the 2023 Turkey National Tourism Strategy Report, it is stated that a thermal tourism master plan will be prepared for the Phrygian region in which Eskisehir is also located and this region will be an important centre for health tourism (Ministry of Culture and Tourism, 2007). According to 2011 data, there are 55 accommodation facilities with 5603 bed capacity and the occupancy rate is around 40% for the years between 2008 and 2010.

Image 1. Some of the touristic places in Eskişehir a) Riverside in City Centre b) Odunpazarı Historical Area c) City Centre d) Sazova Theme Park

In a 2014 study conducted with domestic travellers, the destination of Eskişehir was described as a “student-university and cultural city” (Üsküdar et al., 2014, Tokay-Argan, 2016). In another study, ease of transportation, economic comfort, geographical proximity, existing of cultural places, parks and museums, having a social life and shopping options were found as the factors motivating people to visit Eskişehir (Oyman et al., 2010; Tokay-Argan, 2016). Another study showed that the factors affecting the development of tourism in Eskişehir were marital status, educational levels and income level. These factors caused differences in perception of problems: as the level of income increase, perceptions about the existence of problems increase. When the impact of HSR on Eskişehir was investigated, it was found that there has been an increase in daily trips due to the travel time saving provided by HSR (Seçilmiş, 2011).

Figure 3 shows the number of arrivals and stay overnights between 2007 and 2012, the HSR service having begun in 2009. There has been an increase in total number of visitors from 138,677 to 173,461 between 2007 and 2010, while foreign visitor number had some fluctuations in this period. However, 2012 data indicates that the total number of visitors decreased to 143,526. There has been a 35% increase in the total number of stay overnight between 2007 and 2010 (Akoğlan-Kozak et al., 2011). It decreased to 231, 515 in 2012. The number of stay overnight for foreign visitors increased to from 18,729 to 23,513 between 2010 and 2012. (Note: for Figure 3, 2007-2010 data was gathered from Akoğlan Kozak et al. (2011), 2012 data was compiled from Eskişehir Culture and Tourism Provincial Directorate (2012), and 2011 data is not available.)

4.3.2 Connectivity to HSR station

Due to its locations, Eskişehir is a railway hub, where both conventional trains (both passenger and freight locomotives) and HSR lines serve. Eskisehir Train Station is located at the city centre of Eskisehir, which is also combined with intercity bus service to Bursa, an upcoming HSR city (see Figure 2 and Image 2). For in-city transportation, there are taxi services and bus services close to station. Also, there is a tram stop 400 m walking distance.

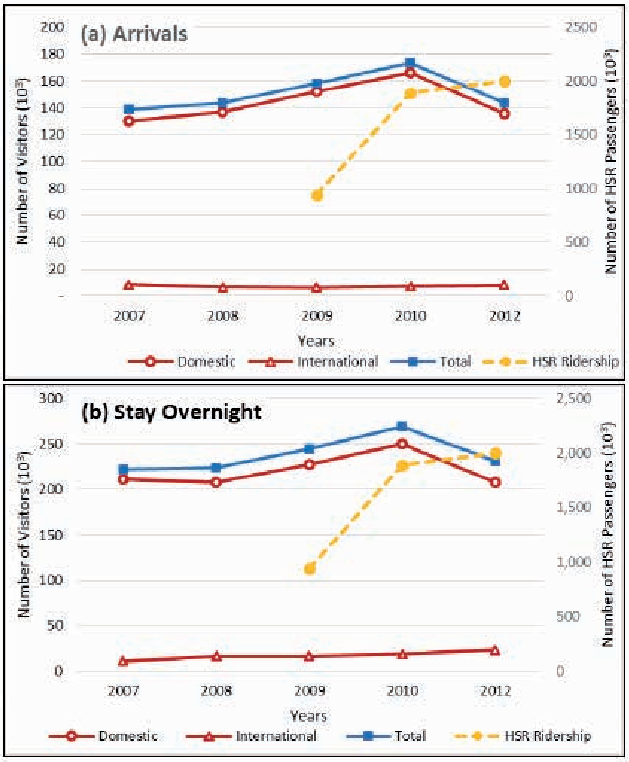


Figure 3. Number of arrivals and stay overnights for different years

Image 2. Eskişehir train station a) outside view (showing the combined bus service stop), b) ticket offices c) railway platform

4.3.3 Observations during the technical trip to Eskişehir

A technical trip to Eskişehir started from the Eskişehir Railway Station on foot and included some touristic places in the city center, mainly around the Odunpazarı region, which revealed that the city had a good image and there were a lot of attractive places. Some other weaknesses were observed as lack of actively working tourist information centre and urban transportation information at train station (see Image 2).

While the tramway network is still in a developing stage, it connects the city centre to the most attractive places in the city such as Odunpazarı Historical Center, intercity bus terminal, universities, etc., though not the HSR station directly; the nearest tramway stop is located within 400 m walking distance (see Figure 2) and there are no directional signs leading HSR users to the city centre.

The Sazova Theme Park was visited, which was not accessible with public transportation service directly. However, there exists the Eskişehir Mobile City Guide, a smart phone application, which includes the location information on cultural places, museums, shopping places, universities, etc. is very comprehensive and informative. An interview with the Eskişehir Provincial Culture and Tourism Directorate revealed that HSR affected the urban economy and the number of daily visits increased after HSR started to operate.

4.4 Konya as a HSR city

4.4.1 Tourism in Konya

Konya is an important centre in terms of religious and culture tourism both in national and international level. It was the place where Rumi (or Mevlana), an important Persian-born Islamic scholar, lived most of his life. His tomb has become a place of pilgrimage today. The Whirling Dervishes perform once a year during the Mevlana Festival in every beginning of December. In 2005, the Mevlana Sama Ceremony in Konya was proclaimed by UNESCO as one of the Masterpieces of Oral and Intangible Heritage of Humanity (Egresi et al., 2012).

In addition to Mevlana Museum, there are a lot of historical places and recreation areas in Konya. Figure 4 shows the locations of attractive places and some of their photos are given in Image 3.

According to the information given by the Konya Provincial Culture and Tourism Directorate, Konya has further potential in health, congress, fair and meeting, nature, sport and hunting tourism as well (Konya Provincial Culture and Tourism Directorate, 2017, interview, 25/03/2017). Currently, Konya has 37 hotels (having 2-5 stars) with 3080 rooms and 6161 beds and 107 accommodation places with 2430 rooms and 5189 bed capacities. However, 12 hotels with 4769 beds have also received investment certificates.

In addition, World Trade Centre and Fair Exhibition area, which will have a conference room with 1000 seats, meeting halls, exhibition areas and restaurants, is in the planning phase. Also, there are 7 first class restaurants having investment certificate in the city (Konya Provincial Culture and Tourism Directorate, 2017).

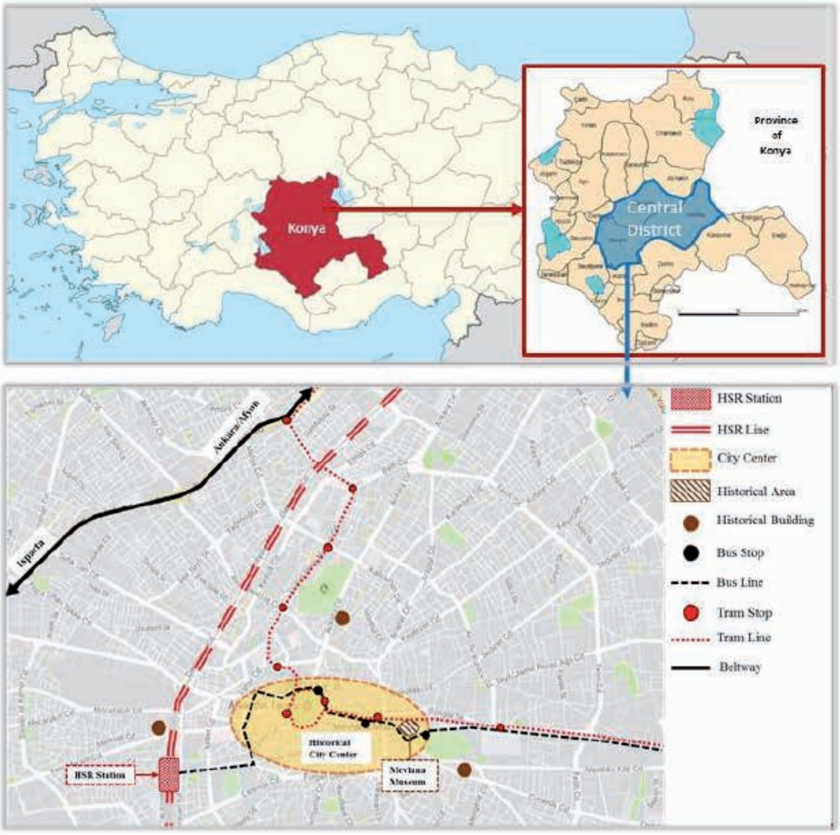


Figure 4. Location of tourist attractive places in Konya

Image 3. Touristic places in Konya a) Mevlana Museum Square b) Mevlana Museum c) Aleaddin Hill and Mosque d) City Centre.

The number of tourists visiting Konya[1] in Figure 5 showed a rapid increase in the number of visitors between 2007 and 2016 (HSR service to Konya begun in 2009). However, there are some fluctuations in the number of visitors in this period: In 2009, 2012, 2014 and 2016 years decreases in the number of visitors have been registered.

The interview with the city Tourism Directorate revealed potential reasons as economic and political situations. Moreover, after the application of Museum Card started in 2011, domestic and foreign visitor data could not be recorded separately. However, when the numbers of foreign tourists in the previous years were examined, it was seen that Konya had attracted tourists from the Republic of China, followed by Japan, Iran, United States and Germany in turn (Konya Provincial Culture and Tourism Directorate, 2017).

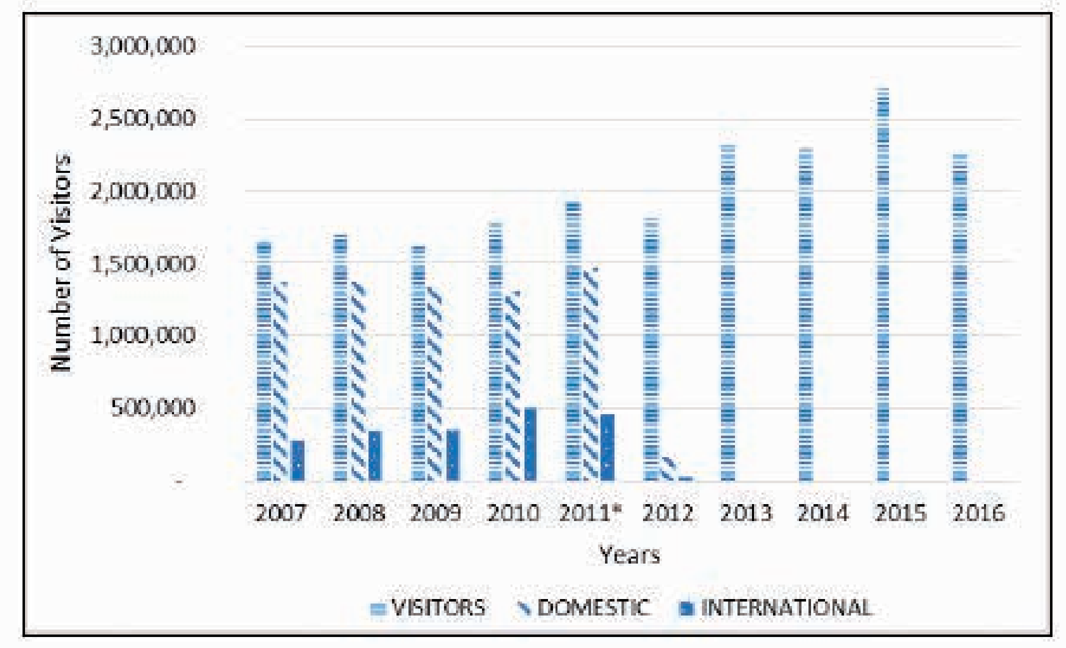


Figure 5. Number of Konya visitors by different years

***As Museum Card application was started in 2011, domestic and international visitor data were not available after 2011.**

Figure 6 shows the number of arrivals and stay overnights between 2011 and 2016, which showed an increase in both total and domestic values. Even though HSR ridership has increased between 2011 and 2014, decrease in the ridership was observed in 2015 and 2016. When the monthly distribution visitors is examined for 2016 shown in Figure 7, it is seen that the number of arrivals and stay overnights increased in March, August, October and December months.

At the beginning of October, there is an annual event called as "International Mystic Music Festival" which attracts more domestic and foreign tourists. Also, Mevlana Sama Ceremonies that are realized in every beginning of December could be a reason for the increase in the number of visitors in December period. Even though, a relationship between HSR passenger arrival and visitor numbers was not observed in the first half of the year, in the second half, it seems that there is a parallel relationship between them.

---

[1] The data included the number of tourists visiting museums and cultural places such as Mevlana Museum, Karatay Museum, etc.

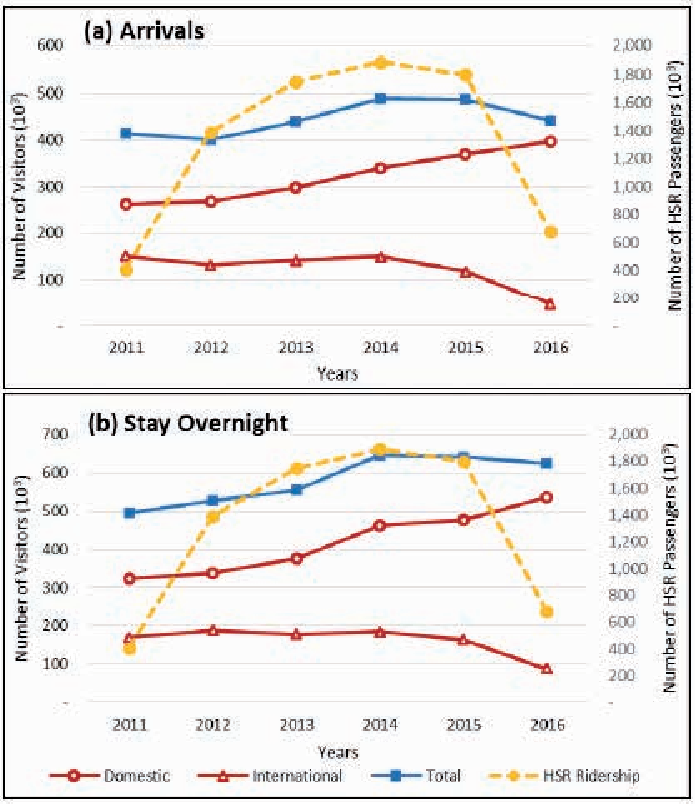


Figure 6. Number of Konya Visitors and ANK-KON HSR Ridership a) Arrivals, b) Stay overnight

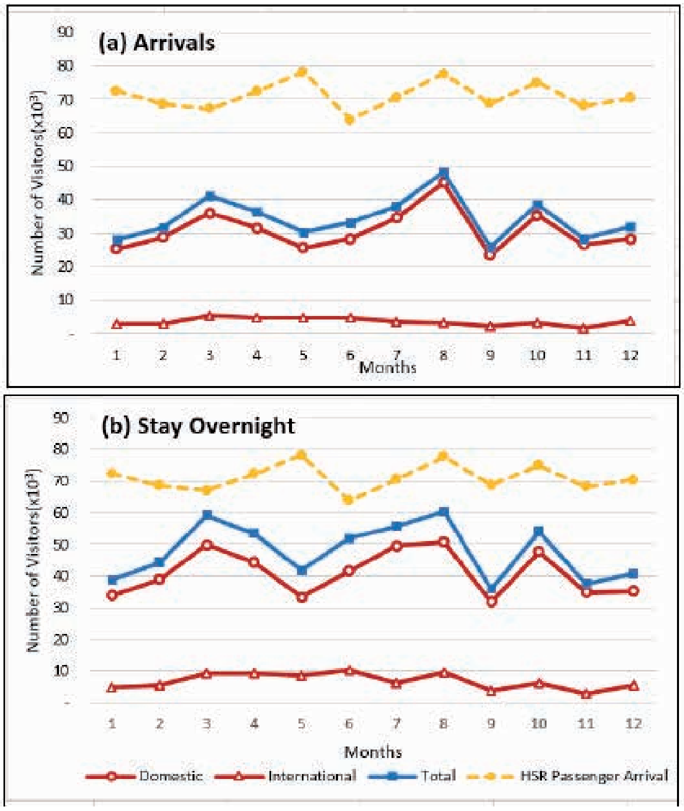


Figure 7. Monthly Distribution of Visitors and HSR Passenger Arrivals to Konya in 2016 a) Arrivals, b) Stay overnight

4.4.2 Connectivity to HSR Station

Currently, there are two HSR lines served at Konya HSR Station, ANK-KON and KON-IST. Konya station is located relatively close to the city centre (around 2 km away). Konya train station serves both HSR (12 departures and 12 arrivals per day) and conventional railways (passenger and freight locomotives). In a study on the regional impacts of HSR, it was determined that 23% of the passengers traveling with Ankara-Konya HSR had tourism purposes. Also, it was stated that people living in Ankara mostly travel to Konya for touristic purposes (MEVKA, 2012). Photos of Konya Train Station are given in Image 4. As it was stated before, there are combined HSR services with intercity busses and their stop is in front of the station for the passengers going Karaman and Alanya (see Image 4a). Also, there is an in-city bus stop in front of the station that provides direct accessibility to city centre (Image 4d).

Image 4. Konya train station a) outside view (showing the intercity buses combined with HSR), b) outside view, c) railway platform d) public bus stop in front of the station

4.4.3 Observations during the technical trip to Konya

During the Konya technical trip, it was observed that tourist information centre located in Konya Railway Station (see Image 5a) and there is bus stop in front of the station that provide direct access to the city center. There is also a shared bicycle system (see Image 5b). Moreover, Konya Metropolitan Municipality's smart phone application was found as quite useful that provides information on touristic places, transportation alternatives, public transportation stops, municipal services etc.

Image 5. Photos of a) tourism Information office in rail station (b) and bike-sharing system in Konya city centre

4.5 Some recommendations concerning HSR and tourism expansion in Turkey

As a result of the first observations made during the technical trips and the analysis of the literature, some recommendations can be put forward to develop tourism by HSR.

First of all, due to the newness of HSR in Turkey, communications concerning its possible use for tourism purpose should be provided in rail stations and airports.

Second, in HSR stations, wireless internet (wi-fi) access should be provided for tourists and in particular for foreign tourists, and tourist information offices and souvenir shops should actively work.

Third, in trains, images related to the touristic areas in the city should be presented on the TV and activities should be done in parallel with the urban identity, information about intra-city transportation should be presented.

Fourth, in order to provide intermodality with in-city transportation options, information about the public transportation should be presented in front of the station and walkways should be provided with visible signs. Walkability of stations should be assessed within 500 m and bicycle sharing options should be evaluated. Also, HSR station and intercity bus terminal connections should be strengthened and combined ticketing options should be increased.

Fifth using intelligent transportation systems, integration of HSR ticket with conventional train and urban transportation tickets can be realized. Also, HSR ticket, other transportation tickets and Museum Card system might be integrated in a single pass with certain discounts for activities in the city. In addition, intelligent shuttle/taxi service might be provided with the provision of prepaid systems (online payment systems, the shortest route representation, etc.).

5. Conclusions

The comfort and the travel time saving provided by the HSR services might contribute to tourism sector (especially daily tourism), as well as attractiveness of the cities. But the literature review shows that the link between HSR and tourism is not systematic. Some conditions are required such as touristic resources, intermodality at the rail station, and local policies. In Turkey, HSR is a newly introduced transportation mode, which has a network developed branching out from Ankara (ANK), the capital city, first to the two nearby cities, Eskişehir (ESK) and Konya (KON). While the network further reached Istanbul, the biggest city in Turkey, and is planned to reach İzmir and Bursa, two major city in Turkey, the biggest share of the current ridership comes from ANK-ESK and ANK-KON lines, which were chosen for major reductions in travel times. Majority of the travels in these lines was for tourism purposes (vacation, family/relative visits, etc.).

Though not measured in detail, HSR has started to affect both the service and tourism sectors, especially in Eskişehir and Konya. When basic tourism statistics for Eskişehir and Konya were evaluated, it was seen that the daily trips made to these cities increased as well as the tourism arrivals. But to measure the contribution of HSR to tourism more clearly and to determine the applications that can be made to increase this contribution further, more detailed researches must be conducted on these cities, but also on the other upcoming HSR cities. These researches are expected to shed light for the HSR services for the upcoming HSR lines in Turkey and in the other developing countries.

6. References

- ALBALATE, D., CAMPOS, J., JIMENEZ, J-L. (2015). Tourism and high speed rail in Spain: Does the AVE increase local visitors? Technical Report, December.

  http://www.researchgate.net/publication/288604283 [Accessed 24 April 2016].

- ALBALATE, D., BEL G. (2010). High-Speed Rail: Lessons for policy makers from experiences abroad. Research Institute of Applied Economics Working Paper,03, 34.

- ALBALATE, D., FAGEDA, X. (2016) High speed rail and tourism: Empirical evidence from Spain, Transportation Research Part A, 85, pp.174–185.

- ALONSO, M., BELLET, C. (2009). El tren de alta velocidad y el proyecto urbano. Un nuevo ferrocarril para la Zaragoza del tercer milenio", in "Scripta nova. Revista electrónica de geografía y ciencias sociale. Barcelona: Universidad de Barcelona, 13, 281.

- AMIARD D. (1997). Le tourisme d'affaire et de Congrès dans l'agglomération mancelle Dans Chevalier, Jean (sous la direction de), Le Mans 6 ans après l'arrivée du TGV, groupe de recherche en géographie sociale, ESO – Espaces géographiques et Sociétés, Université du Maine, Le Mans.

- AGENCE D'URBANISME RÉGION DE REIMS (AUDRR) (2012) TGV Est Approche des impacts socio- économiques sur Reims et sa région. Paroles d'acteurs. Mai

  [www.audrr.fr/centre_ressources/publication/paroles-d-acteurs-tgv-est.html]

- BABALIK-SUTCLIFFE, E. (2007). Pro-rail policies in Turkey: A policy shift. Transport Reviews: A Transnational Transdisciplinary Journal, 27(4), pp.485-498.

- BAZIN-BENOIT, S., BECKERICH, C., DELAPLACE, M. (2013a). Desserte TGV et villes petites et moyennes, une illustration par le cas du tourisme à Arras, Auray, Charleville-Mézières et Saverne. Les Cahiers Scientifiques du Transport, 63, pp. 33–62.

- BAZIN S., BECKERICH C., DELAPLACE M. (2011). High Speed Railway, Service Innovations and Urban and Business Tourisms Development, Chapitre 4 In SARMENTO Manuela, & ALVARO Matias Economics and Management of Tourism: Trends and Recent Developments. Collecçao Manuais, Universidade Luisiada Editora, Lisboa, Portugal.

- BAZIN-BENOIT, S., BECKERICH, C., DELAPLACE, M. (2010). Grande vitesse, activation des ressources spécifiques et développement du tourisme urbain: Le cas de l'agglomération rémoise. Belgeo, 1–2, pp. 65–78.

- BAZIN S., BECKERICH C., BLANQUART C., DELAPLACE M. (2013b). Les enjeux et opportunités des dessertes ferroviaires à grande vitesse en matière de développement local et de développement durable, rapport final, Contrat PREDIT, financement ADEME, mai, 186p.

- Bazin S., Beckerich C., Delaplace M. (2014). Valorisation touristique du patrimoine et dessertes TGV dans les villes intermédiaires à moins d'1h30 de Paris: les cas de Reims", Metz, Le Mans et Tours. Revue d'économie Régionale et Urbaine, 5, pp. 5-23.

- BAZIN S., BECKERICH C., DELAPLACE M. (2013a). Desserte TGV et villes petites et moyennes, Une illustration par le cas du tourisme à Arras, Auray, Charleville-Mézières et Saverne. Les Cahiers Scientifiques du Transport, 63, pp.33-62.

- BAZIN-BENOIT S., DELAPLACE M. (2015). Mise en service des dessertes TGV et gouvernance dans le domaine du tourisme: le cas de villes françaises. Revue Géographique de l'Est, 55(3- 4)
- BONNAFOUS, A. (1987). The regional impact of the TGV", Transportation, 14(2), pp. 127–137.
- BUTTET ET AL.,(2001) Impact de l'arrivée du TGV sur la ville de Tours. Rapports d'atelier, Magistère 2ème année, CESA.
- CELIKKOL-KOCAK, T., DALKIC, G., TUYDES-YAMAN, H.(2017). High-Speed Rail (HSR) users and travel characteristics in Turkey. Procedia Engineering, 187, pp. 212 – 221.
- CHEN, Z., HAYNES, K. E. (2012). Tourism industry and High Speed Rail - Is there a linkage: Evidence from China's High Speed Rail development. GMU School of Public Policy Research Paper No. 2012-14. Available at SSRN:

  https://ssrn.com/abstract=2130830 or http://dx.doi.org/10.2139/ssrn.2130830
- CHEN Z., HAYNES K. E. (2014). Impact of high-speed rail on international tourism demand in China, Applied Economics Letters, 22(1), 57-60, DOI: 10.1080/13504851.2014.925043
- CHEN X. (2013). Assessing the impacts of High Speed Rail development in China's Yangtze River Delta Megaregion. Journal of Transportation Technologies, 3, pp. 113-122.
- CHENG, Y-H. (2009). High-speed rail in Taiwan: New experience and issues for future development. Transport policy, 17(2), pp.51-63.
- CORONADO J-M., GARMENDIA M., MOYANO A., UREÑA J-M. (2013). Assessing Spanish HSR network utility for same-day tourism, Rechs. Transportation Security, 29, pp.161-175.
- Comité Régional du Tourisme (CRT) Provence Alpes Côte d'Azur (PACA), « Modes de transport et Tourisme en Provence-Alpes-Côte d'Azur », Observatoire du Tourisme, CSEF.(2005). La gare TGV : quels impacts sur l'emploi à Liège ? Compte-rendu d'une table ronde organisée par le CSEF de Liège,

  http://csefliege.org/Analyse_du_marche_de_l_emploi_et_de_la_formation/Synthese%20TGV.pdf
- DALKIC G. (2014). High Speed Rail development in Turkey: Government policy, investments and users perspective, Master Dissertation, Middle East Technical University, Graduate School Of Natural And Applied Sciences, Ankara, Turkey.
- DELAPLACE M.; PERRIN J. (2013). Multiplication des dessertes TGV et Tourismes urbains et d'affaires, Regards croisés sur la Province et l'Ile de »rance", Recherche Transport et Sécurité, 29, pp. 177-191.
- DELAPLACE M., BAZIN-BENOIT, S. (2017). High-speed rail services and tourism expansion: The need for cooperation, in ALBALATE D. and BEL G. Eds. "Evaluating High Speed Rail" Routledge studies in transport analysis, pp. 69-81
- DELAPLACE M., PAGLIARA, F., PERRIN, J., MERMET, S. (2014). Can High Speed Rail foster the choice of destination for tourism purpose? Procedia - Social and Behavioral Sciences, EWGT2013 – 16th Meeting of the EURO Working Group on Transportation.

- DELAPLACE M., PAGLIARA F., LA PIETRA A. (2016). Does High-Speed rail affect destination choice for tourism purpose? Disneyland Paris and Futuroscope case studies? Belgeo, 3.

- DELAPLACE, M. (2012). Pourquoi les « effets » TGV sont-ils différents selon les territoires ? L'hétérogénéité au cœur du triptyque « Innovations, Territoires et Stratégies ». Recherche Transports et Sécurité, 28, pp.290-302.

- EGRESI I., BAYRAM, B., KARA, F., KESİK, O. A. (2012). Unlocking the potential of religious tourism in Turkey. GeoJournal of Tourism and Geosites, 1(9).
- ESKIŞEHIR CULTURE AND TOURISM PROVINCIAL DIRECTORATE (2012), Tourism Statistics, Available at:

  http://www.eskisehirkulturturizm.gov.tr/TR,149875/ istatistikler.html. Accessed date: 20/05/2017.

- AKOĞLAN KOZAK, M., KILIÇLAR, A., SARI, Y., KAŞLI, M., SEÇILMIŞ, C., ÇIÇEK, D., ZENCIR, E. (2011), Eskişehir Tourism Master Plan 2011-2015, Eskişehir: Anadolu University, Eskişehir, Turkey.
- FAYE M. (1998). Tours à la conquête du tourisme d'affaires. Norois, 178, pp. 293-300.
- GUIRAO B., CAMPA J-L. (2015). The effects of tourism on HSR: Spanis,h empirical evidence derived from a multi-criteria corridor selection methodology. Journal of Transport Geography, 47, pp. 37–46.

- HAYNES K.E. (1997) Labor markets and regional transportation Improvements: The case of High Speed Trains. An introduction and review. The annals of regional science, Vol.31, N°1, pp. 57-76.INSEE LORRAINE. (2009). La ligne à Grande Vitesse Est-européenne : une évaluation de l'impact sur le tourisme. INSEE Lorraine. 163, 6p.

- KONYA CULTURE AND TOURISM PROVINCIAL DIRECTORATE. (2017). Interview in 23/03/2017.Konya, Turkey.
- KURIHARA, T., WU, L. (2016). The impact of High Speed Rail on tourism development: A case study of Japan, The Open Transportation Journal, 10, (Suppl-1, M4), pp.35-44.
- LEVINSON DAVID M. (2012). Accessibility impacts of high-speed rail. Journal of Transport Geography, 22, pp. 288–291.

- LUMSDON, L. AND PAGE, S. J. (2004). Progress in transport and tourism research: reformulating the transport-tourism interface and future research agendas in Lumsdon, L.; (eds) Tourism and transport: issues and agenda for the new millennium, Oxford, Elsevier, pp.1-27.
- MANNONNE V. (1995). L'impact régional du TGV sud-est, Thèse pour l'obtention du doctorat de géographie, 2 tomes, Université de Provence Aix-MarseilleI.
- MASSON S.; PETIOT R. (2009). Can the high speed rail reinforce tourism attractiveness? The case of the high speed rail between Perpignan (France) and Barcelona (Spain). Technovation, 29(9), pp. 611-617.
- MEVLANA DEVELOPMENT AGENCY-MEVKA. (2012). The impact of HSR on regional development, Technical Support Program of MEVKA 2011. Konya, Turkey.

- MINISTRY OF CULTURE AND TOURISM. (2017). Tourism Statistics, Available at:

  http://yigm.kulturturizm.gov.tr/TR,9851/turizm-istatistikleri.html. Accessed date: 20.05.2017.

- MINISTRY OF DEVELOPMENT. (2014). The tenth development plan. Ankara, Turkey.
- MIZOHATA Y. (1995), "L'impact du Shinkansen sur les villes: Le cas du Shinkansen du Tohoku", in Centre Jacques Cartier, Patier Danièle (Eds). Villes et TGV : Actes des sixièmes Entretiens Jacques Cartier 8-10 Décembre 1993, Lyon, France : LET. 330 p. (Études et Recherches, n° 6), pp. 277-280
- Ministry of Transportation, Maritime and Communication (2016), National Transportation Portal. https://www.ulasim.gov.tr/
- OKABE, S. (1980). Impact of the Sanyo Shinkansen on local communities. Straszak A. and Tuch R. (ed), The Shinkansen High-Speed Rail Network of Japan, Pergamon Press, pp. 105- 129.
- OKUMUŞ, E. (2013). Odunpazarı Evleri'nde Ramazan Gecelerinde Boş Zaman Mekânları. Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi. 14, pp. 103–140.
- OYMAN, M., Yılmaz, H., & Kutlu, Ö. (2010). Motivations for visiting heritage sites: The case of Eskisehir. The International Journal of The Humanities, 8(6), 131–145.
- PAGLIARA F., DELAPLACE M., VASSALO J.M. (2014). High-speed trains and tourists: what is the link? Evidence from the French and Spanish capitals. Urban Transport XX, 17.
- PAGLIARA F., DELAPLACE M. AND VASSALLO J.M. (2015). High-speed rail systems and tourists' destination choice: The case studies of Paris and Madrid. The International Journal of Sustainable Development and Planning, 10 (3), pp. 399–410.
- PAGLIARA F. (2014). High Speed Rail systems. impacts on mobility, on tourism and on mobile workers. LAP Lambert Academic Publishing, Saarbrucken, Germany.
- People's Republic of China. (2014). Regional Economic Impact Analysis of High Speed Rail in China, Main Report June 25, Document of the World Bank.
- ŞAHIN F. (2012). Eskişehir ilinin kültür turizmi potansiyeli: mevcut durum ve öneriler, Uzmanlık Tezi, Kültür ve Turizm Bakanlığı Eskişehir Kültür Varlıklarını Koruma Bölge Kurulu Müdürlüğü.
- SALADIÉ O., ANTON CLAVÉ S., GUTIÉRREZ A. (2016). Measuring the influence of the Camp de Tarragona high-speed rail station on first-time and repeat tourists visiting a coastal destination. Belgeo, 3.
- SAMKAR, H., ALPU, O. (2013). M regression approach for satisfaction of municipality services: The case of Eskisehir. Quality and Quantity, 47(3), pp.1629–1637.
- SANDS B. D. (1993). The development effects of High-Speed Rail stations and implications for California, Institute of Urban and Regional Development University of Berkeley, working paper, http://www.uctc.net/papers/115.pdf
- SEÇİLMİŞ, C. (2011). The problems that affect the development of Eskişehir's tourism from the visitors' perspective, İşletme Araştırmaları Dergisi, 3(3),pp. 37-57

- TEPAV--The Economic Policy Research Foundation of Turkey. (2016). Türkiye'de insani gelişmişlik iller arasında nasıl farklılaşıyor? 81 il için insani gelişmişlik endeksi, Turkey.
- TIMUR, M. N., ÇEVIK, S., & KICIR, G. K. (2014). Etkinlik turrizmi: kültür başkenti etkinliklerinin başarı unsurları üzerine bir değerlendirme. Akademik Sosyal Araştırmalar Dergisi, 2(2/1), pp. 56–83.
- TODOROVICH P., SCHNED D., LANE R. (2011). High-Speed Rail international lessons for U.S. policy makers. Policy Focus Report Series, Lincoln Institute of Land Policy
- TOKAY-ARGAN, M.(2016). Eskişehir, Turkey as a crossroads for leisure, travel and entertainment. Cities 56, pp. 74–84.
- Tourisme- Alsace (2009). Le bilan après 2 ans de TGV Est en Alsace, Tourism Regional Council of the Alsace region,

  http://www.clicalsace.com/fr/thematique/offre-et- frequentation/impact-du-tgv-est-sur-les-activites-touristique-apres-2-ans,[Accessed 15.01.2016].
- TURKISH STATE RAILWAYS. (2016). Annual Statistics 2011-2015. Ankara, Turkey.
- TURKISH STATE RAILWAYS. (2017a). Annual Statistics 2012-2016. Ankara, Turkey.
- TURKISH STATE RAILWAYS. (2017b). Customer Satisfaction Survey Evaluation Report. Ankara, Turkey.
- TURKISH STATISTICAL INSTITUTE. (2016). Annual Population Statistics, Accessed from the web site: http://www.tuik.gov.tr/PreTablo.do?alt_id=1027
- TURKISH STATISTICAL INSTITUTE. (2016). Life Satisfaction Research, Accessed from the web site: http://www.tuik.gov.tr/UstMenu.do?metod=temelist
- TURKISH STATISTICAL INSTITUTE (2016). Basic Statistics, Accessed from the Web site:

  http://www.turkstat.gov.tr/Start.do;jsessionid=GT2VZmTLkTmnv25sGr-MhyWldG6Zw5Dh 6grTJ2gzWSJx2j3hrRH2p!-1124279051
- URENA J.; MENERAULT P., GARMENDIA M. (2009), The high-speed rail challenge for big intermediate cities : a national, regional and local perspective, Cities, Vol. 26, n°5, pp. 266- 279
- ÜSKÜDAR, Ş., CAKIR, M.., TEMIZKAN, S. P. (2014). Yerli Turistlerin Eskişehir'in Kültür Turizmine İlişkin Algıları (the perception of domestic tourists regarding Eskisehir's cultural tourism). Journal of Tourism and Gastronomy Studies, 2(2), pp. 67–76.
- VALERI E., PAGLIARA F., MARCUCCI E. (2012). A destination choice model for tourism purpose. ASRDLF 2012 conference special session on High Speed Rail, Tourism and Territories, 9th-11th July, Belfort, France
- VICKERMAN R., ULIED A. (2006), Indirect and wider economic impacts of high-speed rail

  http://163.117.2.172/temp/agenda/mad2006/papers/12.%20Vi ckerman,%20Roger.pdf

- VILLE DE MARSEILLE. (2011). Le tourisme made in Marseille, Direction de la communication.
- WANG X., HUANG S., ZOU T., YAN H. (2012). Effects of the high speed rail network on China's regional tourism development. Tourism Management Perspectives, 1, pp.34–38.
- WORLD TRAVEL TOURISM COUNCIL (WTTC). (2015). Travel & Tourism: Economic Impact 2015, Turkey.
- ZHAO, D. (2012). The High-Speed Railway network in Yangtze River Delta: an analysis of the accessibility impact.

  http://www.regionalstudies.org/conferences/presentations/european-conference-2012-best-international-paper-early-career [accessed 15 February 2016]

7. Acknowledgements

Authors would like to acknowledge the support of Ministry of Transport, Maritime and Communication, General Directorate of the Turkish State Railways for the organization of technical trip. Also, the authors gratefully acknowledge Konya Provincial Culture and Tourism Directorate for the data provision.

Economic impacts

360.revista de alta velocidad
número 5 - junio - 2018. Pág 435 - 449

Break-even point analysis of the Business Plan for a High-speed line in Egypt as a measure of financial sustainability

Maté Sanz, David

Fernández Gago, Jose Ángel

Chapinal Rivera, Silvia

INECO [1]

Abstract

INECO developed between 2015 and 2017 a study for the Spanish Ministry of Economy and Competitiveness (MINECO) titled “Feasibility Study for a High-Speed Rail between Cairo, Luxor, Aswan and Hurghada in Egypt”. In the framework of this study, a financial model was developed that has led to the identification of break-even points that guarantee the repayment of financial costs under different investment and demand scenarios.

This scenario analysis takes on special relevance in a macro economic, financial and situational context of the real economy such as of Egypt’s during this period.

This paper presents the main conclusions obtained in the specific case of a High-Speed analysis in Egypt. It is a high-speed railway infrastructure to be developed in a middle-income country, in a corridor highly linked to tourism and in a context of economic uncertainty.

Several investment infrastructure alternatives have been evaluated, assessing the suitability or lack of suitability of making this investment in the country. From the point of view of the financial analysis, the report provides the level of demand, or income, from which the project has found its break-even point in the short and medium term, complemented by the environment conditions in which this circumstance has been estimated.

The Break-even point analysis in the short term of the Business Plan for a High-Speed line at which the system will function without an operating subsidy, as a measure of timing financial sustainability (EBITDA > 0). Additionally, the break-even point in the long term, or discounted payback, could be considered as a relevant indicator to choose among different options for the development of a high-speed network in competition with other transport modes. It can help either on how to schedule investment over time or to evaluate the financial sustainability of an investment. The views expressed in this paper are those of the consultant and do not present an official view of the Spanish Ministry of Economy and Competitiveness (MINECO).

Keywords: Transport planning, investor’s return, financial risk, cost benefit analysis, new business plan, long-term economic sustainability

---

[1] *Maté Sanz, David (Corresponding Author). INECO, Ingeniería y Economía del Transporte S.A. Email: david.mate@ineco.com*
Fernández Gago, Jose Ángel (Financial modelling) INECO, Ingeniería y Economía del Transporte S.A. Email: jose.gago@ineco.com
Chapinal Rivera, Silvia (Quality assurance) INECO, Ingeniería y Economía del Transporte S.A. Email: silvia.chapinal@ineco.com

1. Introduction

Commissioning a high-speed line benefits both the cities and the regions it passes through, as well as the rest of the regions within the country, thanks to its impact overall transport system and the potential intermodality benefits that it implies.

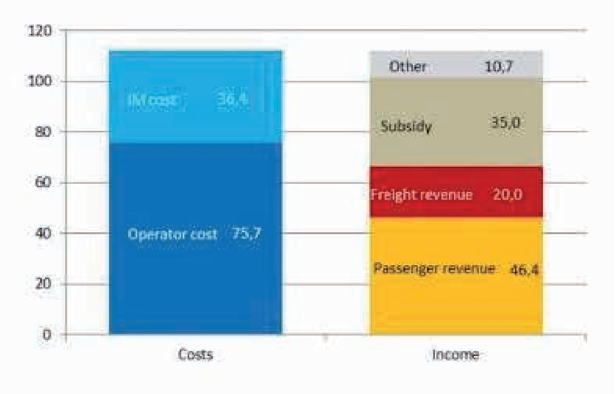
Transport services, along with their infrastructure requirements, occupy a relevant space in the agendas of policy makers. There are many economic methodologies used to evaluate transport infrastructures in general, and High Speed Railway Lines in particular, as a fundamental support in the decision-making processes on which investments and which transport systems are optimal for a country or region.

A commonly asked question among Public Services Senior Management decision makers is which break-even point in high-speed investment would meet financial profitability criteria? Thus, defining the break-even point as the ability to cope with the re-payment of initial investment plus financial return.

2. State of the art

The answer to the question stated above, on whether "to go" or "not to go" for a HSR project based on investors returns and breakeven point depends on numerous elements, profit project vector. As a first approach, they could be synthesized as the cut-off point between the expected revenues once it has started to function against the estimated operating costs. In a more detailed approach, the return required by the financial resources used to make the initial investment should be taken into consideration, as well as the subsequent investments in the expansion of the network capacity and replacement of the assets according to their useful life.

In average, operating costs of the railway system (conventional, freight and HSR lines) rely more on railway undertakings revenues and state subsidies than on the rest of the infrastructure manager's income.



Source: European Commission (2016)

Figure 1. Cost and contribution of the rail sector (billion EUR, 2012 in 2010 prices)

As a result, the break-even point is more precisely defined when complementing calculations with the asset amortization policy or the country's taxation in order to differentiate the effects of indebtedness on the profitability of the different profiles of investors, public of private, that have contributed capital to the project.

Different authors point out that the feasibility of a high-speed line should not be based on deterministic outputs, but on the profitability outcome. Therefore, uncertainty variables should be explicitly considered as a range that would in turn depend on certain scenarios or key indicators or drivers, as it is set out in Article 101 (Information necessary for the approval of a major project), Regulation (EU) No 1303/2013, stating *"a risk assessment must be included in the Cost Benefit Analysis"*, EUROPEAN PARLIAMENT (2013). This is required to deal with the uncertainty that always permeates investment projects. These ranges would be obtained through modelling the profitability generated by the High-Speed project, whether it be socioeconomic or financial, EUROPEAN COMMISSION (2014).

Examples of these drivers or indicators that may have an impact on the break-even point identification, among others, could be the average cost per kilometre of the initial investment and replacements, the anticipated average maintenance costs, the feasible average cost of the ticket, the scheduling alternatives of the investment made and the scenarios of demand considered.

In the process of prioritizing and selecting public investments, the evaluation of a transport project, and particularly of the size of a new high-speed railway line, commonly includes a socioeconomic cost-benefit analysis, which compares the costs and benefits that the project contributes to the whole of society throughout its useful life. In addition, a financial profitability analysis is carried out, allowing the incorporation of the returns generated by the Business Plan of an investment into the decision-making process, regardless of the public or private ownership of its shareholders.

These practises are aligned with the different guidelines that EU Member Countries are setting, such as:

- To draft business plans by infrastructure managers,
- To manage transport companies according to commercial and profitability criteria,
- To implement ex-ante evaluations of investment projects as a support for the decisionmaking process in the public sector
- To identify the degree of financial sustainability, both in the short and long term, depending on the different profiles of investors.

INECO developed between 2015 and 2017 a study for the Spanish Ministry of Economy and Competitiveness (MINECO), titled *"Feasibility Study for a High-Speed Rail between Cairo, Luxor, Aswan and Hurghada in Egypt"*. In the framework of this study, a financial model was developed that has led to the identification of break-even points that guarantee the repayment of financial costs under different investment and demand scenarios. This scenario analysis takes on special relevance in a macro economic, financial and situational context of the real economy such as of Egypt's during this period.

Feasibility Studies are aimed at assessing on the practicality of a proposed project, exploring if it is technically and economically feasible. These studies used to be carried out at the first phase of projects' implementation to support decision-making processes.

The development of the study was entrusted to INECO, a leading transport engineering and consulting company dependent of the Spanish Ministry of Transport and Public Works. INECO has more than 45 years of consolidated experience, contracts in more than 40 countries and over 2.500 employees. It provides professional engineering, consulting and project management services for transport infrastructure projects, giving advice to central, regional and local government and companies around the world.

This paper presents the main conclusions drawn from the specific case of High Speed analysis in Egypt. It is a high-speed railway infrastructure to be developed in a middle-income country, in a corridor highly linked to tourism and in a context of economic uncertainty.

The government of Egypt is considering investing to construct a high-speed railway between Cairo-Luxor and Hurghada. A densely populated corridor in Egypt, with strong economic implications in the Egyptian tourist sector. Several studies have been developed before in Egypt by private and public sector regarding boosting High-speed railway services in the country. Based on Ali (2012) and Alí (2015) studies, it can be observed that the preconditions or success factors for reasonably getting started on a high-speed rail project in Cairo Luxor corridor are:

- Between 25 % and 60 % of the costs of high speed, tracks would in any case be required to build conventional railways with 160 km/h speeds. Therefore, project appraisal should then focus on the incremental costs to achieve this goal.
- The promising indicator of economic potential feasibility is a corridor of a high population density within large cities and high volume of demand with enough economic value to repay the high cost involved in the providing and maintaining the line. It is not only that the number of passengers must be large; however, a high willingness to pay for the new facility is required. Tourist and business passengers in the case of Egypt as law and middle-income country would cover this condition.
- High-speed rail service that can deliver competitive advantage over airlines for journeys of up to about four hours or 700 km, particularly between city pairs where airports are located far from city centers.
- A longer corridor that has a very large number of urban centers located on the line (eg: Cairo- Aswan, 879 km). On these corridors, high-speed rail has the ability to serve multiple cities in Egypt in one line.

3. HSR business case

3.1 Market sizing. Who is the high-speed train’s real competition?

The Area of Study was determined by the corridors affected by the future HSR network under study, namely Cairo-Luxor, Luxor-Aswan and Luxor-Hurghada. Options for intermediate HSR stations location were analysed and discussed with MOT. Transport offer is mainly composed of:

- Road services: private car, shared taxi and bus.
- Rail services (ordinary, express and sleep cars).
- Air services.

The Demand Study provided an O-D matrix for the travel relations between the different areas by transport mode. The analysis of this matrix enabled to extract the main features of current mobility patterns in the Area of study:

- Yearly travel demand along the corridor adds up to 38.5 million passengers (105,000 passengers per day).
- Private car and shared taxi are the most frequent transport modes, representing together a share of 57%. The image below shows the modal split of the different transport modes.

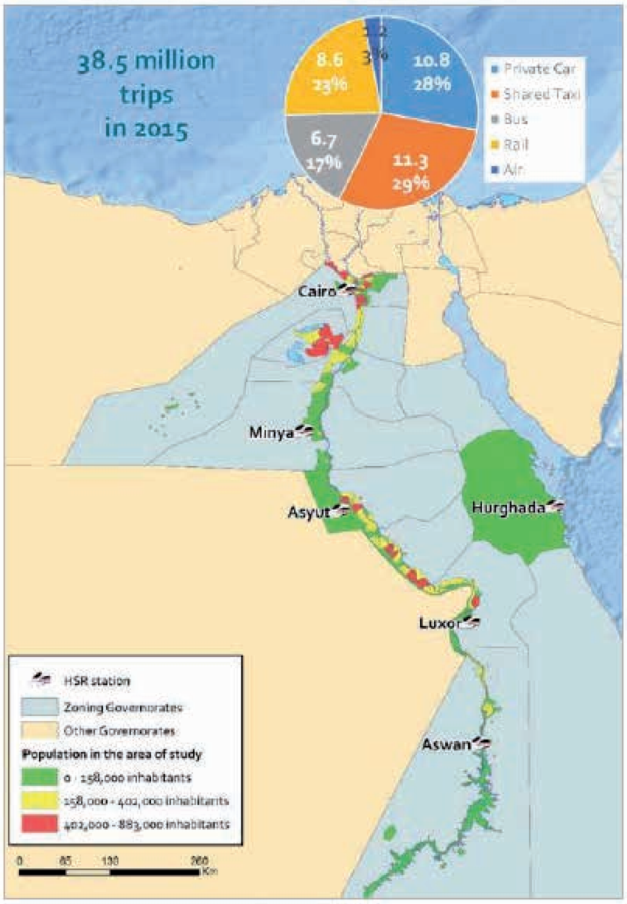


Figure 2. Travel market in the Area of Study (2015)

In terms of travel costs, total passenger demand in the Area of Study represents about 4,100 million yearly Egyptian Pounds. This market size has been estimated with the information obtained from available trip prices and from the consultants' experience in similar studies.

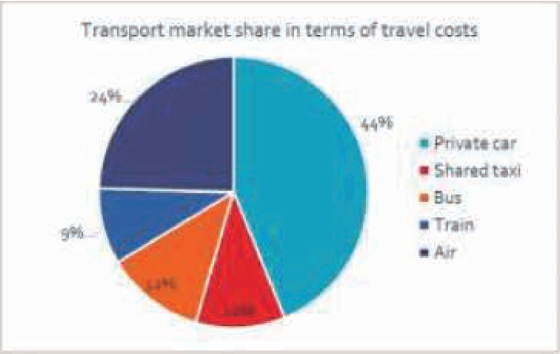


Figure 3. Transport market share in terms of travel costs.

Source: INECO

So, what is the high speed train’s real competition?

- It takes around seven hours to reach Luxor by private car
- It takes more than ten hours to travel between Cairo and Luxor by regular railroad
- The flight time between Cairo and Luxor is three hours
- The high-speed train will take three and a half hours to reach Luxor from Cairo.

Therefore, the real competition for HSR comes from air travel, focusing in tourism travellers and business activity in the area of study.

3.3 Pricing. How much should be charged for high-speed train service?

Based on INECO’s initial studies, the number of captured target customers depends on:

- Competitive door-to-door journey times
- Competitive fare structure
- Attractive services supply
- High comfort standards

However, not everyone will have the means or willingness to pay for a new high speed train service. How much should passengers be charged for high-speed train services? It was also found that they should not be charged more than 800 Egyptian pounds, as corporate and tourism clients are unlikely to be reimbursed a price higher than the flight fare. Therefore, demand studies were performed with a price discount of 35% below air traffic fares.

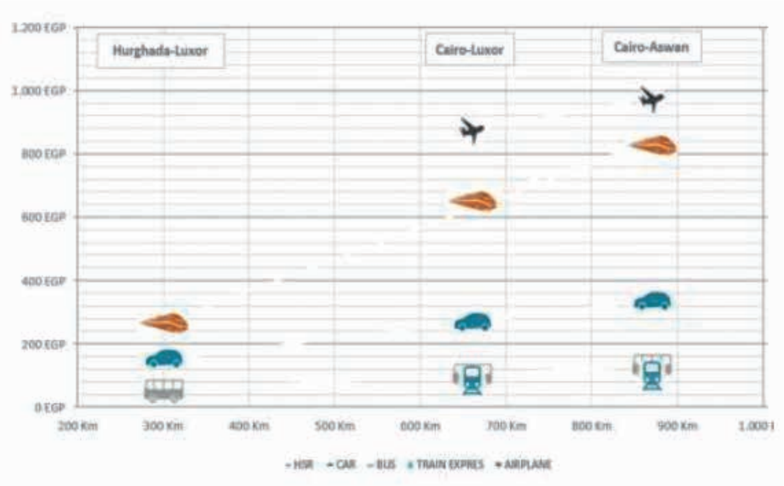


Figure 4. Fare comparison for the different transport modes.

Source: INECO

3.3 Future demand scenarios

Future transport demand scenarios were estimated considering different growth rate assumptions for each demand segment:

- Local mobility.
- International tourism mobility.

Egypt received approximately 10 million international tourists per year at 2015, almost 5 million less than the peak of 14.7 million foreign tourists reached in 2010. Although there is great uncertainty on the recovery of international tourism, it is expected to take place in the medium term.

Since tourism and its related activities are of great importance for the high-speed railway line under study, the analysis considered three different demand recovery scenarios:

- High demand scenario: This scenario foresees a high GDP and population growth with an optimistic tourism growth based on the Ministry of Tourism's objective of reaching 20 million tourists in Egypt by 2020. The annual growth rate obtained by this assumption for the period 2014-2026 is of 9%, which is highly above the expected GDP growth for the same period.
- Medium demand scenario: The foreign tourism levels of 2010 will be regained by 2026, which represents an annual growth rate of 3.4% for the period 2014-2026. The medium scenario growth rate obtained is very similar to the expected GDP growth for the same period.
- Low demand scenario: The foreign tourism levels of 2010 will be regained by 2036, which means an annual growth rate of 1.8% for the period 2014-2026. With regards to GDP, the low demand scenario growth rate obtained is lower than the expected GDP growth rate for the same period.

The following figure shows the forecasts of the HSR demand for the three different scenarios. Section A opening of the Cairo - Luxor section in 2026, Section B opening of the Luxor - Aswan section in 2031 and Section C opening of the Luxor - Hurghada section in 2036.

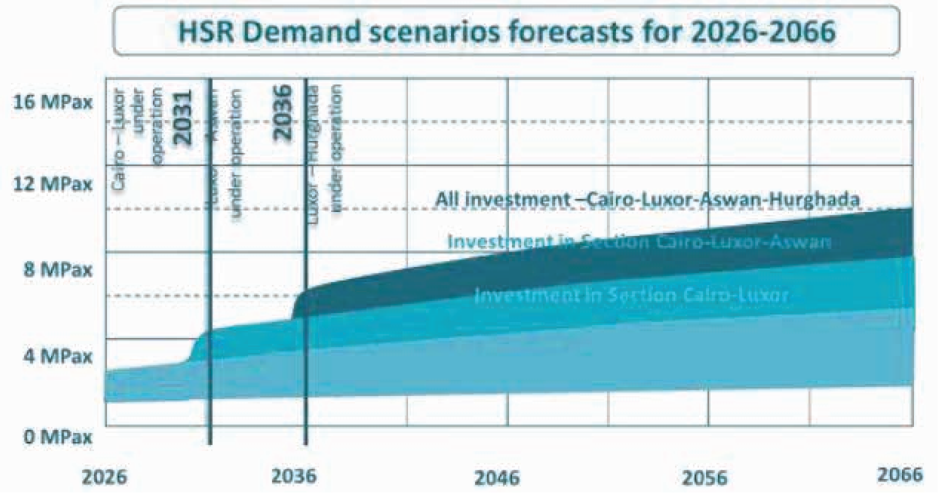


Figure 5. Future demand scenarios in Cairo-Luxor Aswan – Hurghada System

Source: INECO

4. Cost benefit analysis. Project profitability and economic sustainability

To make a “go” or “no-go” decision, it is essential to evaluate the anticipated profitability of the project this being measured by the internal rate of return required by project investors. As stated before, an additional financial project driver is the break-even point: this is considered as the ability to cope with the re-payment of initial investment plus financial return. Yield of this profit project vector will be determined by the profiles of the investors supporting the project.

Therefore, discounted cash flow method (DCF) was used by INECO financial analyst team to assess capital expenditure decisions for this HSR investment. The net present value (NPV) and internal return of return (IRR), widely used in discounted cash flow analysis, were established under different conditions of uncertainty/risk.

In any case, financial and macroeconomic indicators of Egypt will affect the risk appetite and pricing of risk takers under this investment analysis. The main indicators were:

- **2015 Real GDP** still close 4% yearly growth

- **Inflation rate**: Monthly inflation rose to 33 %, March 2017, from 10.35 % on 2015. Prices have been growing steadily from 2006 to 2014 at an average yearly rate of 9%. However, inflation considered was 9.27% according to the historical data from 1960 until nowadays.

- **Sovereign risk - treasury debt**: 10-year Treasury bond in Egyptian pound grew on March 2017 to 17.01 percent from 14.5 percent on 2015.

4.1 Measure of “go” or “no go” decision. Financial indicators

The discount rate to be used in DCF (discounted cash flow method) is the required rate of return based on the investor’s weighted average cost of capital. The weighted average cost of capital (WACC) “may” be adjusted by the risk associated with the project and with risk investors perception. Expected rate of return on equity capital was calculated based on the capital asset pricing model (CAPM), which is the most widely used in risk-return models to calculate the shareholders’ return, and on the Egyptian financial market analysis on a nominal base:

- Private investor profile. Capital market returns: 26.57 % in the case of private sector shareholding, and a priori of 21.74% in the event of a public sector shareholding of the SPV (Special Purpose Vehicle)
- Local institutional investor with Long-term view: 14.91%, which represents the estimation of remuneration of the Egyptian Treasury to the public debt holders in the form of sovereign 10 year bonds (Bonds issued for Suez Channel extension by Government Treasury paid a 12% interest rate)
- International institutional investor profile. Banks and Multilateral or bilateral Longterm debt returns. On this basis, and within the scope of the analysed projects from OCDE and WB databases, it has been assumed that the funding provided by the so-called commercial bank moves around 14.5%, while that provided by multilateral banks moves around 3.2%.

4.2 Project profitability. Financial indicators

A detailed financial model was developed to support the appraisal of the financial performance of the future HSR program as a commercial undertaking. The financial model presents the total financial picture of the future HSR program by compiling all costs and revenues for the evaluation period (construction & operation). An evaluation period of 50 years (from 2021 to 2070) was selected for the financial analysis so that all project cash flows from the project development stage until 2070 were incorporated to the analysis, in line with accepted financial analysis principles.

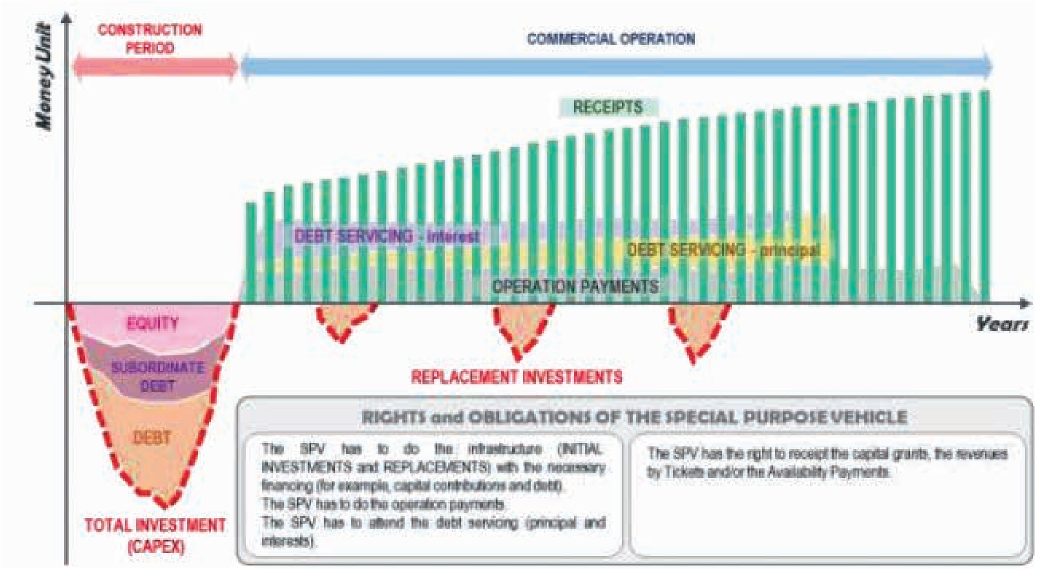


Figure 6. HSR project cash flow financial performance profile

Source: INECO

Mathematically, the **Project IRR** represents the discount rate that makes the net present value of all economic cash-flows of the project equal to zero. The higher the IRR, the greater the net economic returns achieved by a project relative to its capital resource costs and if IRR is greater than the discount rate, then the project delivers a positive net economic benefit (NPV).

Overall, the results of the financial analysis present:

- **A small** positive Project IRR for a medium-income country as Egypt ranging between **2.21% and 11.03%** depending on the infrastructure and demand scenario.
- A positive economic case (**positive Project NPV**) for the introduction of HSR only in the **high demand scenario. A project expected IRR of 7.46%** would be necessary for Capital Investors, without liquidity project constrains, considering a scenario of strong funding by multilateral agencies.

Considering the three different investment options, these being: 1: Cairo-Luxor, 2: Cairo-LuxorAswan, and 3: Cairo-Luxor-Hurghada, the main results derived of the financial assessment modelling, with incomes from HSR service tickets and an additional 5% from commercial activities are the following:

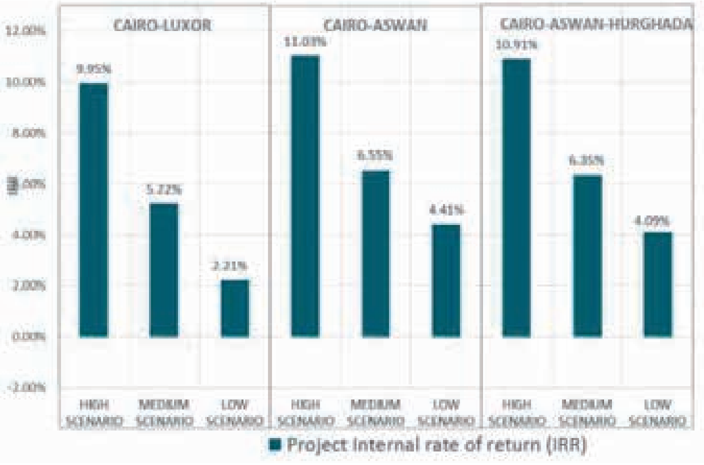


Figure 7. Financial assessment – Results

Source: INECO

Infrastructure yearly investment cash flow fluctuate from yearly 2016 € 0.65 billion in 2021 to 2016 € 1.1 billion in 2024 in Alternative 3 option (construction of all project phases):

HSR PHASE	Section	Total Investment	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035
CONSTRUCTION PHASE	A: Cairo-Luxor	M€ 4,452	60 months														
	B: Luxor-Aswan	M€ 928									24 months						
	C: Luxor-Hurghada	M€ 1,114														24 months	
		M€ 6,494															
% of investment to total amount of the section			15%	15%	20%	25%	25%				45%	55%				45%	55%

This IRR objective would only be achieved on the basis of a "project finance" structure with low equity requirements.

For the case of No contribution from multilateral resources, none of the simulated scenarios would be financially sustainable, although the Cairo - Aswan corridor would be the strongest financial performing HSR program. In this case, the project IRR would be of 10.77%, below the commercial bank interest rate approach, and over the average inflation rate of the last decade. The last expansion of the Canal the Suez by Egyptian Government pays an interest of 12% for their corporates bonds. More than 80% of that total investment came from the Egyptian public's purchase of state-issued bonds.

If we were to compare this project's IRR with the one requested in the USA, the project would not be profitable enough to reach the 11.12% return on capital demanded by the United States Surface Transportation Board. The Railroad Cost of Capital 2012[1] found that the weighted average cost of capital for US railways in 2012 was fixed onto 11.12% (22.56% debt capital X 3.29% return on debt) + (77.44% equity capital x 13.40 return on equity) (Lawrence, Martha & Ollivier, Gerald, 2014). This rate was 8.8% for 2016, which would position the project under profitable standards in the high demand scenario (SURFACE TRANSPORTATION BOARD 2016). In

---

[1] *The cost of capital figure represents the Board's estimate of the average rate of return needed to persuade investors to provide capital to the freight rail industry. Calculated annually, it is an essential component in evaluating the adequacy of an individual railroad's revenues each year. US also uses the figure when determining the reasonableness of a challenged rail rate, considering a proposal to abandona rail line, or valuing a particular railroad operation.). In this paper, it is only used as analytical official reference, taking into account that our corridor initially is only for passenger services.*

the case of the California HSR, the discount rates used in the 2016 Business Plan were 8%, 11% and 14%, depending on the range of risk transfer that will be finally allocated in the future contracts (California High-Speed Rail Authority 2016).

If the project was to be delivered by a European government, the official reference of a minimum project IRR would be the European Commission general recommendations in an scenario of low inflation where the requested IRR would be ranged around 4% (EUROPEAN PARLIAMENT, 2013). In that case, the project would be profitable enough in the medium demand scenario too.

4.3 Measure of "go" or "no go" decision. Break-even point of Cairo-Luxor project (Without Hurghada phase)

INECO feasibility study drafted several investment infrastructure development alternatives that have been evaluated, in order to evaluate the suitability, or lack of suitability, of making this investment in the country. From a financial analysis point of view, the report provides the level of demand, or income, from which the project has found its break-even point in the short and medium term, complemented by the environment conditions in which this circumstance has been estimated.

The objective of setting a Break-even point is to ensure that the **HSR project** is **financially sustainable**, which means that the company will neither experience liquid assets problems nor temporary decapitalisation processes or bankruptcy.

Investors will have to carefully weigh the costs and benefits of the proposed high-speed railway project, and the expected future volatility or evolution (risk analysis).

A **"Profits = Revenues - Costs"** framework seems the right way to start with:

Profits = Revenues - Costs	
Revenues	**Costs**
• **Price**	• Fixed Costs
• **Volume**	• Variable Costs

The break-even point may be measured by a two-pronged approach:

- **Break-even in the short term as the point at which the system will function without an operating subsidy.** From a liquidity perspective, it is the point from which earnings before interests, taxes and amortization should start to be positive (EBITDA > 0). Whereas from an operational point of view, it quantifies the number of passengers or average annual income that would be necessary for the line to obtain operational profits in a given year. A yearly approach would show the turnover available to return and maximize margin on variable costs over which fixed costs will be charged. Based on the margins calculation in the short term, we can establish when sales are able to cover and ensure fix investments and fixed project costs, and when a HSR service or corridor is unsuitable and produces loss to the company.
- **Break-even in the long term**, from a solvency perspective, would be measured by the cumulative number of passengers or necessary income for the line to obtain a return on capital investment that considers the time value of money. This cash flow is commonly updated by the annual variation of the Consumer Price Index or the Gross Domestic Product

deflator, if this data is available in the country where the investment is being analysed. Therefore, the project is judged by its ability to cover non-distributed costs (fixed costs) in a long period of time (payback taking into account Price Index forecast, discounted payback).

For instance, given the **investment alternative** in the **Cairo-Luxor-Aswan** has fixed costs of EUR $5.4 billion for infrastructure, and the Egyptian HSR Railway company shall procure for either 9 or 24 trains depending of demand scenarios, the total variable yearly operation and maintenance costs will be around EUR 575,000 per day.

It would take 0 years to achieve the breakeven point in the short term for medium demand scenario (year 2026). The Gross margin will have yearly variations between 23% of 2026 revenues to 60% in 2070. This project's break-even point in the short term indicates that no operating subsidy will be required under the high or medium demand scenarios.

Based on financial considerations and taking into account yearly operation and maintenance costs as variable costs, this project will be eventually profitable, but it will need to take 40 years to breakeven in the long term for the medium demand scenario in the Cairo- Luxor Investment Alternative. This would be taking into consideration a discount rate of the average Inflation rate of 9.25 %. In the low demand scenario, the discounted payback or break-even point will not be achieved before 2070, the financial analysis period.

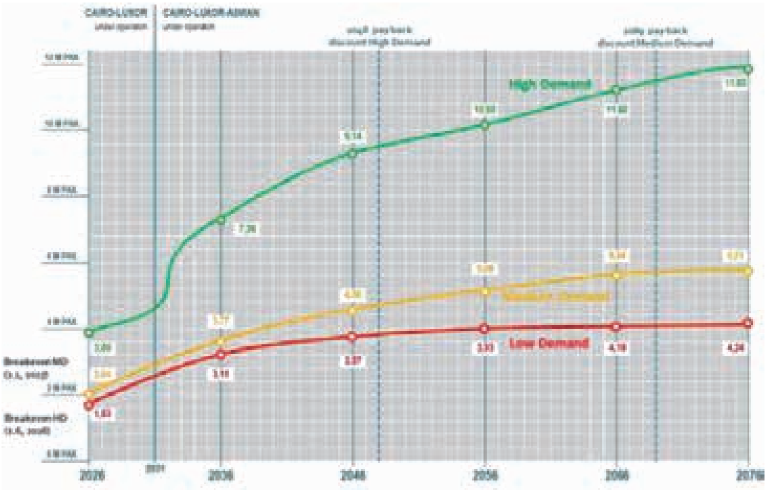


Figure 8. Investment option 2: Cairo-Luxor-Aswan Breakeven points (payback discount)

This is consistent with the results of other high-speed rail projects across the world, like the California HSR project, where the discounted payback period for the total invested capital is 36 years from operations commencement and 45 years from the start of construction. In this case, payback means the break-even point in the long term, and it is expected that collected net cash will equal total expended cash for capital in 2057 (California High-Speed Rail Authority 2012).

Whether this is a "go" or "no-go" decision is totally up to the decision makers. This gross margin has been proved not to be enough to cover all fixed costs including the recovery of capital investment in the long term for medium demand in the Cairo Luxor Aswan infrastructure investment alternative.

5. Conclusions

The Feasibility Study for a High-Speed Rail between Cairo, Luxor, Aswan and Hurghada in Egypt evaluated the operational viability, cash flows, profitability and breakeven for the project.

The analysis has been developed for three different investment scopes. For each of these options, and before any consideration of financing or any particular source of funding, various indicators for the project have been obtained: an internal rate of return, NPV and breakeven points for different periods, in the short and long term. The IRR for the project has been estimated to fluctuate between 4% and 12% depending on the selected investment alternative and demand scenario.

This total project return would be insufficient to attract private capital to pay for the entire project. These results illustrate that the project can generate a positive net cash flow from operations but that it would require government or multilateral funding for construction.

Whether this is a “go” or “no-go” decision is totally up to the decision makers. The breakeven point of the investment alternative Cairo Luxor Aswan proves that yearly gross margin will not be enough to cover all fixed cost including the recovery of capital investment in the long term in the case of such a long contract as that of 50 years.

The project profit vector (IRR; NPV, Breakeven in the short-term, breakeven in the long-term) will provide decision makers with additional approaches to set up “timing strategies” in the HSR Business Plan under different risk scenarios. This would allow to escalate and prioritize investment alternatives and sources of capital for the funding of each section of the project.

This analysis approach is similar to the one followed for the funding and financing analysis of the California HSR project. In this case, the timing strategy of the expansion system was correlated with the finance transactions / indicators. Prioritized investments according to their capacity to generate revenues was expected to attract private sector participation and extend the system beyond the initial line, so that additional segments can be constructed as new funding becomes available (California High-Speed Rail Authority 2016).

Breakeven indicators highlight how the project profitability changes if it is delivered earlier or later. This will help decision makers to consider the edge between an impressive infrastructure project and a possible white elephant project if it was to be deliver too early. The timing indicator for high fixed costs in public investments can be presented by the breakeven point as a way to check how relevant variables can affect the profitability of a project (ie: interest and inflation rates, investment overruns and delays, demand shortcuts).

It is worth noting that this analysis focuses in the purely financial aspects, leaving aside the fact that marketing strategies in competitive markets can always shift the balance of a business plan with actions such as lowering the prices to increase sales volume, direct marketing to educate clients or heavy promotion to increase market penetration.

Finally, from the decision-making point of view, other interesting technical points suggested to be considered for future discussion would be:

- Could it have any effect on the final profitability of the project to add a project profit vector analysis (IRR; NPV, Breakeven in the short-term, breakeven in the long-term) to the decision-making process?
- Should public investors follow either a proactive or a reactive timing strategy and are there any financial indicators that will allow them to create, monitor and react to a predictable rhythm for expansions?

- How could it be avoided that a failure of a big infrastructure project damaged the national budget rating and therefore the country image because of the irreversibility of the public investment?

- Future growth options are often the main argument for undertaking unprofitable investments or speeding up projects. In circumstances when the uncertainty is high, could the right timing save from some bad outcomes?

- Are decision-makers considering the use of a dynamic investment approach based on the monitoring and control of the project profit vector? Is the Business Plan consequently updated regularly in order to assure social and financial profitability considering unexpected events along the different stages of the project?

6. References

- Ali, Mahmoud Ahmed Mousa (2012). “Opportunities for High-Speed Railways in Developing and Emerging Countries: A case study Egypt”, PhD Thesis, Technical University of Berlin – Germany.

- Ali, Mahmoud Ahmed Mousa (2015). “Lessons for Policy Makers in Non-High Speed Rail Countries: A Review”.

- California High-Speed Rail Authority (2012) California High-Speed Rail Program. Revised 2012 Business Plan.

- California High-Speed Rail Authority (2016) 2016 Business Plan.

- EUROPEAN COMMISSION (2014). Guide to Cost-Benefit Analysis of Investment Projects. Economic appraisal tool for Cohesion Policy 2014-2020.

- EUROPEAN COMMISSION (2016). Fifth report on monitoring developments of the rail market.

- EUROPEAN PARLIAMENT (2013) REGULATION (EU) No 1303/2013 of 17 December 2013 laying down common provisions on the European Regional Development Fund, the European Social Fund, the Cohesion Fund, the European Agricultural Fund for Rural Development and the European Maritime and Fisheries Fund and laying down general provisions on the European Regional Development Fund, the European Social Fund, the Cohesion Fund and the European

- Maritime and Fisheries Fund and repealing Council Regulation (EC) No 1083/2006

- SURFACE TRANSPORTATION BOARD (2016). Decision determining the railroad industry's cost of capital for 2016. US Government. Docket No. EP 558 (Sub-No. 20).

- Lawrence, Martha & Olivier, Gerald (2014). Private Capital for Railway Development. World Bank.

7. Acknowledgements

The author sincerely thanks the Spanish Ministry of Economy and Competitiveness (MINECO) for providing the necessary facilities and opportunities to prepare this paper, and RENFE and ADIF staff commentaries to Egyptian HSR marketing and income forecast during MINECO study development.

In addition, thanks to the INECO's technical team who developed the full study, Jose Sesma as Business Manager, Carolina Sanz as Project manager, and the business plan, marketing studies, financial risk analysis and cost benefit analysis specialists that the corresponding author coordinated during 2016 to finally archive the Report 4: Socioeconomic and financial assessment & Management models.

The author would also like to thank the reviewer Silvia Chapinal for her valuable quality review that helped to improve this paper and the financial modelling that Jose Angel Fernandez Gago performed to present specific outputs for this paper.

The views expressed in this paper are those of the consultant and do not present an official view of the Spanish Ministry of Economy and Competitiveness (MINECO).

360.revista de alta velocidad

360. revista de altavelocidad pretende servir de foro de discusión serena y plural, a la vez que profundiza en todos los temas relacionados con la alta velocidad ferroviaria: planificación, efectos económicos y sociales, explotación, tecnología, etc.

Junio 2018
número 5

INTERNATIONAL CONGRESS ON HIGH-SPEED RAIL (Vol. I)

Mobile Material

High Speed Rail Superstructure

Safety and Signalling Systemas

High Speed Rail Infrastructure

Economic Impacts

Fundación de los Ferrocarriles Españoles