

An Assessment of Railway Capacity

M. Abril,^{a,*} F. Barber,^a L. Ingolotti,^a M.A. Salido,^a
P. Tormos,^b A. Lova^b

^a*Department of Information Systems and Computation*

^b*Dept. of Applied Statistics and Operational Research, and Quality
Technical University of Valencia*

Camino de Vera s/n, 46022, Valencia, SPAIN

Abstract

In this paper, we review the main concepts and methods to perform capacity analyses, and we present an automated tool that is able to perform several capacity analyses. Capacity is extremely dependent on infrastructure, traffic, and operating parameters. Therefore, an in-depth study of the main factors that influence railway capacity is performed on several Spanish railway infrastructures. The results show how the capacity varies according to factors such as train speed, commercial stops, train heterogeneity, distance between railway signals, and timetable robustness.

Key words: Capacity, Railway, Timetable, Train, Scheduling.

1 Introduction

One consequence of the globalization of the economy and the increasing integration of the international economies is a considerable growth in the entire transport sector. During the 1990's, many countries began to suffer from congestion in certain areas and on certain routes. Nowadays, there is no doubt about the congestion of the transport situation in some countries. The problem is now beginning to threaten economic competitiveness. Greater economic development cannot take place in the current transport scenario unless ambitious measures are taken. Revitalizing the railways is one of the principal measures

* Corresponding author. Tel.: +34 963877000 (ext:73550); fax: +34 963877359
Email addresses: mabril@dsic.upv.es (M. Abril), fbarber@dsic.upv.es (F. Barber.), lingolotti@dsic.upv.es (L. Ingolotti.), msalido@dsic.upv.es (M.A. Salido.), ptormos@eio.upv.es (P. Tormos.), allova@eio.upv.es (A. Lova).

proposed in European transport policy. The priority is to open up markets with the deregulation of the rail transport sector as one of the principal aims. The objective is to provide railway companies with access to the railway network on equal terms; this access will be determined by infrastructure managers. This has led many railroads to reevaluate their capacity. Capital expansion is a very costly means of increasing capacity. A more cost-effective solution is to manage the existing capacity more effectively using computer-based decision support systems.

Optimizing the use of railway infrastructure is a complex and difficult task. Therefore, numerous capacity studies must be performed in order to work out what part of extra traffic can be absorbed by the existing infrastructure and how much investment will be required for new infrastructure. The results of these studies must be rapid and precise in order to know how many route slots can be offered to railway operators and how much railway traffic can be supported by the current network. They must also provide regional and national authorities and the owners of the infrastructure with information that proves that investing in the development of the rail network is necessary and worthwhile financially.

Capacity, whose definition is a classical problem, has long been a significant issue in the railway industry. The goal of capacity analysis is to determine the maximum number of trains that would be able to operate on a given railway infrastructure, during a specific time interval, given the operational conditions. Numerous approaches and tools have been developed to address this problem; they are based on traffic patterns (Forsgren, 2003), single-track analytical models (Petersen, 1974), or algebraic approaches (Egmond, 1999). Several international companies are also working on similar computer-based systems:

- DEMIURGE (SNCF and Eurodecision, 2004) is a software program designed to assist in making rail network capacity studies. This software can evaluate a network's capacity to absorb additional traffic, to locate bottlenecks, to assist in making decisions about infrastructure investments, to optimize current and future timetables, and to calculate the residual capacity of a timetable.
- CMS (AEA Technology Rail) provides a system to plan the effective utilization of the railway capacity. It offers an easy "what-if" scenario evaluation, automatic generation of timetables, simulation of operations to predict performance and identify remedies, identification of capacity available for sale, and usage forecasts based on improved timetables. However, CMS needs to be calibrated using updated punctuality data to ensure that its predictions are valid.
- RAILCAP (Stratec) measures how much of the available capacity is used by a given operation program in a straightforward way, and it offers a very de-

tailed analysis of bottlenecks. However, it has one major disadvantage since the modelling requires a great deal of effort. RAILCAP requires detailed descriptions of the tracks, switches, crosses, signals and speed limits.

- VIRIATO (SMA and Partner) is mainly used for adapting infrastructure to future service concepts and coordinating several operators or products that share the same infrastructure. It allows the user to determine the amount of saturation of a specified line. It compresses a given timetable, and determines the saturation rate of a line or a part of a line as a percentage. This method leads to varying results for the same line, depending on the length of the section under consideration.
- CAPRES (Lucchini and Curchod, 2001) is a model for the elaboration and saturation of timetable variants. Through the use of iterations, this model determines all available extra trainpaths, given all the constraints and interconnections between lines. A disadvantage of this model is that the traditional network and operational data have to be completed with the information about where, when and how the network capacity must be used.
- FASTTRACK II (Multimodal Applied Systems) is a computer-based train dispatching and meet-pass model that is capable of producing a feasible train dispatching plan for a user-selected corridor, given a set of proposed train schedules and a corridor's track configuration. It can be used to examine the feasibility of a set of proposed train schedules, test the impact of proposed changes in operating policies on train service, and measure both the theoretical and practical line capacity.

More information about these systems and other similar railway management systems can be found in (Barber et al., 2007).

Previous works have shown that several factors can affect capacity, and several methods have been proposed to assess it. However, since there is no commonly accepted measure of capacity, additional analytical analyses of these factors following recent recommendations of the Union of the European Railway Industries are needed. There is also a need for automatic tools that analyze different empirical methods.

In this paper, we review the main concepts and methods of assessing railway capacity, and we analyze the main factors that can affect it. In addition, we present a computer-based tool, the MOM system (Barber et al., 2006a), which has been developed along the same lines as the above-mentioned tools. The MOM system is a highly functional tool that helps railway managers to provide efficient and reactive management of railway infrastructures. The MOM system can generate optimized railway schedules both off-line and on-line (when disruptions occur). It also provides information on railway network capacity and on timetable robustness, helping managers to make better decisions. In this work, we limit our study to the railway capacity module. This module provides several analytical and empirical methods that can be used

to obtain conclusions about the capacity of railway networks and that support the process of adapting the railway infrastructure to traffic needs. The MOM system project has been developed according to the requirements of the Spanish Administration of Railway Infrastructure, ADIF¹.

Section 2 presents the problem description: an introduction to railway capacity, types of capacity, and the factors that affect capacity. Section 3 presents a survey of methods to evaluate railway capacity. Section 4 presents our automated tool that is able to obtain conclusions about railway capacity. In section 5, we perform several analytical and empirical capacity analyses on real networks. In section 6, we analyze the effect of different parameters on railway capacity. Section 7 presents the new signaling and management system for the European Rail Network, and Section 8 presents our conclusions.

2 Problem Description

An efficient utilization of the existing railway infrastructure is an essential component of a high-quality transportation system and has become a central task for railway infrastructure managers. The definition of standards and robust methods for its assessment are very important. Line capacity is, in essence, what the infrastructure managers have to sell as their final product.

Although capacity seems to be a self-explanatory term in common language, its scientific use may lead to substantial difficulties when it is associated to objective and quantifiable measures. It is a complex term that has numerous meanings and for which numerous definitions have been given. When referring to a rail context, it can be described as follows:

“Capacity is a measure of the ability to move a specific amount of traffic over a defined rail line with a given set of resources under a specific service plan.”
(Krueger, 1999).

2.1 Types of Capacity

Different types of capacity are usually used in the railway environment:

- **Theoretical Capacity:** It is the number of trains that could run over a route, during a specific time interval, in a strictly perfect, mathematically generated environment, with the trains running permanently and ideally at minimum headway (i.e. temporal interval between two consecutive trains).

¹ <http://www.adif.es>

It is an upper limit for line capacity. Frequently, it assumes that traffic is homogeneous, that all trains are identical, and that trains are evenly spaced throughout the day with no disruptions. It ignores the effects of variations in traffic and operations that occur in reality. Theoretical Capacity is calculated using an empirical formula. This number is relatively easy to generate (it depends on the longest distance between crossing stations for single-track lines or the minimum headway for double-track lines). It is not possible to actually run the number of trains that can be worked out mathematically.

- **Practical Capacity:** It is the practical limit of “representative” traffic volume that can be moved on a line at a reasonable level of reliability. The “representative” traffic reflects the actual train mix, priorities, traffic bunching, etc. If the theoretical capacity represents the upper *theoretical* bound, the practical capacity represents a more realistic measure. Thus, practical capacity is calculated under more realistic assumptions, which are related to the level of expected operating quality and system reliability, as shown in Figure 9. It is the capacity that can permanently be provided under normal operating conditions. It is usually around 60%-75% of the theoretical capacity, which has already been concluded by Kraft (1982). Practical Capacity is the most significant measure of track capacity since it relates the ability of a specific combination of infrastructure, traffic, and operations to move the most volume within an expected service level.

FIGURE 1

- **Used Capacity:** It is the actual traffic volume occurring over the network. It reflects actual traffic and operations that occur on the line. It is usually lower than the practical capacity.
- **Available Capacity:** It is the difference between the Used Capacity and the Practical Capacity. It is an indication of the additional traffic volume that could be handled in the route. If it allows new trains to be added, it is a useful capacity; otherwise, it is lost capacity.

2.2 Parameters of Capacity

Railway capacity is not static. It is extremely dependent on how it is used. The physical and dynamic variability of train characteristics makes capacity dependent on the particular mix of trains and the order in which they run on the line. Furthermore, it varies with changes in infrastructure and operating conditions. These are some of the fundamental factors that affect it:

- **Infrastructure Parameters**

- **Block and signalling system:** The signals help extend the train driver’s visibility, so it allows greater speeds. The role of signalling is to keep trains at a safe distance. There are two types of systems: fixed and moving block signaling. In a fixed block signaling system, the position of each train is

known only by the block section(s) that it occupies. The separation between trains is maintained by imposing the restriction that each block section be occupied by at most one train at a time. Block section lengths, train speeds, and train lengths are, therefore, important parameters. In a moving block signalling system, which is a modern technology, the position of each train is known continuously, thus permitting better regulation of the relative distances. This requires an efficient communication system between line signals, cabs, and control centres. In addition, speeds that can be achieved by trains can be limited in practical situations by different factors other than track geometry and signalling, such as catenary or power traction constraints. In section 7, we present the new European signalling system called ERTMS.

- **Single/double tracks:** This has a major impact on capacity. It is not as simple as multiplying the number of tracks: two tracks usually have around four times more capacity than a single track; however, a four-track line rarely increases capacity by more than 50% over a double line (Kittelsohn and Associates, 2003). Furthermore, adding a second track may not eliminate the problem because the station is the real bottleneck. Auxiliary tracks at crossing stations also increase the capacity of a single track because they allow trains to carry out crossings and overtakings.
- **Definition of lines, routes:** In assessing the line capacity, the railway line itself must be defined, namely the list of stations and halts along the line and their characteristics. The line capacity should develop into a network capacity problem as soon as other parallel routes are available to provide the same kind of services. This is specifically the case of freight transport on some main international corridors. In this case, the total capacity can be estimated by summing up the contribution of independent lines and finding alternate route sections to overcome bottleneck problems that are detected on the main line.
- **Network effects:** A single line cannot be considered as a fully independent part of the whole network due to crossing and overlapping lines, which can be true bottlenecks. As a consequence, the capacity of a line cannot be defined without considering what happens on the interfering lines.
- **Track structure and speed limits:** The condition of the rails, ties, and ballast dictate the weight and type of equipment that can be used on the line, as well as the speeds allowed on the line. They have an important influence on the capacity. The speed limits are regulated by means of speed profiles, which take into account physics, safety, comfort, train types, etc.
- **Length of the subdivision:** The track connection between two dependencies is a line section. Lengths of line sections between stations (where crossing and overtaking are allowed) are important, mainly, on single track sections. In general, if the length of the subdivision increases, so does the transit time of trains. In this context, the concept of “bottleneck” (or the most constrained line section, e.g., the longest one along the line) is tra-

ditionally introduced as the bounding factor to the overall capacity. In general, the traffic flow can potentially increase and be better managed as the line sections between the stations become shorter, and more auxiliary tracks are available at the station to allow crossing and overtaking operations.

- ***Traffic Parameters***

- **New or existing lines:** The solution is generally different when new (i.e., to be designed) or current (i.e., available) lines are considered. In the second case, which is by far the most common, several constraints are already set, and the traffic to be accommodated often becomes an available capacity problem, which is more constrained.
- **Train mix:** Railway capacity is very much dependent on traffic mix. The ideal case is when all trains are the same or have the same speed. As the mixture of different trains increases, more interferences are generated. Thus, trains require overtakings and crossings, and the mixture reduces the traffic flow. Besides maximum speed, other rolling stock characteristics such as acceleration and deceleration are also important.
- **Regular timetables:** A specific target of the line planning and train scheduling problems is the construction of periodic timetables, which means train services of the same class departing at fixed intervals. It often represents a strong commercial constraint for train scheduling.
- **Traffic peaking factor:** It is a measure of concentration of traffic within a short time frame. It has a significant impact on capacity because it can result in traffic levels that are higher than the network could reliably sustain. When this occurs, the effects are felt for considerable time periods into the future as the system recovers from the overload.
- **Priority:** The priorities of trains play a vital role. Train priorities decrease capacity because priority trains are given preferential treatment over lower priority trains, which results in increased delays. This basically allows the priority traffic to move as if it were the only traffic in the network. As a rule, the greater the number of priority classes, the less capacity is available.

- ***Operating Parameters***

- **Track Interruptions:** These can be planned events (maintenance times, commercial stops, etc.) and unplanned events (train failures, broken rails, etc.) that take a track out of service. These interruptions affect the capacity of a line because they directly reduce the number of hours available in the day to move trains. Track outages usually occur due to maintenance, which is scheduled to minimize train impact; however, it can also be attributed to infrastructure and train failures such as broken rails, accidents, etc.
- **Train stop time:** This parameter accounts for the amount of time trains spend stopped on a line. It is a delay that directly increases the amount of time a train takes to traverse a subdivision.
- **Maximum trip time threshold:** This parameter represents the upper

limit of acceptable trip time. It affects capacity because it restricts the number and/or length of the train stops.

- **Time window:** The interval or unit of time that is taken as a reference for computing the desired line capacity should be defined. Traditionally, this is set either to one hour or to the whole working day.
- **Quality of service, reliability, or robustness:** Since train operations are not perfect, and random minor disturbances and failures occur in the real management of trains during the railway operation of certain lines reducing the theoretical capacity, some buffer times must be taken into account in order to design a robust timetable. This stochastic effect is often difficult to take into account in line capacity evaluations. However, a balance between service reliability and maximum physical capacity is needed to find the economically optimal level of capacity. In the 1970's, Akyilmaz had already carried out some trade-off analyses of alternative transportation plans (Akyilmaz, 1976), more recently, Landex et al. have also dealt with this subject in their work (Landex et al., 2006). Figure 9 shows how the average delays increase exponentially when the number of trains exceeds the saturation level and how network reliability is lost. Therefore, increasing the number of trains over the *saturation level* is not profitable.

FIGURE 2

3 Methods to Evaluate Railway Capacity

Numerous approaches have been developed to evaluate railway capacity. The most relevant methods can be classified in three levels: Analytical Methods, Optimization Methods, and Simulation Methods.

- **Analytical Methods:** These are very simple models aimed at determining a preliminary solution. These methods can also be used for reference or comparison. They are designed to model the railway environment by means of mathematical formulae or algebraic expressions. They usually obtain theoretical capacities and determine practical capacities either as a percentage of the theoretical capacity or by including regularity margins when they calculate the theoretical capacity. An example of these methods, which perform probabilistic analyses of dispatch patterns, was presented by Petersen and Taylor (1982) and was subsequently extended by Kraft (1988). Martland (1982) introduced the PMAKE analysis, which aims to understand the probabilistic distribution of time that a train spends during its intermediate stops on its way to its destination. A more recent reference about analytical methods is Malaspina and Reitani (1995), whose work discusses how to compute the capacity of a railway line and compare their results with previous ones. This method accounts for possible delays and introduces probabilistic

priority coefficients for all combinations of train types.

The International Union of Railways, more generally known as the UIC (from its French name, Union Internationale des Chemins de fer), proposed the UIC method (UIC, 1983); it calculates capacity in line sections to identify bottlenecks. It takes into account the order of trains, and a buffer time is inserted to achieve an acceptable quality of service. This method was officially dropped some years ago and is no longer recognized as a standard. It has been superseded by more general recommendations (UIC, 1996) that establish a link between railway capacity and railway quality.

Petersen (1974) presents a single-track analytical model assuming that the departure times are uniformly distributed over the time period of interest and that three distinct train speeds are possible. Canciani (1991) proposes a method based on train types of two speeds, with simplifying assumptions about the line (stations regularly spaced over the line), which offers good insight into the problem of diagramming “high” and “low” speed trains. Among the analytical methods are the attempts to estimate the delays of a given traffic flow by a stochastic approach, as reported in Chen and Harker (1990). The model was enhanced by Harker and Hong (1990) to include a partially double-track corridor.

One of the most recent references about railway capacity is Burdett and Kozan (2006). They develop several approaches to calculate theoretical capacity (they call it “absolute capacity”) for railway lines and networks. These approaches take numerous railway aspects into account; for example, mix of trains, signal locations, or dwell times. They are based on given train proportions. These proportions are used to weight the average sectional running time. This time is the key to determining the lower and upper bounds of capacity on the critical section. Therefore, these approaches are based on the concepts of critical section and train proportions, which are also introduced in the UIC method (UIC (1983), UIC (1996)).

Analytical methods for computing railway line capacity may be a good start for identifying bottlenecks and major constraints; however, analytical results vary from one method to another depending on what type of parameters they model. Furthermore, analytical models are very sensitive to parameter input and train mix variations.

- **Optimization Methods:** They are designed to provide more strategic methods for solving the railway capacity problem and provide much better solutions than purely analytical formulae. Optimization methods for evaluating railway capacity are based on obtaining optimal saturated timetables. These optimal timetables are usually obtained by using mathematical programming techniques (Mixed Integer Linear Programming formulations and Enumerative algorithms). A particular method of optimization is Saturation. This method obtains line capacity by scheduling a maximum number of additional train services in a timetable (starting with either an empty timetable or with an initial base timetable).

References before the 1980’s can be found in the survey by Assad (1980).

Szpigel (1972) is the first to propose a branch-and-bound algorithm for train scheduling on a single-track line, given the train departure times. He considers a variant of the models mentioned above in which the order of the train departures from a station is not represented by binary variables but by disjunctive constraints. The problem is then solved by branch-and-bound for small size instances by computing bounds through the relaxation of these disjunctive constraints. Only in the 1990's does the problem seem to have attracted the attention of operations researchers. In particular, Jovanovic and Harker (1991) solve a mixed integer linear programming model, calling for a feasible schedule rather than for the optimization of a suitable objective function using branch-and-bound techniques. Timetable scheduling that also considers economic impact (i.e. the determination of the values of train paths assigned to several operators willing to use the same infrastructure) is studied by Harker and Hong (1994), who introduce a computable equilibrium model of an internal market for track resources following a game theoretic approach. Cai and Goh (1994) illustrate a constructive, greedy heuristic driven by one of these models. Carey and Lockwood (1995) describe a heuristic algorithm for double-track lines. It is a heuristic that considers the trains one at a time (in appropriate order) and solves a mixed integer linear program for each train in order to schedule the train optimally, keeping the path of the previously scheduled trains partially fixed. Higgins et al. (1996) define local search, tabu search, genetic and hybrid heuristics, and they find feasible solutions with these methods.

More recently, Oliveira and Smith (2000) model the problem as a special case of the Job-Shop Scheduling Problem. They consider trains to be jobs to be scheduled on tracks that are regarded as resources. They present a hybrid algorithm devised under the Constraint Programming paradigm and show how to adapt this framework in some special, real-life applications. Different models that are based on a graph theoretical representation of the problem are presented by Caprara et al. (2001). In this work, times are discretized and expressed in minutes, and Lagrangian relaxation is used to derive bounds on the optimal solution value as well as to drive a simple heuristic procedure. In the framework of the EUROPE-TRIS project, two scheduling algorithms were developed: FLOU (Flow Line Optimal Utilization), which attempts to integrate the maximum trains at the minimum cost in a given timetable; and TCM (Traffic Capacity Management), which searches for the best solution at the lowest cost from a given timetable and from additional train-paths desired by operators.

The International Union of Railways proposes a new method that is included in the framework of the optimization methods (UIC, 2004), which is based on a timetable *compaction* method. By modifying the base timetable, existing train paths are scheduled as close as possible to each other. Modifying the travel times, the overtakings, the crossings, and the commercial stops is prohibited during the process of compaction. The remaining blank (empty, or unused) time left in the timetable represents the maximum spare

time during which additional train services may theoretically be scheduled. In practice, if too much of the unused capacity is taken to run more trains, this may cause serious reliability problems because it reduces the buffer times that allow minor incidents to be absorbed. An example of this method is shown in Landex et al. (2006), where they describe how the *compaction* method is applied in Denmark.

- ***Simulation Methods:*** A simulation is the imitation of an operation of a real-world process or system over time. It is the representation of dynamic behavior of a system by moving it from state to state in accordance with well-defined rules. Simulation methods provide a model, which is as close as possible to reality, to validate a given timetable. For train scheduling, simulation has often been used in combination with other methods, originating what could be defined as “hybrid models”. Since the 1980’s, Petersen (1974) are quoted for their work, which uses combined techniques, such as dynamic programming and branch-and-bound in a simulation context. Welch and Gussow (1986) use simulation and heuristics to evaluate the relative effect of the many factors that influence line capacity. A composite simulation and optimization method also appears in the work of Jovanovic and Harker (1991). A model called *Strategic Capacity Analysis for Network* (SCAN) was developed by Kaas (1991), who defined factors at different levels of detail that together determine the capacity of a network. Besides purely academic products, various simulation environments have been produced and are commercially available in the rail industry. The general performances of these simulators are comparable, and their main technical differences concern interface design, user interaction and flexibility, track infrastructure, other data management processes, and integration with company information systems. These tools normally generate timetables by time-stepping simulation using train motion differential equations. Alternatively, these tools can be used to validate a given timetable, which is provided by optimization methods. They can detect delays and analyze interferences on the preliminary timetable. A complete analysis of the commercial simulation environments is outside the scope of this paper. Nevertheless, we give a short description of some main simulation environments:
 - *MultiRail* (Multimodal Applied Systems) has a full train simulation capability, which shows train conflicts in detailed graphical and tabular formats. The user can define the conflict rules within the track manager, when defining the various characteristics of the track arcs and switches. Furthermore, *MultiRail* has a very fast seven-day simulation algorithm called *SuperSim* that produces a complete trip plan for each traffic record.
 - *OpenTrack* (OpenTrack Railway Technology) is a simulation tool to answer questions about railway operations. It calculates train movements under the constraints of the signaling system and timetable. It also handles simulation where random generators produce different initial delays and station delays.
 - *SIMONE* (Incontrol Enterprise Dynamics) is a simulation model for rail-

way networks. It can support: determination of the robustness of a timetable, trace and quantification of bottlenecks in a network, and analysis of cause and effect relations when delays emerge.

More information about these systems and other simulation environments can be found in Barber et al. (2007).

In conclusion, analytical methods may be a good start for optimization methods. The optimization and simulation methods need to be adapted to each application environment. The three levels represent a general methodology for capacity management, where the first level represents a preliminary solution and the second level obtains a desired train schedule which is validated by means of simulation (the third level). Therefore, the current trend is to develop tools with an integrated methodology that embed analytical, optimization and simulation approaches.

The following section describes a system that is able to evaluate both analytical methods (UIC (UIC, 1983), Periodic Trains (Salido et al., 2004), (Abril et al., 2005)) and optimization methods (*Compaction* (UIC, 2004), Saturation) and verify the results obtained.

4 A System for Capacity Analysis

The new policy of the European Union is to encourage open access to railway networks. This process has already begun in the Spanish Administration of Railway Infrastructure, ADIF, which is interested in using advanced computer tools to improve railway management. In collaboration with ADIF, the authors have developed a tool called MOM (acronym of the Spanish name: Modulo Optimizador de Mallas) that embeds analytical and optimization approaches in this context. This integrated system helps railway managers perform capacity studies to optimize their timetables and evaluate their track and station capacity in order to satisfy the demands of their customers.

4.1 General System Architecture

The MOM system (Barber et al., 2006a) is a flexible and useful tool for the automated and optimized management of railway running maps (timetables) in accordance with railway infrastructure, traffic rules, user and commercial requirements, and optimization criteria.

The MOM system obtains the required data from standard databases with information on railway infrastructures, types of trains, journeys, scheduled

running-maps, commercial requirements, etc., and obtains the optimization parameters from friendly interfaces. Afterwards, it transforms the problem into mathematical models, applies advanced algorithmic techniques and heuristic criteria (Barber et al. (2004), Barber et al. (2006b), Ingolotti et al. (2004), Salido et al. (2004)) and automatically obtains optimized, robust, reactive and flexible railway timetables. The MOM system do the following: it can optimize existing running maps; schedule new periodic and/or not periodic trains, making them compatible with the existing scheduled trains; validate and perform theoretical and practical capacity analyses; simulate the effects of delays or incidences; measure robustness of timetables; and reschedule running maps according to incidences and delays. Therefore, the MOM system can be considered as one of the most complete Decision Support Software Systems for railway timetabling since it can support managers of railway infrastructure in the difficult decision-making process they face.

The MOM system is a very efficient software tool. The computational effort to obtain a satisfactory optimized solution depends on the number of trains, the railway infrastructure and its capacity (tracks in stations, single/double-way tracks), the required traffic operations due to the load of the network, etc. Typically, the computational time required for very complex and real problems varies between a few seconds and 2-5 minutes. For instance, a railway timetabling problem that implies the scheduling of 95 new trains, with 37 trains already in circulation (with fixed timetables), on a line of 271.1 Kms, with 51 single-way track sections is optimized by the MOM system in less than 60 seconds. This scheduling problem implies the solution of 136 crossing conflicts. In this case, the search space is composed of $8.7E+40$ possible solutions. In order to reduce the search space, the MOM system makes use of several pre-processes, operations research techniques, and powerful, intelligent heuristics (Ingolotti et al., 2004). Moreover, several levels of optimization can be selected by the user. The minimum level obtains a feasible but semi-optimized timetable in a few seconds, so that the user can easily and efficiently evaluate different scheduling parameters and alternative timetables. The maximum optimization level obtains higher quality timetables when the computational time is longer.

The general outline of the capacity module of the MOM system is presented in Figure 9. It shows several steps, some of which require direct interaction with a human user in order to select the line and trains to be studied, to choose the method, and to insert the required parameters according to the method. The MOM system accesses the databases to obtain information about the line to be studied and the trains that use or intend to use the line.

FIGURE 3

4.2 Types of Capacity Analyses that can be performed

The system is able to perform both analytical and empirical studies.

4.2.1 Analytical Methods

The system can carry out several analytical capacity studies, which are described below.

- (1) The first is a theoretical study that is based on the UIC 405 Formula (UIC, 1983), which has been improved in accordance with the general recommendations of UIC (1996):

$$L = T / (t_{fm} + t_r + t_{zu})$$

where L is the capacity of a section in number of trains in period T ; t_{fm} is the average time span at minimal sequence of trains; t_r is the running time margin; and t_{zu} is the added time.

This method allows the MOM system to efficiently obtain an estimated capacity. The first version of this formula is strongly dependent on the defined sequence of trains. Thus, the timetable is required to know the number of sequence cases where one train of type i follows one train of type j . However, we use the revised version where only the proportion of trains of each type is required. The parameters t_r and t_{zu} help take into account infrastructure characteristics and network effects.

- (2) A second type of analytical studies for periodic trains (Abril et al. (2005), Salido et al. (2004)) can also be performed. These studies are based on characteristics such as Infrastructure Topology or Operational Constraints. These operational constraints are: (i) Minimum distance between one train and the next one on the line. (ii) Overtaking between trains, which can only occur within a station. (iii) Minimum distance between two consecutive arrivals/departures in a station. (iv) Two trains going in opposite directions, which must not simultaneously use the same one-way track. (v) Reception time, which is the time required to detour a train from the main track so that crossing or overtaking can be performed; and Expedition time, which is the time required to put a detoured train back on the main track and exit from a station. (vi) The maximum number of trains that can simultaneously be present in each station, which is given.

There are two methods for studying periodic trains:

- (a) The Traffic Pattern method: The main idea of Salido et al. (2004) is to generate a traffic pattern for each set of stations such that the union of these contiguous traffic patterns determines the journey of each train. These traffic patterns will be periodically repeated to compose the entire timetable and obtain the line capacity.

- (b) The Fractional Method: The method proposed in Abril et al. (2005) is based on first performing an individual study for each line section in order to evaluate all possible crossing scenarios. Then, the method analyzes this information and detects a critical line section. Finally, it evaluates whether the capacity of the critical line section can be completely absorbed by the other line sections. This method is detailed in section 5.1.

Thus, we can obtain reports that provide relevant data about capacity for the studied line. With these data, we can obtain conclusions about the capacity of the line (maximum capacity, bottlenecks, saturated line sections, etc.), we can help in the planning of train scheduling, and we can provide alternative planning, etc.

4.2.2 Optimization Methods

The results obtained from the analytical methods can be validated in the optimization module by means of the following optimization methods:

- Compacting a previous timetable and studying the available capacity.
- Obtaining alternative saturated and consistent timetables
- Saturating a previous timetable with new trains.

5 Applying Methods to Evaluate Railway Capacity

In this section, we evaluate the performance of our tool with several real problems. We show some of its capabilities in real capacity studies.

5.1 Capacity Analysis on Periodic Trains

The objective of this example is to demonstrate the use of the MOM system to obtain the maximum capacity of a railway infrastructure which is used by periodic trains. We use the periodic train analytical method described in Abril et al. (2005) to determine the maximum line capacity. This method is implemented in the capacity module of the MOM system. Later, we verify the achieved result obtaining a saturated timetable with Periodic Trains using the optimization module of the MOM system.

Figure 9 shows the line selected in this assessment. It is a railway line between two Spanish cities: “*Santa Cruz - Linares*”. It has about 75 kilometers, 10 stations and 4 halts. In Figure 9, the first column shows the station kilometers,

the second column shows the number of platforms in each station, and the third column indicates the station name. This line only has single tracks. Furthermore, in this example, an exceptional situation is reflected: it has a special line section (“*Venta de Cardenas - Calancha*”). The two stations inside this section have one track and, therefore, it is not possible to perform a crossing in this station. The *time period* is 24 hours. The routes of the selected trains are shown in Figure 9.

FIGURE 4

The maximum capacity for these periodic trains can be calculated according to the periodic train method described in Abril et al. (2005). It follows the formula:

$$Capacity_{jks} = (Period - TI_{1j} - TV_{nk} - TI_{kn} - TV_{j1}) / Ts_{jk}$$

where $Capacity_{jks}$ is the capacity of the line section delimited by the stations j and k when the crossing type s (see Figure 9) is carried out. Ts_{AB} is the global travel time of the line section delimited by the stations A and B when the crossing type s is carried out. TI_{jk} is the time needed by the train traveling in the down direction to go from station j to station k (without technical stops). TV_{kj} is the time needed by the train traveling in the up direction to go from station k to station j (without technical stops). n is the last station of the line.

FIGURE 5

Figure 9 shows crossing types for a line section delimited by two contiguous stations, j and k , where rec_I and exp_I are the reception and expedition times, respectively, of the train traveling in the down direction, and rec_V and exp_V are the reception and expedition times, respectively, of the train traveling in the up direction.

The results achieved with the above method are shown in Table 9. This table shows the maximum capacities classified by line section and crossing type. To determine the critical line section and its maximum capacity, we select the line section with the minimum capacity as the initial line capacity.

TABLE 1

The algorithm in Figure 9 verifies whether the capacity of the critical line section remains feasible in the neighboring stations due to the performed crossing types; otherwise, it updates the capacity value.

FIGURE 6

According to the data shown in Table 9 and the Algorithm in Figure 9, the critical line section is “*Venta de Cardenas - Calancha*”. This is a very long line section compared with the rest of the sections. The maximum capacity is 36 trains per day for each direction. This result is empirically validated.

Figure 9 shows a valid and saturated timetable of the “*Santa Cruz - Linares*” line, which was generated with the optimization module with 36 trains per day for each direction. Therefore, the maximum capacity calculated is valid and can be scheduled in a feasible timetable.

FIGURE 7

Very restrictive line sections (such as “*Venta de Cardenas - Calancha*”) are a clear example of bottlenecks that limit the global line capacity.

In addition to this conclusion about the capacity of the line, the tool performs a “what-if” analysis and gives some alternatives for improvement. If a new platform were included at intermediate stations, we could conclude that the capacity increases considerably, by making automatic additional calculations. Specifically:

- If a new platform is added at the *Las Correderas* station, then the maximum line capacity is 46 trains for each direction.
- If the new platform is added at the *Santa Elena* station, then the maximum line capacity is 47 trains for each direction.
- However, if a new platform is added at both stations, the maximum line capacity will not increase further; it remains at 47 trains for each direction.

According to this information, the most profitable solution for increasing the line capacity is to include a new platform in the *Santa Elena* station.

5.2 UIC Analysis about Mixed Trains

The objective of this example is also to demonstrate the use of the MOM system to obtain the maximum capacity of a railway infrastructure, however, in this case, mixed trains run in the railway line. We first carry out an analytical study to determine the maximum line capacity. We use the UIC 405 Formula (UIC, 1983) implemented in the capacity module of the MOM system. Later, we validate the theoretical results using the optimization module of the MOM system to obtain a saturated timetable with these mixed trains. We show how the MOM system is able to give alternative timetables for the same mixed trains.

The first step in obtaining a saturated timetable is to choose the railway line

and the train types to schedule. Figure 9 shows our choices in a real situation. The railway network selected is the “*Etxarri - Pamplona*” line. This line has 42 kilometers, 5 stations, and 2 halts. We have chosen 6 different train types (3 in each direction with different speeds). Their routes are shown in Figure 9.

FIGURE 8

Now, we must decide the number of trains of each type in order to obtain a saturated timetable. This number can be determined by a market study, a forecast of future demand, an analysis of the current timetable, etc. Table 9 shows the percentage of trains (%) that run along this line; these percentages were calculated according to real-world operator data, after analyzing the current timetable for the studied line (the first eight hours). Table 9 also shows the capacity forecast based on the UIC 405 Formula (UIC, 1983).

TABLE 2

The next step is to validate the capacity data. The MOM system optimization module carries out different schedulings and provides alternative timetables with the required capacity. Figure 9 shows two different valid timetables generated with 46 trains each.

The results indicate that the capacity shown in Table 9 is feasible, but it does not necessarily have to be the maximum line capacity. Figure 9 shows a new timetable with 4 new trains. There is still available capacity for other new trains, but the trains are more bunched. Therefore, the capacity will depend on operative constraints. User intervention is crucial in the choice of the best saturation strategy.

These results show some of the weaknesses of the UIC 405 Formula. Estimating the capacity of a line on the basis of an extrapolation of the critical section capacity is clearly not a good method.

FIGURE 9

FIGURE 10

5.3 *Compacting Timetable*

In their last recommendation, the International Union of Railways presented the *compaction* method as the best way of performing a capacity study (UIC, 2004). In this section, we use the *compaction* method implemented in the capacity module of the MOM system to determine the available capacity of a

real network.

FIGURE 11

The *compaction* method begins with the selection of a base timetable (Figure 9). The network selected is the “*Noain - Castejon*” line. This line has about 75 kilometers, 10 stations, and 4 halts. There are several trains running on this network which have different speeds. The time window chosen is 12 hours.

Two compaction types can be carried out in the MOM system:

- (1) The first type performs the compaction respecting only the order in which the trains run on the line and the commercial stops (it does not take into account overtakings or crossings). That is:

$$\forall i \in [2..n] : TDep_{i,0} = TDep_{i-1,0} + Suc_{i-1,i}$$

where $TDep_{i,0}$ is the departure time of train i at station 0, $Suc_{i-1,i}$ is the safety time between trains $i-1$ and i , and n is the number of trains. Figure 9a shows the base timetable for Figure 9 which is compacted using this compaction type. The available time for new trains is approximately 8 hours. The available time in the compacted timetable of Figure 9a allows us to triple the current *used capacity* (Figure 9) with the same types and sequence of trains.

- (2) The second type performs the compaction respecting the order in which the trains run on the line, as well as the overtakings, the crossings and the commercial stops. Figure 9b shows this compaction type. The available time is almost 4 hours. Furthermore, new trains can be scheduled into the compacted timetable, which represents more available capacity.

It is important to remark that the compacting method eliminates buffer times between trains. These schedule slacks could absorb normal operational variances, so their elimination may cause serious reliability problems in the management of the infrastructure.

FIGURE 12

6 How Different Parameters Affect Railway Capacity

This example illustrates the influence of several factors that affect capacity.

FIGURE 13

The network selected for our example is the “*Reinosa - Torrelavega*” line. It has about 59 kilometers with 9 stations and 9 halts. The real network has a

single track, but in this example, we will also simulate this line with a double track because it is the most common situation in many real railway lines. The chosen time window is 24 hours. Currently, several types of trains run on this line. We choose two fast trains that run in opposite directions whose route is shown in Figure 9 (this is the basic case).

The maximum line capacity of the basic case according to the UIC 405 Formula (UIC, 1983) is 82 trains for each direction in the single-track case, and 360 trains for each direction in the double-track case. In section 5, we have verified that these analytical results are not immutable. However, we use them in our comparative study because they are a good reference.

Just as we expected, according to the work of Kittelson and Associates (2003), *the double-track line capacity (360) is greater than twice the single-track line capacity (82)*.

Now, we study the effect on railway capacity of four main traffic or operating parameters. In each case, we also determine the maximum line capacity with the UIC 405 Formula (UIC, 1983). The variation of these parameters increases or decreases railway capacity in a different way (Figure 9). We can increase capacity by increasing speed or decreasing robustness, commercial stops, and heterogeneity. Clearly, increasing capacity decreases reliability (see Figure 9).

FIGURE 14

Train speed.

Figure 9 shows how the speed of the trains affects capacity. Slow trains decrease the capacity because they need more travel time. Furthermore, the *Headway Time* between consecutive trains depends on two main factors: the line section length (or the distance between signs) and the train speed (Figure 9). These two factors affect the *Travel Time*, the *Braking Time*, and the *Release Time*. In general, the *Headway Time* can be estimated by summing up these three times plus a constant *Operating Time*.

FIGURE 15

FIGURE 16

In this example, we decrease the speed of the trains in Figure 9 by 50%. As Figure 9 shows, in the single-track case, *when speed decreases, the maximum line capacity decreases*. Specifically, there is nearly 65% of capacity when the train speed decreases by 50%. However, in the double-track case, the UIC 405 Formula always obtains the same capacity (100%, Figure 9). This is due to the fact that it does not modify the *Headway Time*. If we used the *Headway Time* according to the train speed, then the capacity would be different. An

example of this is shown in case 6 (Commercial Stop).

FIGURE 17

Train heterogeneity.

This case includes mixed trains. We mix two trains with different speeds and different proportions. The fast trains used in this example are shown in Figure 9. The slow trains used in this example are 50% slower than these fast trains.

Figure 20 shows how the heterogeneity in double-track and single-track cases affects capacity. In the single-track case, there are four types of heterogeneity. These heterogeneities take into account the sequence of trains. The sequence heterogeneity increases from $Heter_1$ up to $Heter_n$ (Figure 9). $Heter_1$ first arranges all the trains in one direction and later arranges the trains in the opposite direction. This type of sequence is called *train fleeting*. It gives the maximum capacity; however, it significantly reduces the operating flexibility. $Heter_n$ alternates one train in the down direction with one train in the up direction. This type of sequence is called *even spacing*, i.e., trains are evenly spaced during the period with other trains of the same type. It gives the minimum capacity in saturated timetables.

As can be observed, in Figure 20 the *double-track* and $Heter_1$ cases have the same performance. This is due to the fact that $Heter_1$ simulates a double-track line dividing the *time period* into two parts: the first part of the period is used only by trains in one direction, and the second part is used only by trains in the opposite direction.

FIGURE 18

In Figure 20, the symbol X marks an abrupt change of capacity evolution. This decrease in capacity is caused when the first train with a different speed is placed among the other trains. This effect is shown in Figure 9. The effect is similar when we insert a slower train (Figure 19a1) or a faster train (Figure 19b1). However, when we insert the second train with a different speed, then only one (Figure 19b2) or two (Figure 19a2) more trains are eliminated.

FIGURE 19

The *double-track* and $Heter_1$ cases reach their lowest boundary of capacity when the heterogeneity is maximum (symbol k in Figure 20), that is, we have the same number of fast and slow trains. Thus, in double-track lines, the capacity is affected by heterogeneity more than by train speed.

In single-track lines, $Heter_1$ is not the most common sequence of trains. The circulation is usually alternate. Figure 20 shows three different alternate cir-

culations: $Heter_2$, $Heter_3$, and $Heter_4$, where $Heter_3$ is more heterogeneous than $Heter_2$, and where $Heter_4$ is more heterogeneous than $Heter_3$. In these cases, the lowest boundary of capacity is reached when there are 100% slow trains (symbol r in Figure 20). In conclusion, in single-track lines, capacity is more affected by train speed than by heterogeneity.

FIGURE 20

Quality of service, Reliability, or Robustness.

We chose to increase the *Headway Time* between consecutive trains in order to obtain more robust timetables, thereby increasing the quality of service. The current *Headway Time* is 3 minutes, but, in this example, the *Headway Time* is increased from 3 minutes to 6 minutes at intervals of 30 seconds. As can be observed in Figure 21, *when the introduced buffer times are increased, the capacity clearly decreases*, especially in double-track lines. In other words, introducing scheduled slacks among trains will decrease capacity but will increase reliability, allowing absorb delays due to typical operational incidents.

FIGURE 21

Commercial stops.

The UIC 405 Formula obtains small deviations of capacity when new commercial stops are introduced. We do not show these results because these deviations are virtually nil. This is due to the fact that the UIC 405 Formula studies each section line independently and new commercial stops represent a very small increase in the *Travel Time* of each section line. However, with other analyses we have verified that the commercial stops do affect the capacity.

When a train has to stop at a station, it starts to decelerate enough line section or signal sections before arriving to the station. Consequently, the train increases its *Travel Time*, and the *Headway Time* between consecutive trains must also be increased. Figure 22 shows an example of this situation.

FIGURE 22

More information about the influence of commercial stops on *Headway Time* is shown in 7.4.

7 Capacity Analyses on The European Rail Traffic Management System

The European Rail Traffic Management System (ERTMS)² is the initiative from the European Commission to create a unique signaling standard as a cornerstone for the achievement of the interoperability of the trans-European rail network and is likely to be adopted by the rest of the world as well.

To account for the diversity of functional/operational requirements, three system levels of reference have been identified. There are already ERTMS commercial projects in several countries, most of which have already established ERTMS level 2. In this level, ERTMS does not require trackside signals, and the train position is controlled with virtual signals and radio control messages.

To ensure the safe operation of the train and to enable the optimization of the line capacity, the specifications of ERTMS include the calculation method of the *Headway Time* between consecutive trains: Formula (1). This theoretical method is the same in ERTMS level 1 and ERTMS level 2. Thus, the theoretical capacity is the same in both levels, although, in practice, ERTMS level 2 is more efficient. ERTMS determines the *Headway Time* by summing up the following four times (see Figure 23):

$$\textit{Headway Time} = \textit{Travel Time} + \textit{Braking Time} + \textit{Release Time} + \textit{OT} \quad (1)$$

FIGURE 23

where:

- *Travel Time* is the time required to cover the distance between two consecutive virtual signals. *Travel Time* depends inversely on the train speed and directly on the distance between consecutive virtual signals.

$$\textit{Travel Time} = F(\textit{Distance}/\textit{Speed})$$

- *Braking Time* is the time needed to cover the braking distance, that is, the distance required to stop a train before a virtual signal. *Braking Time* depends directly on the train speed and inversely on the maximum deceleration.

$$\textit{Braking Time} = F'(\textit{Speed}/\textit{Deceleration})$$

- *Release Time* is the time required for the entire length of a train to cross a virtual signal. *Release Time* depends on the train speed and the train length.

$$\textit{Release Time} = F''(\textit{Length}/\textit{Speed})$$

² <http://www.ertms.com>

- *Operating Time* (OT) is a safety time. It is a constant, and it is set by the infrastructure managers.

In our tool, train speed is calculated for each meter of train path by means of an underlying dynamic model.

The capacity of a double-track line in a fixed *time period* depends on the *Headway Time* between consecutive trains. This *Headway Time* is the maximum of all *Headway Times* between consecutive virtual signals of the line. In each line section, capacity is:

$$Capacity = \frac{Time\ Period}{Headway\ Time} \Rightarrow \quad (2)$$

$$Capacity = \frac{Time\ Period}{F\left(\frac{Distance}{Speed}\right) + F'\left(\frac{Speed}{Deceleration}\right) + F''\left(\frac{Length}{Speed}\right) + OT} \quad (3)$$

We assume a continuous *Operating Time* without interruptions (see Figure 24a). However, with discontinuous *Operating Times*, the *time period* decreases due to the journey time of the first train (see Figure 24b).

FIGURE 24

Formula (3) shows that capacity is strongly dependent on the train speed: It is directly proportional to speed due to the *Travel* and *Release Times*, but it is indirectly proportional to speed due to the *Braking Time*. When speed is constant, Formula (3) can be simplified as:

$$Capacity = \frac{Time\ Period}{\frac{Distance}{Speed} + \frac{Speed^2}{Deceleration} + \frac{Length}{Speed} + OT} \quad (4)$$

The proposed computer-based tool is used to analyze the influence of several parameters in ERTMS lines. All the data and figures have been automatically obtained by the MOM system.

7.1 How several parameters influence capacity in ERTMS lines

Figures 25 and 26 show the *Headway Times* of a railway line whose line sections have constant lengths. Specifically, in Figure 25, the line sections are 6000 meters long, which mean that the distance between consecutive virtual signals is 6000 meters. The train speed increases from 200 Km/h up to 500 Km/h.

Figure 25 shows the dependency of capacity on train speed. As can be observed, when the train is slow, the *Travel Time* influences capacity more

than the *Braking Time*. However, when the train is fast, capacity depends mainly on the *Braking Time*. The *Operating Time (OT)* and the *Release Time (Length/Speed)* are much smaller than the *Travel* and *Braking Times*.

FIGURE 25

FIGURE 26

In Figure 26, we analyze the influence of train speed on capacity with different line section lengths. For short line sections, when the train speed increases, the *Headway Time* also increases. However, for large line sections, the minimum *Headway Time* is obtained with a medium train speed. This is due to the fact that the *Braking Time* and the *Travel Time* are balanced. (See Figure 25). As Formula 2 indicates, the minimum *Headway Time* gives the maximum capacity.

Figures 27 and 28 show the influence of another important factor: the distance between consecutive virtual signals. At a given speed, as the distance gets bigger, the *Headway Time* increases. Furthermore, the distance has a large influence on the *Headway Time* when the speed is slow. Figure 27 shows that when the speed is 200 Km/h, the *Headway Time* grows faster than when the speed is 500 Km/h. This is due to the fact that the *Travel Time* depends on the distance between consecutive virtual signals. As Figure 28a shows, when the speed is slow, the *Travel Time* tends to be bigger than *Braking Time*. However, when the speed is fast, the *Travel Time* tends to be smaller than the *Braking Time*, showing that the influence of *Travel Time* is less significant (Figure 28b). This is why the lines cross in Figure 27.

FIGURE 27

FIGURE 28

7.2 Trains with Constant Acceleration

This section shows an ideal line with ERTMS and double track. We assume a constant acceleration until the train reaches the maximum speed of 500 Km/h, which is the maximum speed in ERTMS level 2. However, it should be taken into account that in a real line, a constant acceleration is nearly impossible because of different aspects of infrastructure, speed limit, very steep slopes, frictions, etc.

Figure 29 shows the *Headway Times* of a line with ERTMS where the virtual signals are uniformly located every 3000 meters and the train acceleration is constant ($0.3m/s^2$). Clearly, the *Headway Time* for the first meters (when

the train goes slow) is more influenced by the *Travel* and *Release Times*. Furthermore, the *Braking Time* of the first section is 0 since the train speed is 0 Km/h at the beginning of the section. However, as the train goes faster, the *Braking Time* exerts greater influence on the *Headway Time*. Finally, when the speed is constant, the *Headway Time* is also constant.

FIGURE 29

Figure 30 shows the maximum capacities according to the *Headway Time* of each section. The line capacity is limited by the sections with the greatest *Headway Time*; which are the sections with the maximum speed. Thus, in this case, the maximum line capacity is 16 trains.

FIGURE 30

We can conclude that the *Braking Time* is a prominent time in the *Headway Time*. Therefore, the ability to decelerate is an important factor in increasing capacity. Likewise:

- Section distances between signals inversely affect capacity, mainly at lower train speeds (Figure 27).
- Train speed has a *complex* influence on *Headway Time* in large sections (Figure 26). In shorter sections ($\approx < 4$ Km), the train speed *directly* affects the *Headway Time*. As the *Headway Time* increases, trains should be more separated and capacity should decrease. However, with a discontinuous *time period* (Figure 24b), if the speed increases, more trains can be scheduled in a given *time period* and capacity increases (Figure 31). Broadly speaking, in a *time period*, we can put n trains, such that:

$$n = (\text{TimePeriod} - \text{JourneyTime}) / \text{HeadwayTime}.$$
Therefore, assuming an ideal case of trains running at regular speed between two locations that are separated by 300 Km, Figure 32 shows the maximum capacity in a *time period* of 8 hours, depending on the distance between signals and train speed.
- Factors like train length or *Operating Time* do not affect the capacity significantly. However, train length could be important in very large trains.

FIGURE 31

FIGURE 32

7.3 Standard Line with ERTMS level 2

This section shows a standard line (which has infrastructure speed limits, very steep slopes, tunnels, frictions,...) with ERTMS level 2 and double track.

FIGURE 33

In the vertical right axis of Figure 33, it can be observed that the maximum speed limit (which in ERTMS level 2 is 500 Km/h) is never reached. The most crucial time along the line for calculating the *Headway Time* is the *Braking Time*, with the exception of the section where a commercial stop is carried out. This section has the greatest *Headway Time* because of the *Travel* and *Release Times*. Therefore, this section has the smallest capacity, which is 16 trains (see Figure 34).

FIGURE 34

7.4 The Introduction of a New Stop

We show how the line capacity decreases when a new commercial stop is introduced into a line with ERTMS. This is mainly due to the changes in speed.

Figure 35 shows two different train speed curves for one railway line. Figure 35a shows the speed curve when the train does not stop, and Figure 35b shows the speed curve when the train performs a stop. This last figure shows how the train decelerates and accelerates.

FIGURE 35

Figures 36 and 37 show the *Headway Time* difference between a train without a stop or with a stop. When the train decelerates and accelerates, the *Headway Time* of the affected line sections increases or decreases. Specifically, in the line section where the train stops (in this example, the third section), the *Headway Time* increases a lot (Figure 37). Therefore, the capacity of this line section decreases. Figure 38 shows how the capacity of the third line section decreases from 29 trains (Figure 38a) to 14 trains (Figure 38b).

FIGURE 36

FIGURE 37

FIGURE 38

8 Conclusions

In this paper, we have reviewed the term “*capacity*”, its different types, the methods to evaluate it, and the main parameters affecting it, as well as its importance in the railway sector.

Infrastructure authorities need software tools that help them better understand the economic trade-offs associated with trying to place ever more trains down an ever more congested rail asset. This includes lost ridership due to slower and less reliable schedules, operating cost increases particularly for lower priority services, displacement of freight onto less economical routes or onto the highways, loss of economic development opportunities in cities that are served by overly congested transportation system as industries relocate out of those areas in search of other locations where they can actually operate more economically.

With these ideas in mind, there are currently a great variety of computer-based railway systems that incorporate functions to do capacity estimates. However, each one has its peculiarity regarding the evaluated characteristics of infrastructure, rolling stock, and railway traffic, making it difficult to compare their results.

The standards to evaluate railway capacity are based on either analytical models (spreadsheet, the UIC Formula, etc.) or optimization processes (*compaction*, saturation, etc.). These methods obtain theoretical capacity estimates that allow us to bound the maximum number of trains that can be scheduled. Theoretical capacity is clearly affected by several railway parameters such as infrastructure (number of tracks, signals, etc.), traffic features (speed of trains, heterogeneity, etc.), and operating requirements (commercial stops, maximum trip time, etc.). Specifically, we show the great influence of train speed and signal position over capacity in the new signalling and management system for the European rail networks (ERTMS). Hence, it is essential to know how railway parameters influence theoretical capacity because it allows us to adjust them in order to analyze their effect on network capacity.

FIGURE 39

A clear distinction should be made between theoretical and practical capacity (see Figure 39). Practical capacity is a more realistic measure of railway capacity, which takes into account reliability and operational aspects. Clearly, there is a trade-off between capacity and reliability/robustness. In other words, between the “physical maximum” level of capacity and the “economically optimal” level of capacity. This issue is a key point in operational management. However, there are no clear standards about how this trade-off can be modelled. It requires in-depth analysis to model railway reliability and to stan-

standardize the desired levels of robustness.

Railway managers know that scheduling as many trains as theoretical capacity is not viable. However, they must make the best use of the expensive railway infrastructures. For this purpose, the application of simulation methods seems to be a convenient way to evaluate the trade-off between capacity and reliability. However, an in-depth analysis using simulation methods must be done to obtain practical capacity measures.

To carry out the analytical and empirical capacity studies shown in this paper, we have used the MOM system, a computer-based tool that has several analytical and empirical methods embedded in it. The MOM system contains an optimization module for obtaining feasible and optimized timetables that performs the compaction of a previous timetable, obtains saturated timetables to evaluate the maximum capacity, and performs several other analytical and empirical capacity assessments based on the UIC model, ERTMS lines, etc. All the data in this paper correspond to real-world cases. Likewise, all the results have been automatically obtained by our system. Additional features of the MOM system for railway capacity analysis can be obtained from the authors, as well as from <http://www.dsic.upv.es/users/ia/gps/MOM/>.

9 Acknowledgements

The authors are grateful to the anonymous referees for their helpful comments and valuable suggestions to improve the manuscript. We also appreciate the assistance of Jose Estrada of the Spanish Administration of Railway Infrastructure (ADIF). This work has been partially supported by the research projects TIN2004-06354-C02-01 (Min. de Educacion y Ciencia, Spain-FEDER), FOM-70022/T05 (Min. de Fomento, Spain), GV/2007/274 (Generalidad Valenciana) and by the Future and Emerging Technologies Unit of EC (IST priority - 6th FP), under contract no. FP6-021235-2 (project ARRIVAL).

References

- Abril, M., Salido, M.A., Barber, F., Ingolotti, L., Lova, A., Tormos, P., 2005. *A Heuristic Technique for the Capacity Assessment of Periodic Train*, Frontiers in Artificial Intelligence and Applications (IOS Press) 131, 339-346.
- AEA Technology Rail, *CMS: Capacity Management Systems*, VISION, VAMPIRE. <http://www.aeat.co.uk/rail/>.
- Akyilmaz, M.O., 1976. *Trade-off analyses for multi-objective transportation plans*, PhD thesis, University of Pennsylvania.

- Assad, A., 1980. *Models for Rail Transportation*, Transportation Research A 14, 205-220.
- Barber, F., Abril, M., Salido, M.A., Ingolotti, L., Tormos, P., Lova, A., 2007. *Survey of automated Systems for Railway Management*, Technical Report DSIC-II/01/07, Department of Computer Systems and Computation, Technical University of Valencia.
- Barber, F., Salido, M.A., Ingolotti, L., Abril, M., Lova, A., Tormos, P., 2004. *An Interactive Train Scheduling Tool for Solving and Plotting Running Maps*, Current Topics in Artificial Intelligence LNAI 3040, 646-655.
- Barber, F., Tormos, P., Lova, A., Ingolotti, L., Salido, M.A., Abril, M., 2006a. *A Decision Support System for Railway Timetabling (MOM): The Spanish Case*, Computers in Railways X 235-244.
- Barber, F., Tormos, P., Salido, M.A., Abril, M., Ingolotti, L., Lova, A., 2006b. *New heuristics to solve the CSOP railway timetabling problem*, Advances in Applied Artificial Intelligence. LNCS-LNAI 4031, 400-409.
- Burdett, R.L., Kozan, E., 2006. *Techniques for absolute capacity determination in railways*, Transportation Research Part B 40, 616-632.
- Cai, X., Goh, C., 1994. *A Fast Heuristic for the Train Scheduling Problem*, Computers and Operations Research 21, 499-510.
- Canciani, G., 1991. *Criteri Progettuali di Rinnovo e Potenziamento delle Linee Ferroviariarie: Modello di Calcolo e di Verifica della Potenzialita' di Circolazione*, PhD thesis, University of Trieste.
- Caprara, A., Fischetti, M., Toth, P., 2000. *Modeling and Solving the Train Timetabling Problem*, Research Report OR/00/9 DEIS.
- Carey, M., Lockwood, D., 1995. *A Model, Algorithms and Strategy for Train Pathing*, Journal of the Operational Research Society 46, 988-1005.
- Chen, B. and Harker, P., 1990. *Two Moments Estimation of Delay on Single-Track Rail Lines with Scheduled Traffic*, Transportation Science, 24, 261-275.
- Egmond, R.J. van., 1999. *Railway Capacity Assessment, an Algebraic Approach*, TRAIL Studies in Transportation Science, S99/2, Delf University Press.
- Forsgren, M., 2003. *Computation of Capacity on Railway Networks*, Technical Report T2003-12, Swedish Institute of Computer Science.
- Harker, P.T. and Hong, S., 1990. *Two moments estimation of the delay on a partially double-track rail line with scheduled traffic*, Journal of the Transportation Research Forum 31 (1) 38-49.
- Harker, P.T. and Hong, S., 1994. *Pricing of Track Time in Railroad Operations: An Internal Market Approach*, Transportation Research 28B (3) 197-212.
- Higgins, A., Kozan, E., Ferreira, L., 1996. *Optimal Scheduling of Trains on a Single Line Track*, Transportation Research B 30, 147-161.
- Incontrol Enterprise Dynamics. *SIMONE: SIMulation MOdel for NETworks*, <http://incontrol.nl/files/simone.pdf>
- Ingolotti, L., Barber, F., Tormos, P., Lova, A., Salido, M.A., Abril, M., 2004. *An Efficient Method to Schedule New Trains on a Heavily Loaded Railway*

- Network*, Advances in Artificial Intelligence. LNCS/LNAI 3315, 164-173.
- Jovanovic, D., Harker, P., 1991. *Tactical Scheduling of Rail Operations: the SCAN I System*, Transportation Science 25, 46-64.
- Kaas, A. H., 1991. *Strategic Capacity Analysis of Networks: Developing and Practical use of Capacity Model for Railway Networks*, ScanRail Consult, Technical University of Denmark.
- Kittelson and Associates, Inc., 2003. *Transit Capacity and Quality of Service Manual 2nd Edition*, Transportation Research Board Washington, D.C..
- Kraft, E.R., 1982. *Jam Capacity of Single Track Rail Lines*, Proceedings of the Transportation Research Forum 23 (1) 461-471.
- Kraft, E.R., 1988. *Analytical Models for Rail Line Capacity Analysis*, Journal of Transportation Research Forum 29 (1) 153-162.
- Krueger, H., 1999. *Parametric Modelling in Rail Capacity Planning*, Proceedings of the 1999 Winter Simulation Conference, 1194-2000.
- Landex, A., Kaas, A. H., Schittenhelm, B. and Schneider-Tilli, J., 2006. *Practical use of the UIC 406 capacity leaflet by including timetable tools in the investigations*, Computers in Railways X, 643-652.
- Lucchini, L., Curchod, A., 2001. *Transalpine Rail Network: A Capacity Assessment Model (capres)*, First Swiss Transport Research Conference.
- Malaspina, R., Reitani, G., 1995. *Un Criterio di Calcolo della Potenzialita di Circolazione Ferroviaria su Linee a Doppio Binario*, Ingegneria Ferroviaria, CIFI review, Italian Railway Engineers Association.
- Martland, C.D., 1982. *PMAKE Analysis: Predicting Rail Yard Time Distributions Using Probabilistic Train Connection Standards*, Transportation Science 16 (4), 476-506.
- Multimodal Applied Systems, *MultiRail and FastTrack II*. <http://www.multimodalinc.com>.
- Oliveira, E., Smith, B.M., 2000. *A Job-Shop Scheduling Model for the Single-Track Railway Scheduling Problem*, Research Report 2000.21, University of Leeds.
- OpenTrack Railway Technology, *Railway Simulation*. <http://www.opentrack.ch>.
- Petersen, E., 1974. *Over-the-Road Transit Time for a Single Track Railway*, Transportation Science 8, 65-74.
- Petersen, E.R., Taylor, A.J., 1982. *A Structured Model for Rail Line Simulation and Optimization*, Transportation Science 16, 192-206.
- Salido, M.A., Abril, M., Barber, F., Ingolotti, L., Tormos, P., Lova, A., 2005. *Topological Constraint in Periodic Train Scheduling*, Frontiers in Artificial Intelligence and Applications 117, 11-20.
- SMA and Partner, *Viriato: A Software Package for Net-Wide Railway Planning*. <http://www.sma-partner.ch>.
- SNCF and Eurodecision, 2004. *Demiurge: A Tool for the Optimisation and the Capacity Assessment for Railway Infrastructure*.
- Stratec, *Railcap: A Computer Tool for Studying Capacity Problems of Railway Networks*. <http://www.stratec.be/FichesEtudes/PlanGBrailcap.htm>.

- Szpigel, B., 1972. *Optimal Train Scheduling on a Single Track Railway*, in Operational Research 72, 343-352.
- UIC Leaflet 405-1, 1983. *Method to be Used for the Determination of the Capacity of Lines*, International Union of Railways.
- UIC Leaflet 405 OR, 1996. *Links between Railway Infrastructure Capacity and the Quality of Operations*, International Union of Railways.
- UIC Leaflet 406, 2004. *Capacity*, International Union of Railways.
- Welch, N., Gussow, J., 1986. *Expansion of Canadian National Railway's Line Capacity*, Interfaces 16, 51-64.

FIGURE LEGENDS

FIGURE 1: Practical Capacity involves the desirable reliability level.

FIGURE 2: Saturation level according to average delay.

FIGURE 3: General scheme of the capacity module of the MOM system.

FIGURE 4: Trains and railway infrastructure selected to study railway capacity with periodic trains.

FIGURE 5: Crossing types for a railway line section delimited by two contiguous stations, j and k .

FIGURE 6: Saturated timetable with periodic trains in the “*Santa Cruz - Linares*” line.

FIGURE 7: The algorithm to evaluate railway capacity with periodic trains.

FIGURE 8: Train types and the infrastructure selected to study the railway capacity.

FIGURE 9: Two different timetables with 46 trains in each one.

FIGURE 10: A new timetable with more trains than the capacity obtained by the UIC Formula (50 trains).

FIGURE 11: Trains and the infrastructure selected to study the *compaction* method.

FIGURE 12: Compacted timetables using two different compaction types.

FIGURE 13: Trains and the infrastructure selected to study the influence of different parameters.

FIGURE 14: The main parameters that affect capacity.

FIGURE 15: Hourly capacity with fast and slow trains.

FIGURE 16: *Headway Time* between consecutive trains.

FIGURE 17: Influence of train speed on capacity, in single and double track cases.

FIGURE 18: Different sequence of trains modifying the heterogeneity.

FIGURE 19: Capacity decrease due to speed heterogeneity.

FIGURE 20: Influence of the heterogeneity on capacity.

FIGURE 21: Influence of the robustness on capacity.

FIGURE 22: Suppression of route slots due to stops.

FIGURE 23: *Headway Time* diagram.

FIGURE 24: Difference between continuous and discontinuous *time periods*.

FIGURE 25: *Headway Time* vs. *Braking* and *Travel Time*; line section length = 6000 meters.

FIGURE 26: *Headway Time* with several constant distances between virtual signals.

FIGURE 27: *Headway Time* with several constant speeds.

FIGURE 28: *Headway Time* with constant speed.

FIGURE 29: *Headway Time* with constant acceleration.

FIGURE 30: Hourly capacity with constant acceleration.

FIGURE 31: Reduction of *time period*.

FIGURE 32: The maximum capacities in a *time period* of 8 hours, and the *Headway Times*.

FIGURE 33: *Headway Time*, standard ERMTS line.

FIGURE 34: Hourly capacity, standard ERTMS line.

FIGURE 35: Speed curve: without stop vs. with stop.

FIGURE 36: *Headway Time* without stop.

FIGURE 37: *Headway Time* with stop.

FIGURE 38: Maximum hourly capacity: without stop vs. with stop.

FIGURE 39: Reliability is the most influential parameter between theoretical capacity and practical capacity.

<i>Line Section</i>	<i>Crossing s=1</i>	<i>Crossing s=2</i>	<i>Crossing s=3</i>	<i>Crossing s=4</i>
Santa Cruz de Mudela - Pealajo	70	58	58	58
Pealajo - Viso del Marques	78	64	64	54
Viso del Marques - Venta de Cardenas	59	50	51	44
Venta de Cardenas - Calancha	36	33	33	30
Calancha - Vilches	61	52	52	45
Vilches - Vadollano	53	46	46	40
Vadollano - Linares	55	47	55	47

Table 1

Maximum capacities for each crossing type.

<i>Train Type</i>	<i>Proportion</i>	<i>Capacity</i>
Type 1	11%	5
Type 2	11%	5
Type 3	16.5%	8
Type 4	16.5%	8
Type 5	22.5%	10
Type 6	22.5%	10

Table 2

Current train proportions and capacity forecast for each train type.

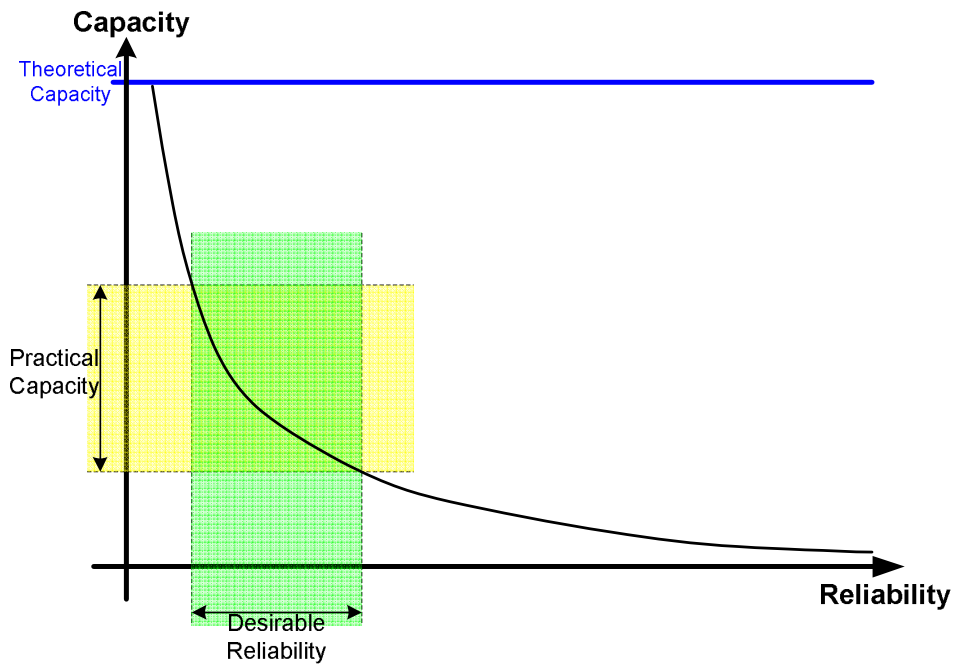


Fig. 1. Abril, M.

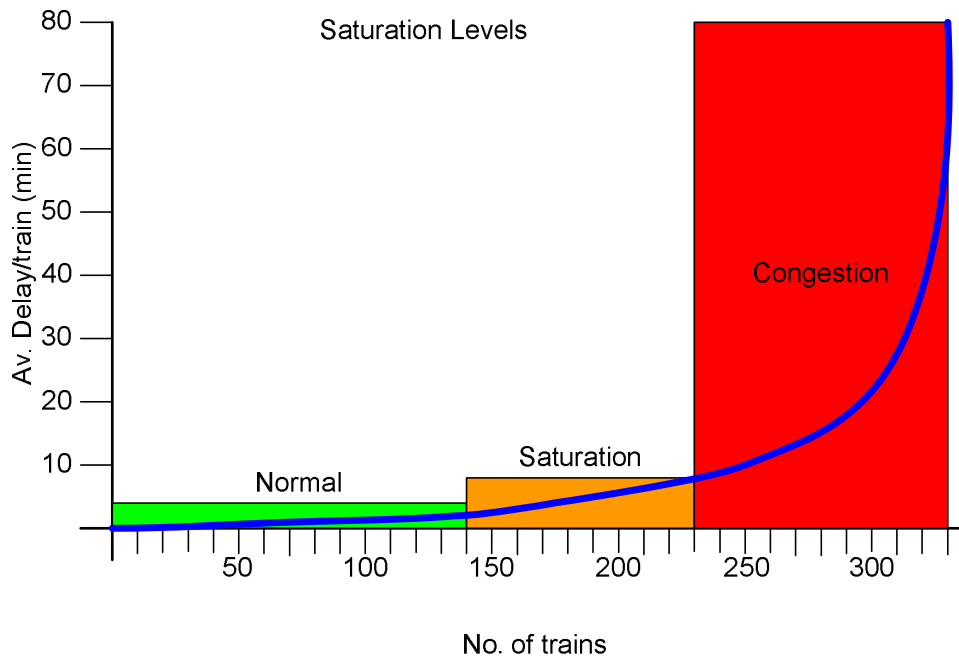


Fig. 2. Abril, M.

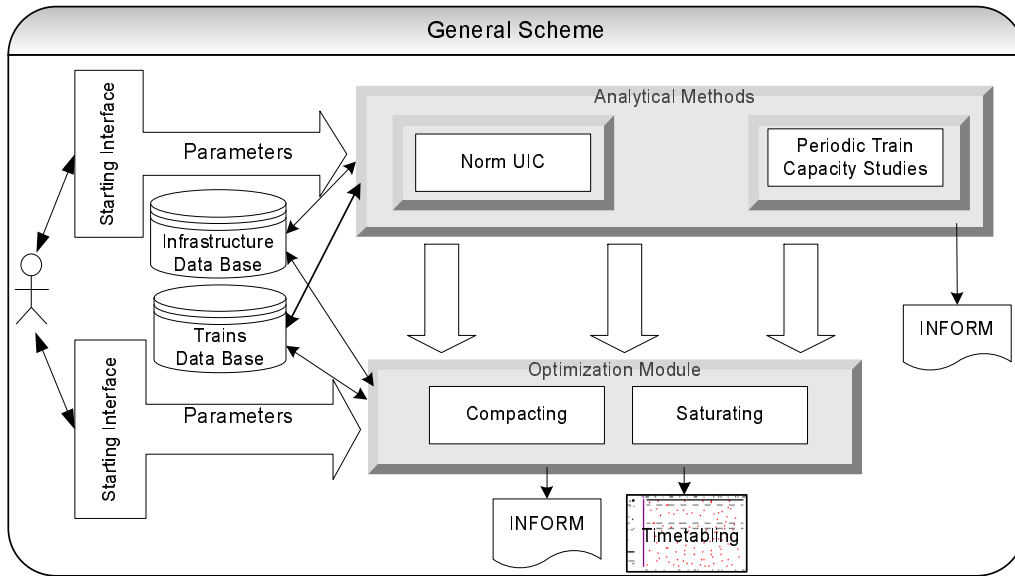


Fig. 3. Abril, M.

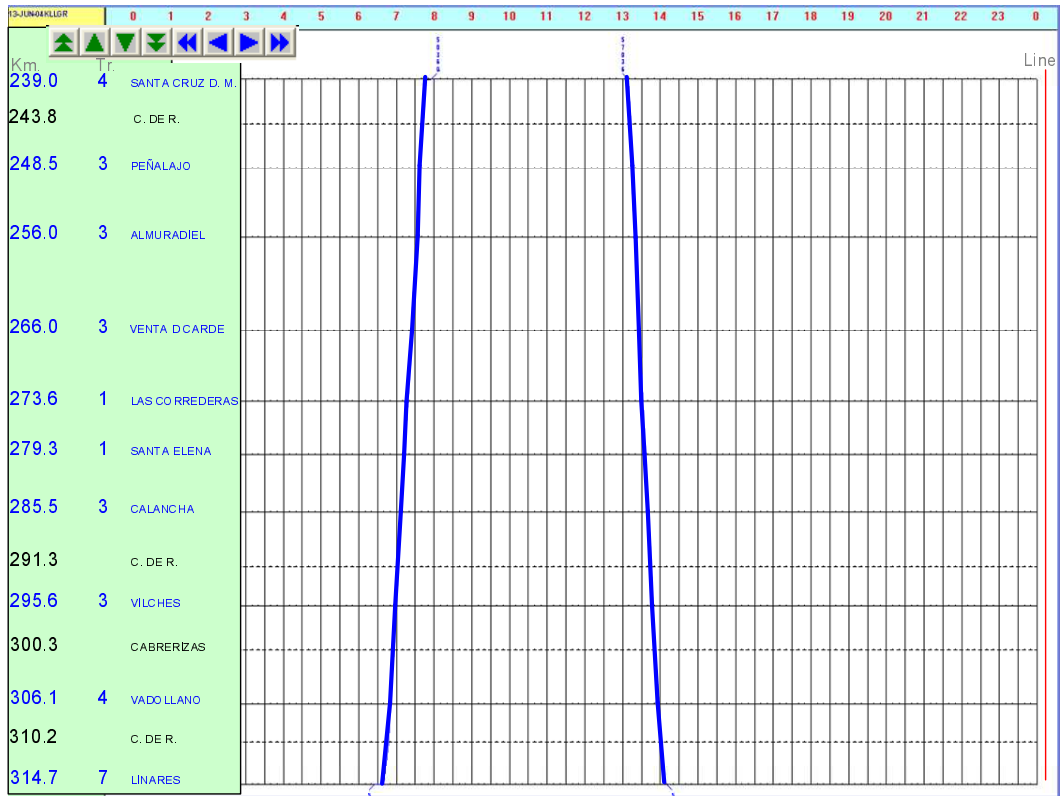


Fig. 4. Abril, M.

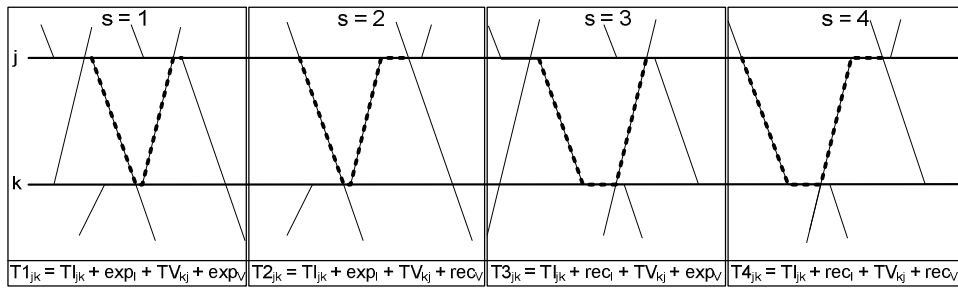


Fig. 5. Abril, M.

```

EvaluateCapacity(max_line_capacity,j,k,s)
{
    capacity = max_line_capacity;
    for(i = j, i > 1, i--)
        capacity = AdmissibleCapacityUP(capacity, i-1,i,s);

    for(i = k, i < n, i++)
        capacity = AdmissibleCapacityDOWN(capacity, i, i+1,s);

    return capacity;
}

AdmissibleCapacityUP(capacity,j,k,s)
{
    switch s
    case 1:
        sectionCapacity = min(capacity, max( $T_{3jk}$ ,  $T_{4jk}$ ));
        ActualizarCrossingType(s);
    case 2:
        sectionCapacity = min(capacity, max( $T_{1jk}$ ,  $T_{2jk}$ ));
        ActualizarCrossingType(s);
    case 3:
        sectionCapacity = min(capacity, max( $T_{3jk}$ ,  $T_{4jk}$ ));
        ActualizarCrossingType(s);
    case 4:
        sectionCapacity = min(capacity, max( $T_{1jk}$ ,  $T_{2jk}$ ));
        ActualizarCrossingType(s);
    return sectioncapacity, s;
}

AdmissibleCapacityDOWN(capacity,j,k,s)
{
    switch s
    case 1:
        sectionCapacity = min(capacity, max( $T_{2jk}$ ,  $T_{4jk}$ ));
        ActualizarCrossingType(s);
    case 2:
        sectionCapacity = min(capacity, max( $T_{2jk}$ ,  $T_{4jk}$ ));
        ActualizarCrossingType(s);
    case 3:
        sectionCapacity = min(capacity, max( $T_{1jk}$ ,  $T_{3jk}$ ));
        ActualizarCrossingType(s);
    case 4:
        sectionCapacity = min(capacity, max( $T_{1jk}$ ,  $T_{3jk}$ ));
        ActualizarCrossingType(s);
    return sectioncapacity, s;
}

```

Fig. 6. Abril, M.

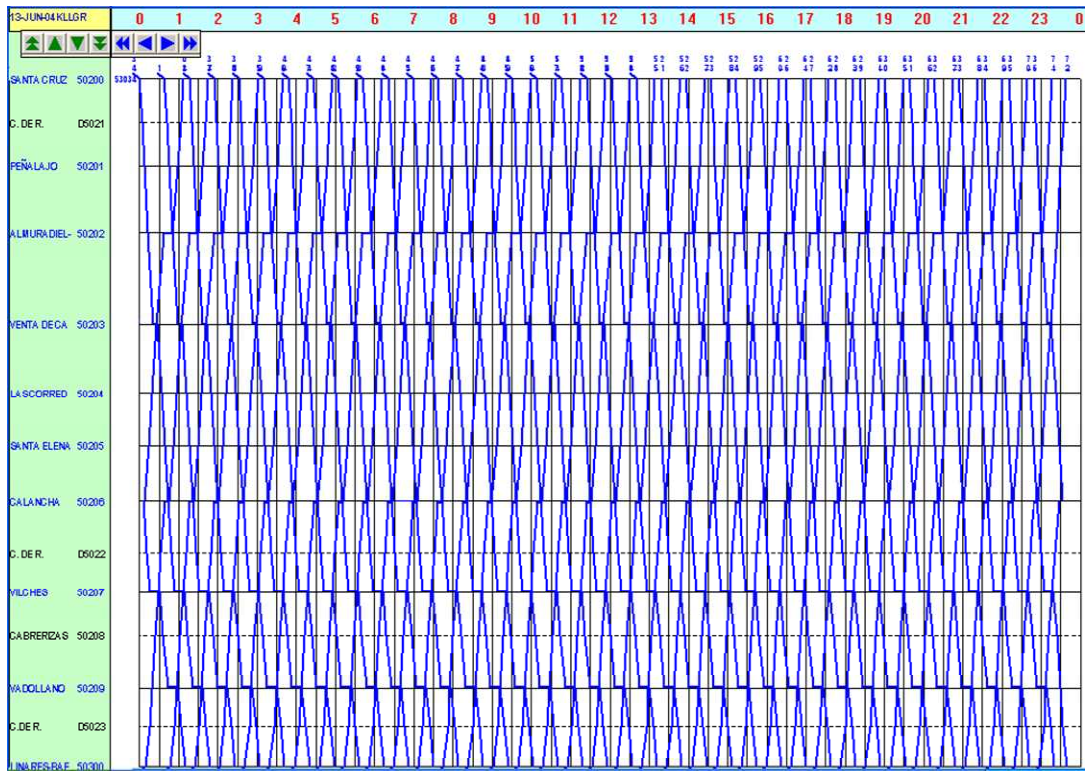


Fig. 7. Abril, M.

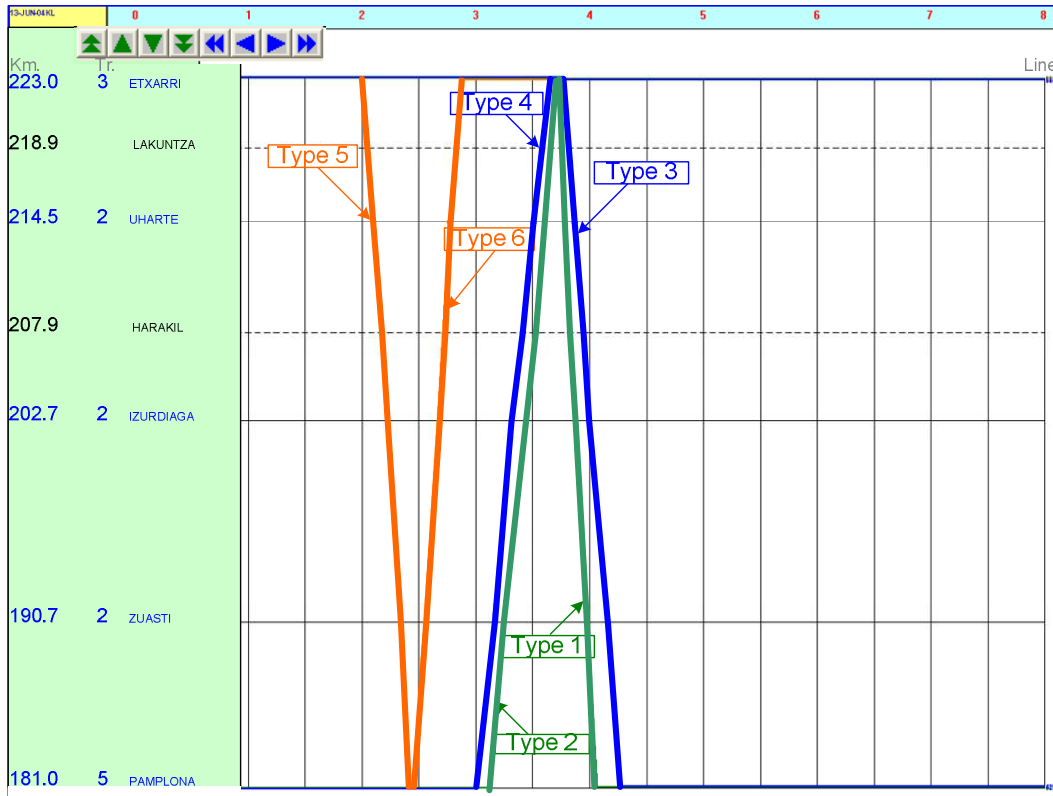


Fig. 8. Abril, M.

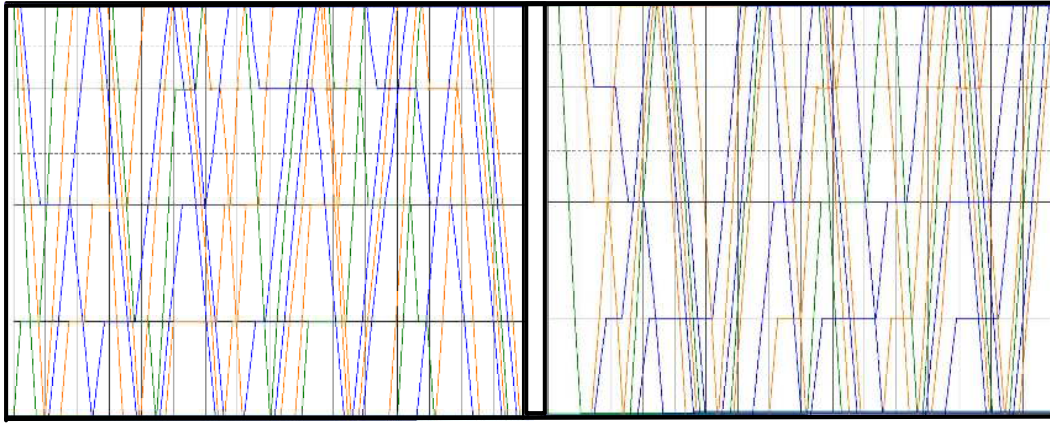


Fig. 9. Abril, M.

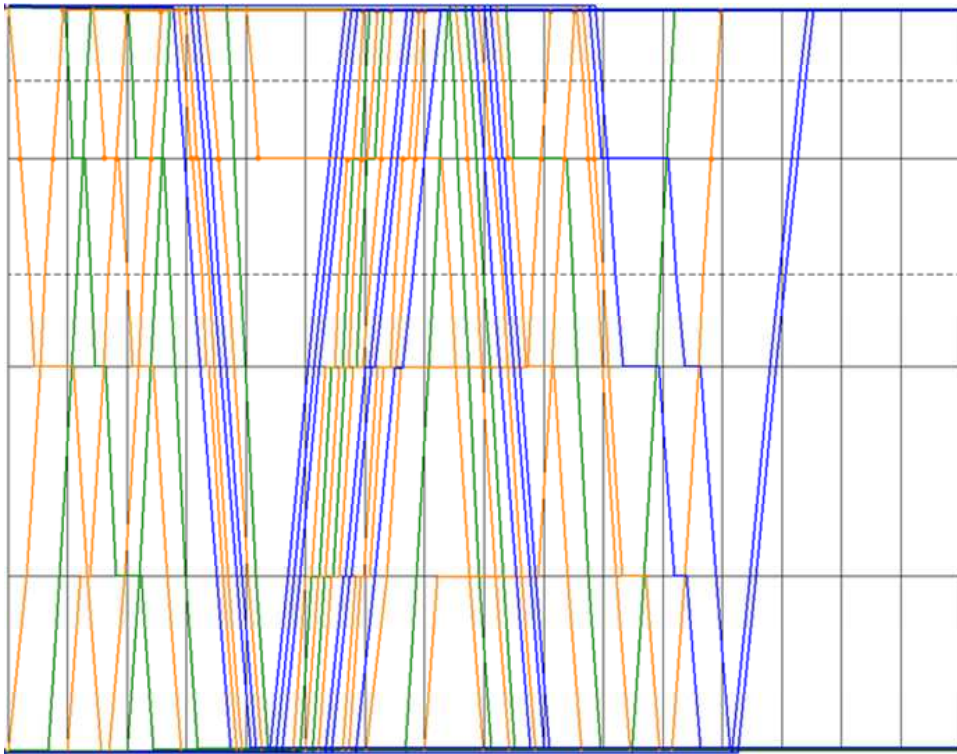


Fig. 10. Abril, M.

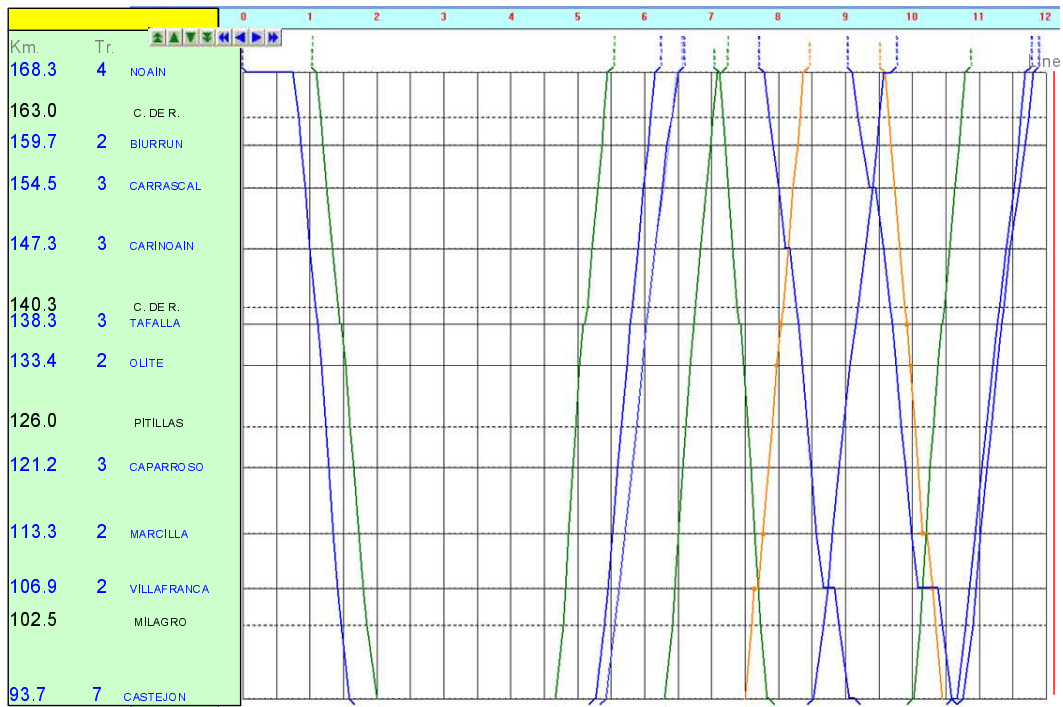


Fig. 11. Abril, M.

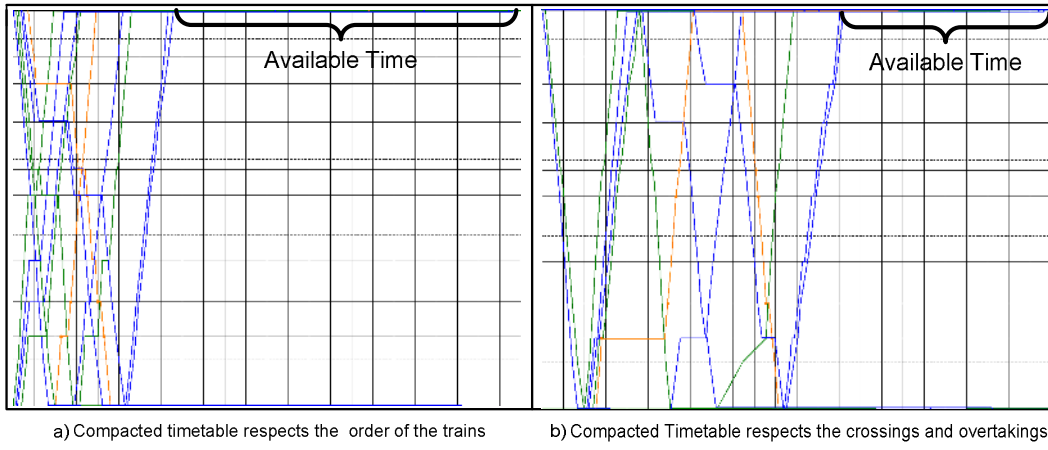


Fig. 12. Abril, M.

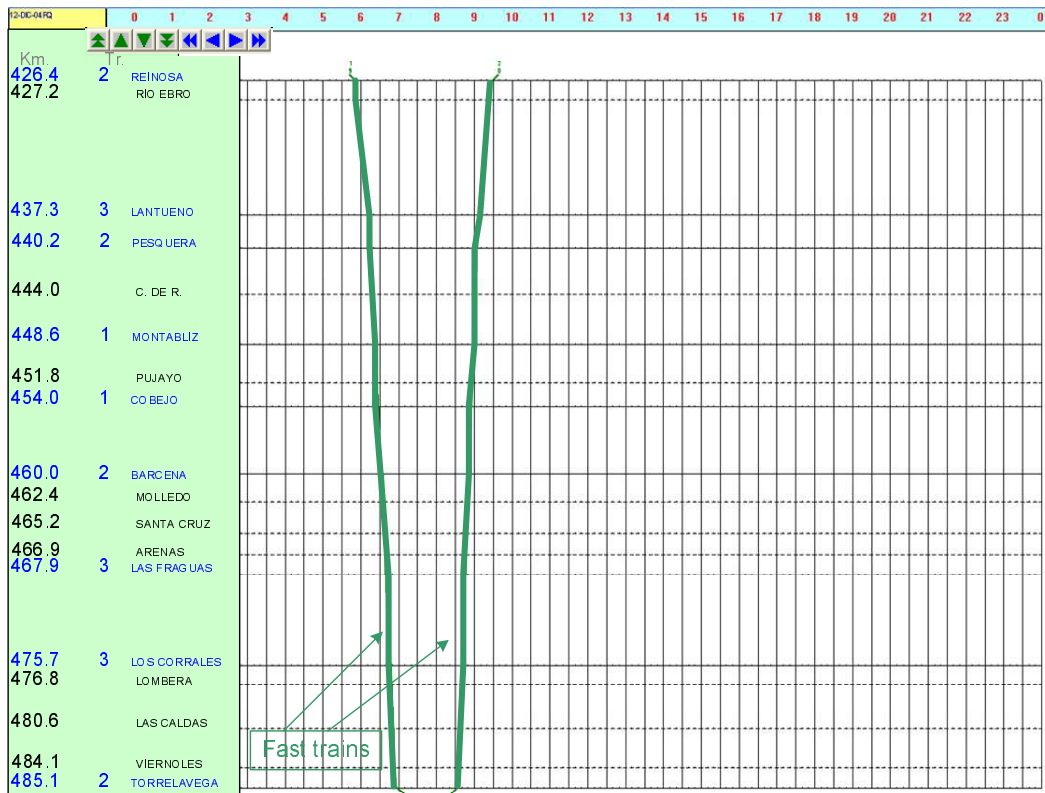


Fig. 13. Abril, M.

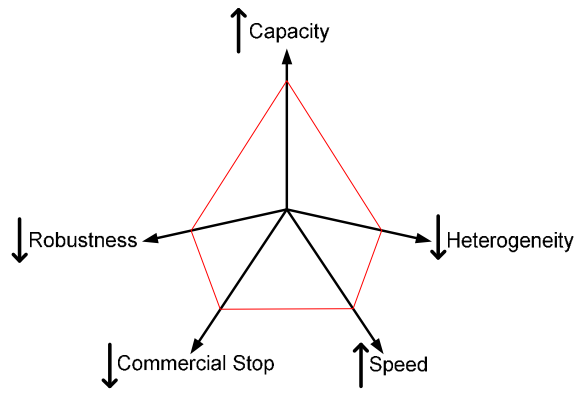


Fig. 14. Abril, M.

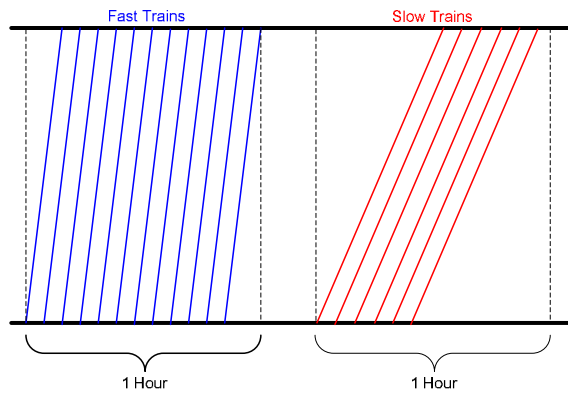


Fig. 15. Abril, M.

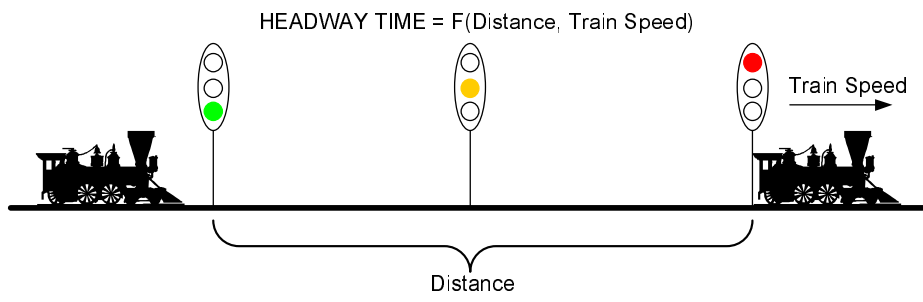


Fig. 16. Abril, M.

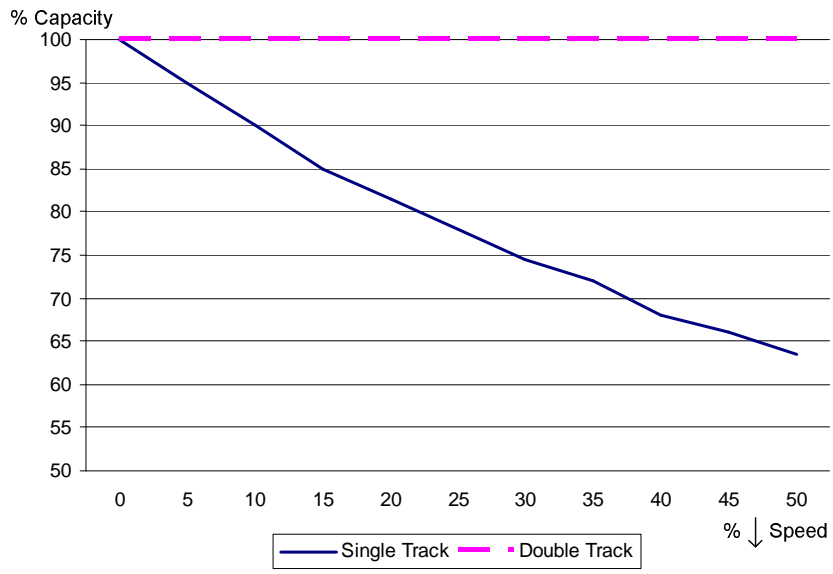


Fig. 17. Abril, M.

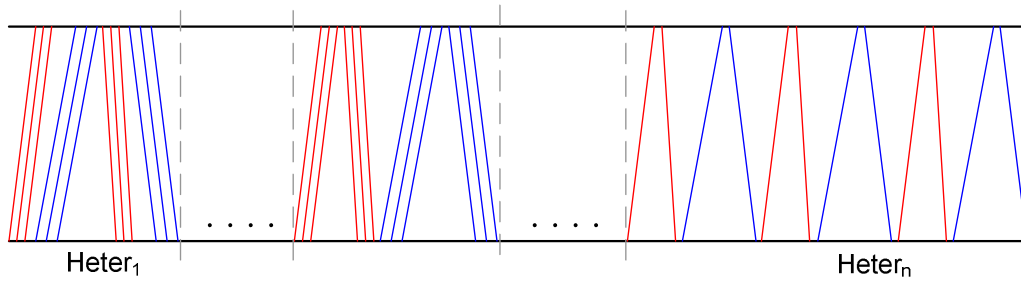


Fig. 18. Abril, M.

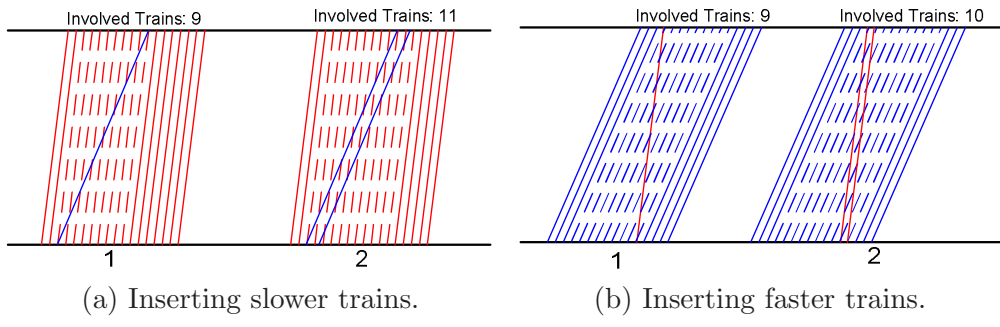


Fig. 19. Abril, M.

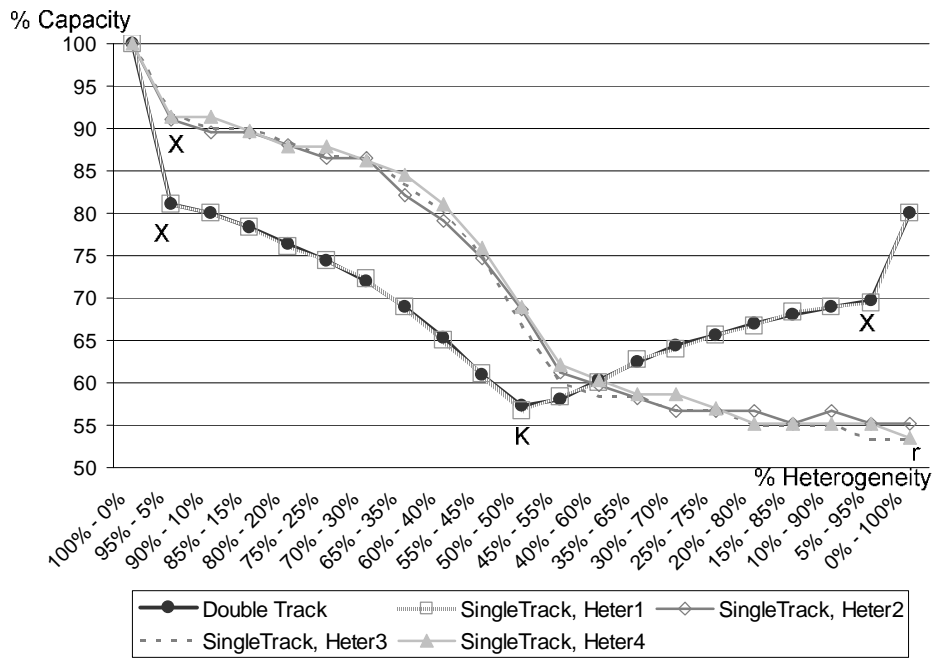


Fig. 20. Abril, M.

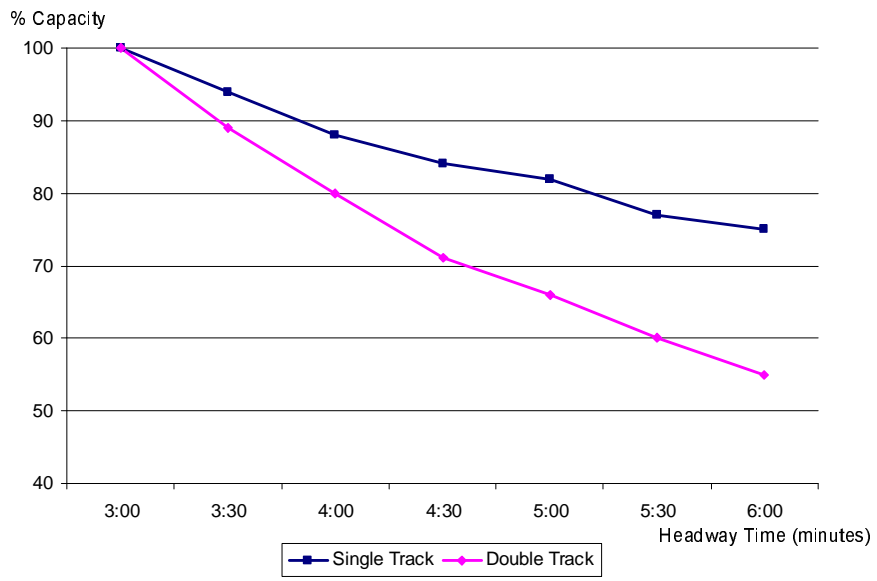


Fig. 21. Abril, M.

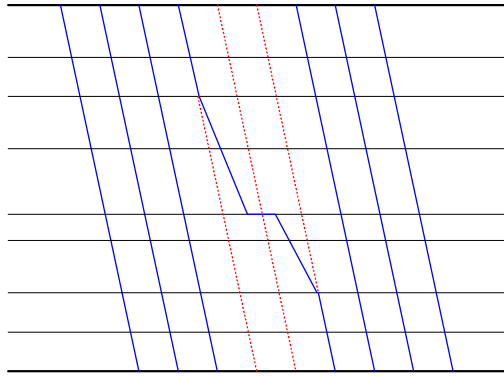


Fig. 22. Abril, M.

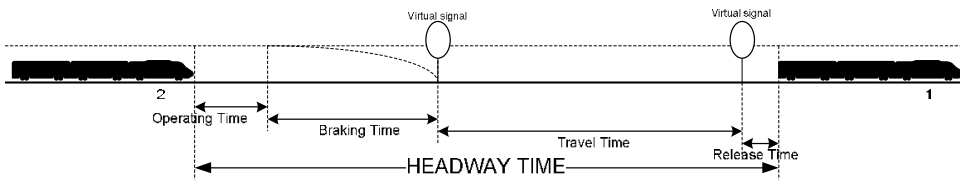


Fig. 23. Abril, M.

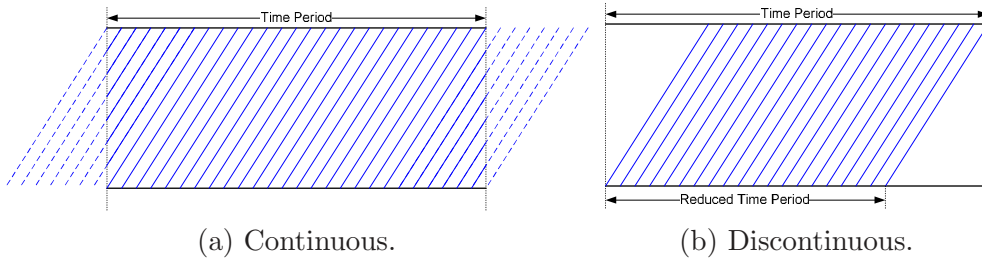


Fig. 24. Abril, M.

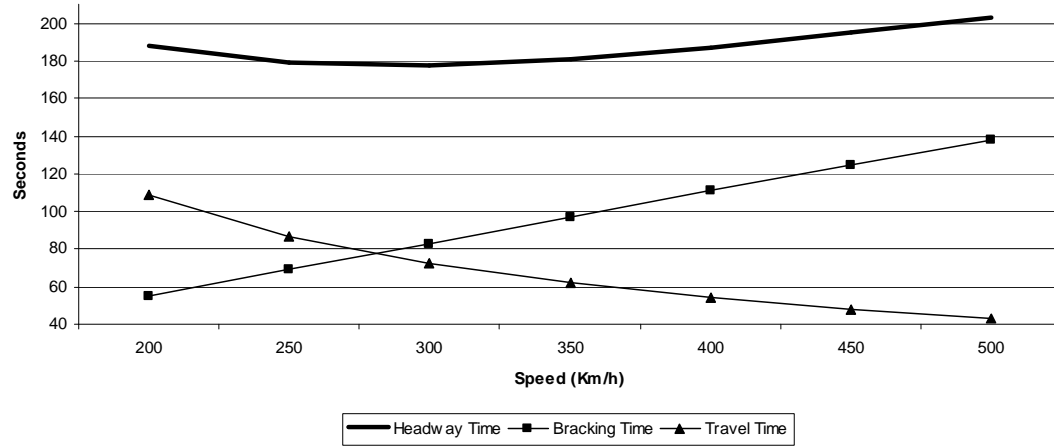


Fig. 25. Abril, M.

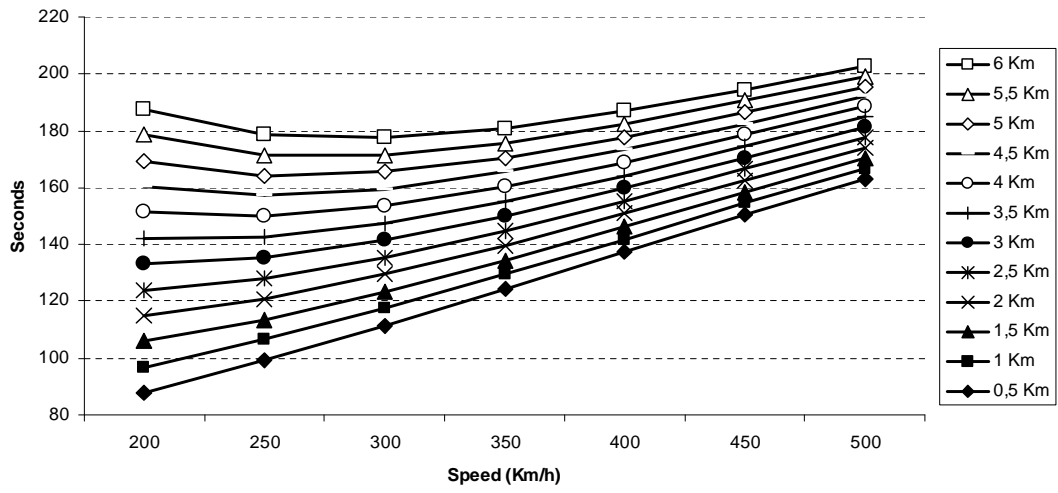


Fig. 26. Abril, M.

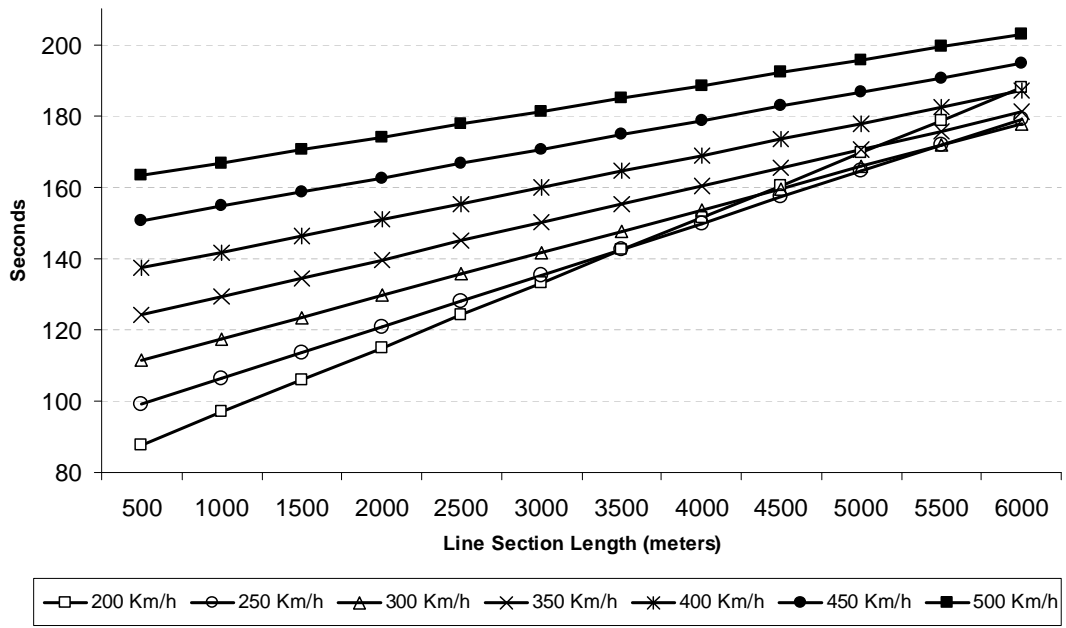
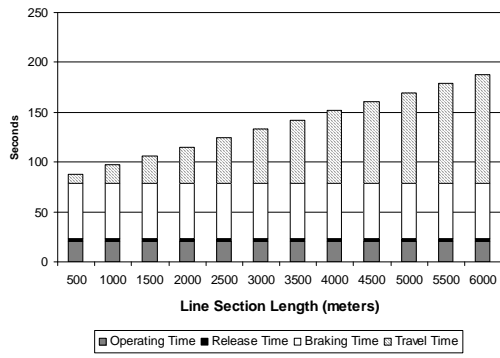
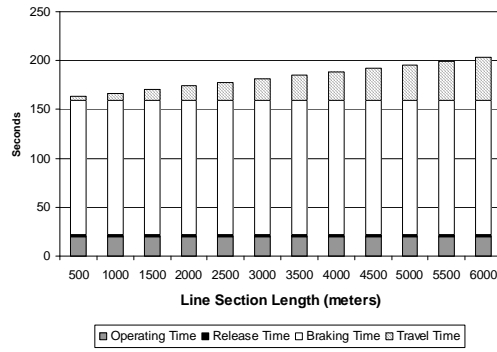


Fig. 27. Abril, M.



(a) Speed = 200 Km/h.



(b) Speed = 500 Km/h.

Fig. 28. Abril, M.

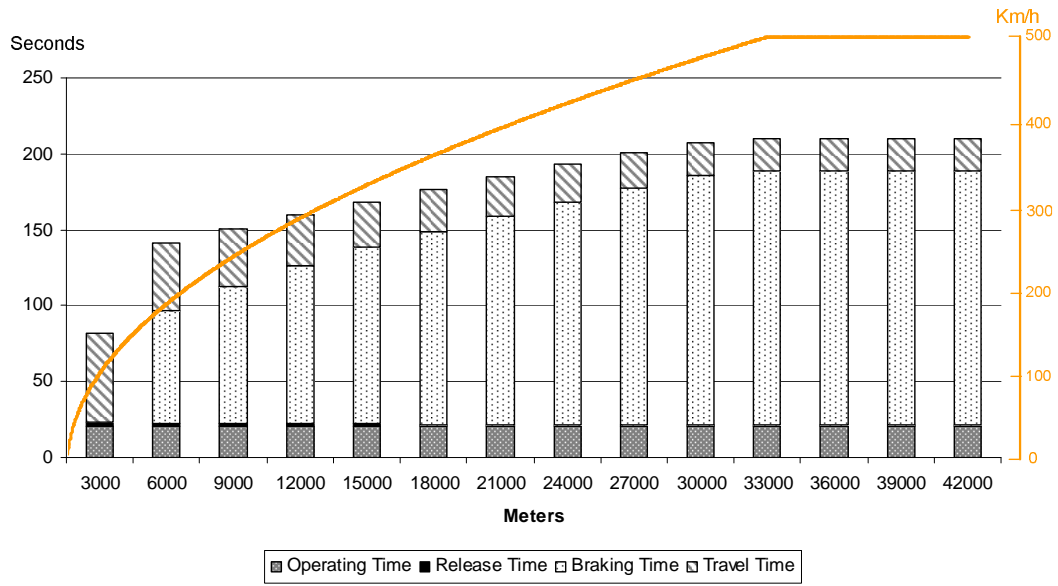


Fig. 29. Abril, M.

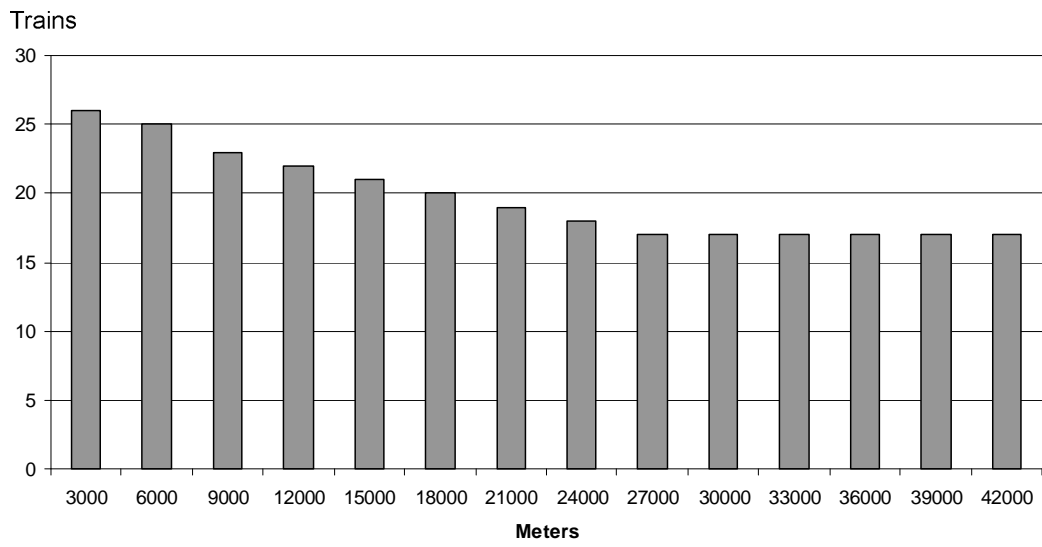
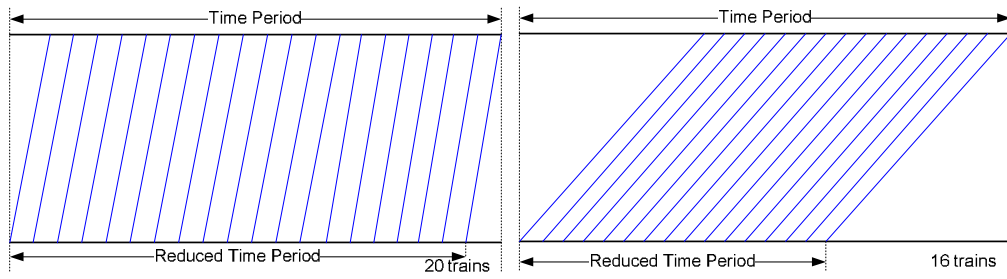


Fig. 30. Abril, M.



(a) Faster trains. More trains in the *time period*, but more separated. (b) Slower trains. Fewer trains in the *time period*, but closer together.

Fig. 31. Abril, M.

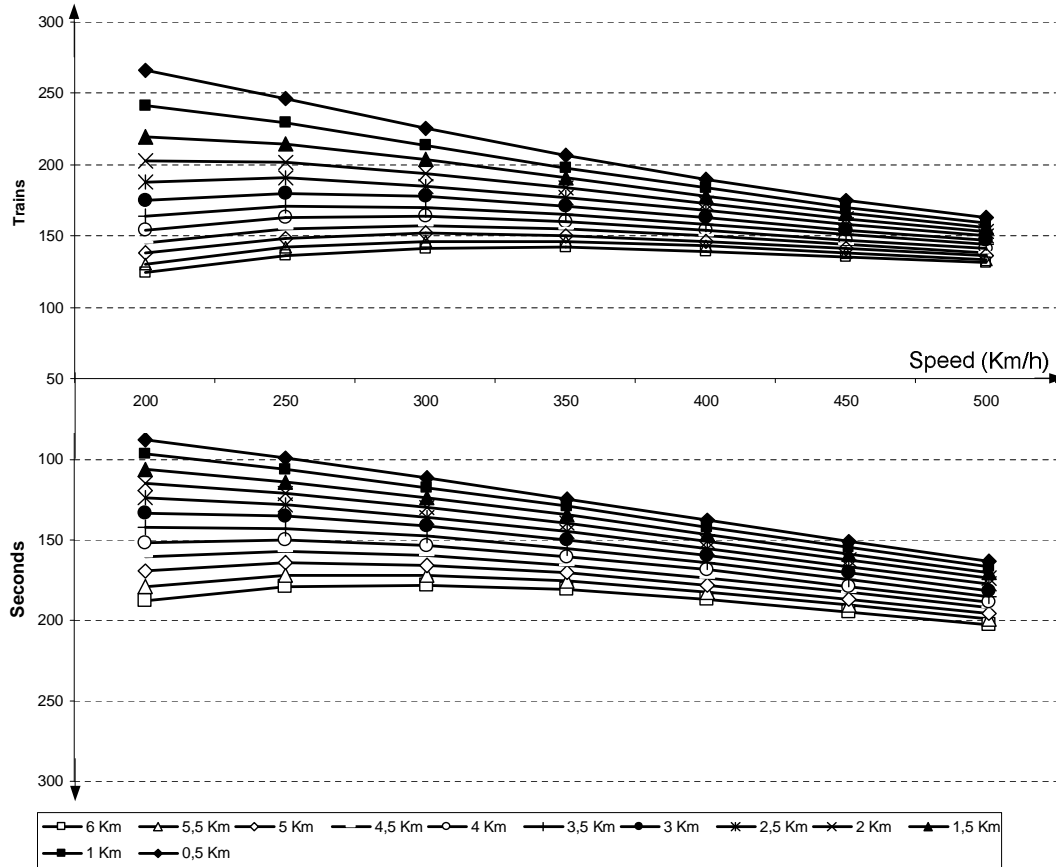


Fig. 32. Abril, M.

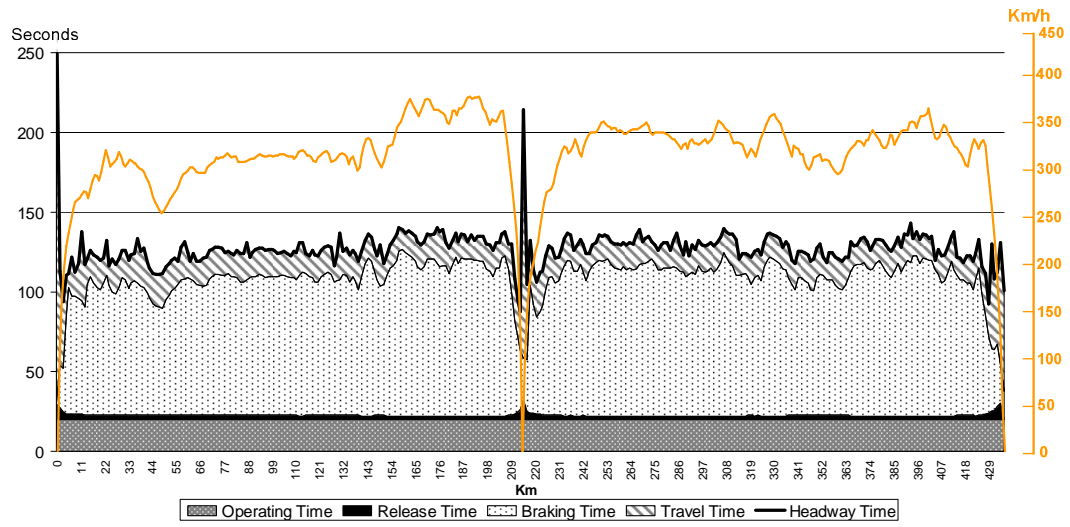


Fig. 33. Abril, M.

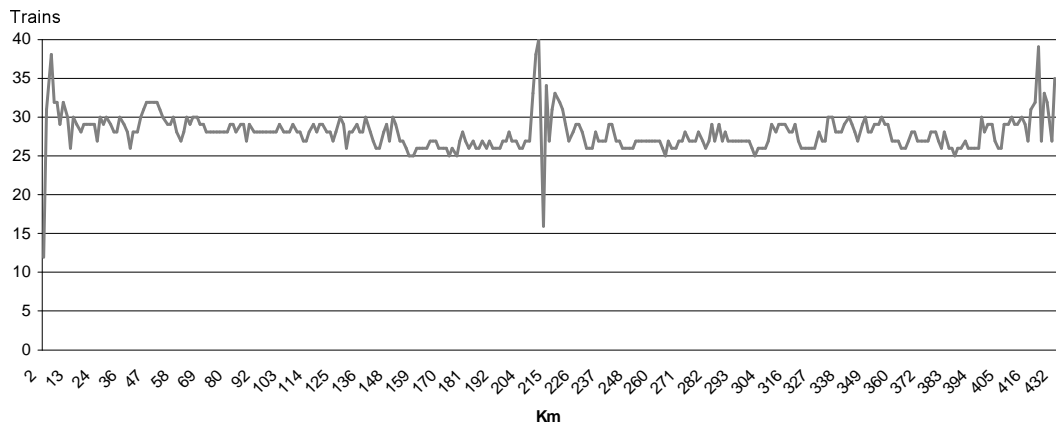
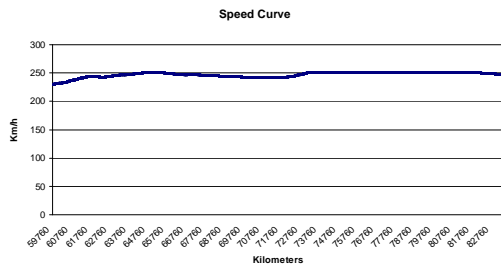
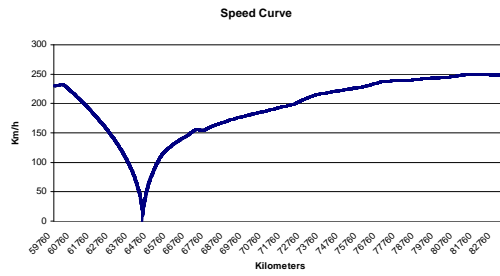


Fig. 34. Abril, M.



(a) Speed curve without stop.



(b) Speed curve with stop.

Fig. 35. Abril, M.

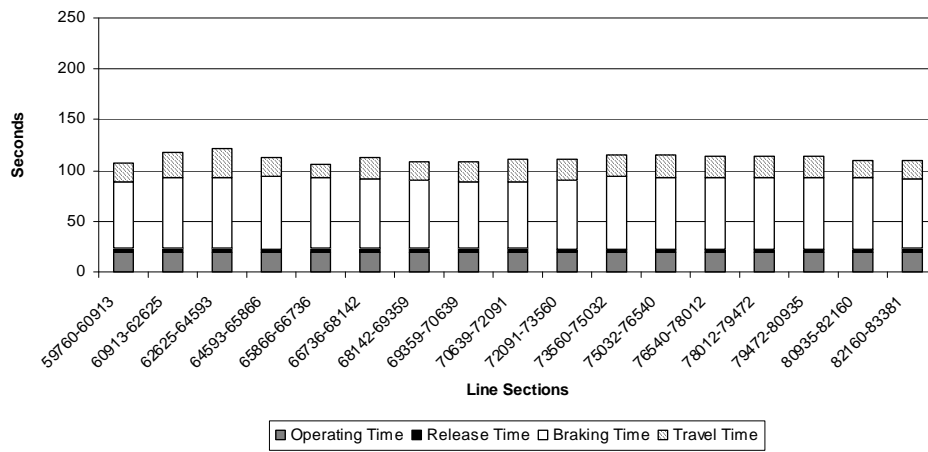


Fig. 36. Abril, M.

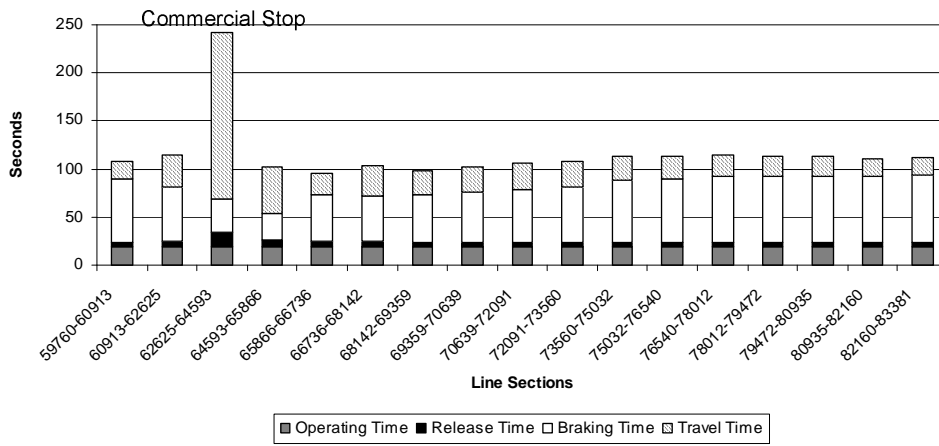
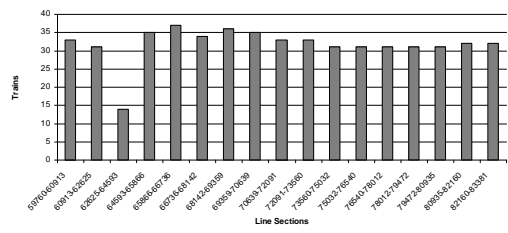
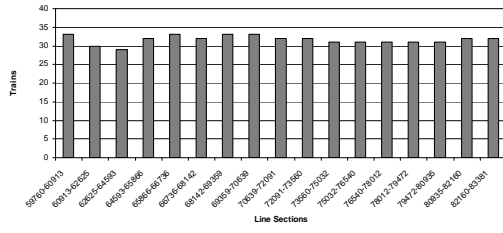


Fig. 37. Abril, M.



(a) Maximum capacity without stop.

(b) Maximum capacity with stop.

Fig. 38. Abril, M.

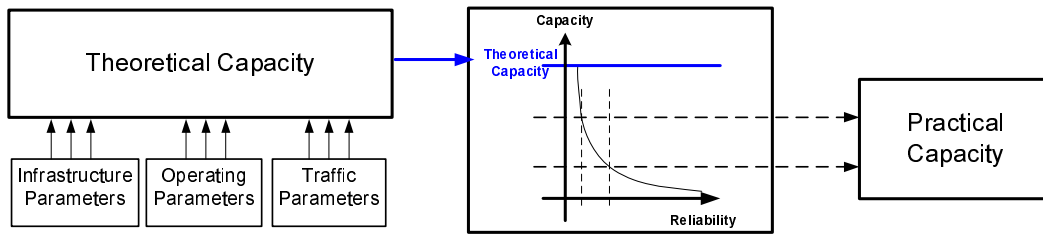


Fig. 39. Abril, M.