UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II

POLO DELLE SCIENZE E DELLE TECNOLOGIE

SCUOLA DI DOTTORATO IN INGEGNERIA CIVILE

DOTTORATO DI RICERCA IN

“ INGEGNERIA DEI SISTEMI IDRAULICI, DI TRASPORTO E TERRITORIALI (XXVII CICLO) ”

The definition of a model framework for the planning and the management phases of the rail system in any kind of service condition.

Antonio Placido

Coordinatore di dottorato/PhD course coordinator:

Prof. Elvira Petroncelli

Tutor/Supervisor:

Prof. Luca D’Acierno

Controrelatore/Examiner:

Prof. Gennaro Bifulco

CONTENTS

CHAPTER 1: INTRODUCTION

1.1 Research motivation

Railway systems represent the backbone of public transportation systems. Indeed, thanks to their main characteristics (i.e. the use of exclusive lanes, the constrained driving and the signalling system), trains manage to keep higher travel speeds and lower headways than other public transportation systems. As a consequence, especially in high density contexts, rail and metro networks are vital to fulfil travel demand needs whilst generating relatively low negative externalities. However, due to new European regulations, the constant increase in passenger flow and the simultaneous deep economic crisis, the planning and the management phase of railway systems are becoming more and more difficult especially in degraded operational conditions. Recently, many researchers belonging to different fields of transportation engineering and Operations Research proposed innovative methods to improve the design of the service (i.e. stable and robust timetables) as well as the management of the network in perturbed conditions (Cacchiani et al., 2014). In particular, different approaches were proposed to tackle the so-called 'rescheduling problem', namely the definition of intervention strategies whose aim is to re-establish ordinary conditions after the occurrence of disturbances or disruptions to the service. In this context, the adoption of simulation models is extremely useful to find solutions for promptly reacting to unforeseen events. However, the majority of the models presented in the literature focuses on the possibility to determine feasible strategies in short time trying to optimise operational aspects (e.g. minimising the number of delayed trains) thinking that this corresponds to the optimal solution also from passengers' standpoint. Unfortunately, this strategy often results in a lower service quality offered and neglects customers' needs (D'Acierno et al., 2012).

Moreover, most models are based on macro-optimisation procedures which, despite providing outputs in short time, require numerous approximations concerning infrastructure representation and train service. As a consequence, the intervention strategies obtained could not be feasible or stable during the real implementation phase.

Stability and robustness is another important issue. Indeed, rail service is affected by numerous random events and any activity, although planned, could differ notably from what is expected (Quaglietta et al., 2013). Therefore, effects of stochasticity on the service, especially on recovery solutions, should be considered so as to evaluate also the effectiveness of rescheduling strategies.

The simulation of travel demand is a further key element which is necessary to take into account. Not only does it enable to assess the quality perceived by passengers in the case of both ordinary and disrupted conditions but, particularly in high density contexts where passenger flows are very high, it gives the possibility to analyse the interaction with the service (Kunimatsu et al., 2012). In particular, passengers on the platform influence dwell times at stations and this represents one of the main disturbances of high frequency service lines. As a result, exhaustive stochastic analyses of rail operations should consider the effects of delays resulting from the randomness of both kinematic parameters and travel demand influence.

In practise, the management of railway services in perturbed conditions is still based on the experience of dispatchers and actual implementation of models proposed in the literature are scarce. Basically, new methodologies do not manage to give reliable response to any kind of event and do not assure a stable control of the network.

1.2 Thesis contributions and objectives

The target of this thesis is the definition of an off-line procedure for managing the rail network in any kind of service conditions focusing on failure events. In particular, the methodology is based on a microscopic simulation approach which considers both rail operations and passenger flows. Basically, the idea is to simulate the network with the higher level of details without neglecting travel demand which is indeed assigned to the service.

The benefits provided by this approach are numerous, namely:

1. it is possible to look for intervention strategies, which optimise passenger satisfaction and do not focus just on operational aspects;
2. by simulating passenger behaviour on the platforms, the procedure enables the assessment of the dynamic interaction between service and user flows. In other words, the model estimates the dwell time at stations as flow dependent providing important information about the influence of passengers on the service;
3. another advantage is the possibility to evaluate crowding levels within the trains or at stations resulting in a more suitable planning of the service as well as a better estimation of the comfort experienced by travellers.
4. the dynamic assignment, although increases the complexity of the model, is extremely useful. In fact, in this way demand peaks, temporary capacity variations, temporary over-saturation of supply elements, and formation and dispersion of queues can be considered;
5. the adoption of proper sensitivity analyses provides the assessment of robustness and effectiveness of planned recovery solutions not only in terms of operational service but also considering customer satisfaction.

The research work can be mainly divided in two phases. The first one concerns the specification of the decision support system and all models which are part of it. The second by contrast, is related to the definition of an application for the dynamic assignment of passenger flow to the rail service and the definition of dwell times depending on the number of passengers at station.

As regards the first phase, the whole procedure is formulated as a bi-level multidimensional optimisation model which is composed of four sub-models: a Failure Model, a Service Simulation Model, a Supply Model and a Travel Demand Model. In order to increase the service quality, the objective function is expressed through the user generalised cost (Cascetta, 2009) perceived by customers during their travel and, evidently, it has to be minimised.

The Failure Model evaluates the failure scenarios which are worth analysing. In particular, through the adoption of RAMS (Reliability, Availability, Maintainability and Safety) techniques (Cenelec, 1999), it gives the possibility to select the breakdown contexts with the higher probability of occurrence.

The Service Simulation Model analyses rail traffic and system performance during both ordinary and perturbed conditions by means of a microscopic simulation of the network. According to the target of the analysis, the simulation can be either deterministic or stochastic.

The Supply Model is instead dedicated to the definition of performances of all public transportation systems within the study area. In fact, rail and metro lines, particularly within cities, are part of the public transportation system and cannot be considered individually. Hence, knowing the characteristics of the other transport modes can also provide a better estimation of the arrival rate at each station.

The Travel Demand Model is the most innovative part of the whole procedure. It is divided into other two sub-models, namely a Pre-Platform Model and On-Platform Model. The first one estimates the number of passenger arriving at stations as a result of the interaction with the Supply Model. This causes a fixed point problem which has been largely dealt with in the literature (see Cantarella, 1997; Cascetta, 2009). Basically, the Pre-Platform Model reproduces the choice process made by passengers who evaluate among all possible alternatives (i.e. different transport modes) the one which maximises their utility.

The On-Platform Model works on the dynamic assignment of passenger flows to the rail service. In particular, the model simulates passenger behaviour on the platform considering the maximum capacity of each train and estimating the dwell time necessary to complete the boarding/alighting process. In this way, travel demand is simulated dynamically according to rail service performances which, in turn, are influenced by passenger flows. As a consequence, this interaction generates another fixed point problem.

Hence, the resolution of the whole procedure consists in solving a double fixed point problem which has required an in-depth analysis about the mathematical assumptions and the resolution techniques to solve it.

However, all railway microscopic simulation software packages focus just on the simulation of train movements within the network and neglects travel demand. Therefore, the second phase of the thesis has concerned the definition of an application developed in C++ language for assigning travel demand to the rail service working in combination with microscopic simulation software. To this purpose the architecture of the OPM 1.0 (On-Platform Model) tool and its internal module DwTE 1.0 (Dwell Time Estimation) has been presented.
Both require input data as text files related to infrastructure, rolling stock, travel demand and operational service. In particular, OPM 1.0 is composed of the following modules:

- a '*Travel demand module*' for the definition of passenger flow on the platform at each station;
- a '*Rolling stock module*' which describes the main features of rail convoys in terms of fleet composition, number and capacity of coaches, number of doors per coach and so on;
- a '*Rail service module*' which includes information about the simulated rail service such headways, running times, empty movement etc.

Additionally, in case also DwTE 1.0 is launched further modules must be considered, that is:

- a '*Passenger flow module*' which considers the number of passengers who can actually board the train according to trains' capacity (this information is obtained by OPM 1.0);
- a '*Station configuration module*' specifying station characteristics (i.e. location of stairs and elevators);
- a '*Dwell time estimation module*' which defines the time trains has to stop within the station as function of the number of boarding/alighting passengers per door.

As outputs, the application provides information about passenger trips, load diagrams, platform congestion and crowding levels within trains as well as dwell time values at stations.

In order to validate the tools and verify the benefits of the proposed procedure, in the last part of this thesis several applications are presented. The majority of them has concerned the Line 1 of Naples (Italy) metro system. Results have shown the importance of considering service quality during the management of the rail service, especially when failures or breakdowns occur. In fact, considering just operational aspects can bring to the definition of intervention strategies which are often far from satisfying customer needs. In this context, the proposed approach gives the possibility to have a precise estimation of user generalised costs and select the alternatives which maximise the utility perceived by passengers. In addition, it also provides indications about how these strategies can be affected by fleet compositions, breakdown contexts and travel demand levels.

Another important result is the possibility to adopt this procedure for the robustness and stability evaluation of recovery solutions. Indeed, by means of stochastic simulations of the network, sensitivity analyses based on the variability of both kinematic parameters (i.e. acceleration and speed) and dwell times (as function of travel demand) are also presented.

However, especially in the case of concession regimes where public authorities try to pursue the difficult task of considering both public (sustainability, accessibility, employment, etc.) and commercial interests (profit, return on investment, growth), the analysis of just user generalised costs is not sufficient to achieve good levels of efficiency, effectiveness and productivity of the service at the same time. Hence, a more complex objective function is proposed which takes into account also efficiency (e.g. operational costs) and effectiveness (e.g. number of passengers/number of offered seats) indexes.

Due to the complexity of the problem, the whole procedure is not feasible to obtain results in short time and, as already said, it is therefore proposed as an ‘off-line’ methodology. Furthermore, in real dimension networks, the number of solutions, which have to be investigated could be so high that it is impossible to evaluate all of them through the adoption of an exhaustive approach. Thus, the combination of the proposed microscopic method with a macro-optimisation model is provided so as to show the benefits in terms of computational time when dealing with rescheduling problems.

Finally, the last application concerns the feasibility of the procedure also in the case of conventional railway lines. Indeed, since this kind of system can be considered as closed network (i.e. the interaction with other public transportation system is lower than metropolitan lines), by introducing some changes to the On-Platform Model, different disrupted events on the regional line ‘Napoli-Formia’ (in Italy) are analysed focusing on users’ perspective.

1.3 Thesis outline.

This section gives a short introduction to each chapter of this thesis.

Chapter 2 provides a general overview of railway systems through the description of infrastructural and operational components as well as rules which regulate the system. The first part is dedicated to the analysis of the European and Italian legislation concerning railways showing the opening process of the market proposed by the European Union and the importance given by Member States to service quality. Then, the chapter explains in detail all elements related to infrastructure, rolling stock, computer-based control systems (e.g. signalling equipment and interlocking) and train operations (e.g. timetable, capacity consumption, etc.).
This analysis is necessary for implementing microscopic simulation models which require a high detailed network representation.

Chapter 3 deals with the state of art of railway traffic simulation models and their application in practical fields. To this purpose, the difference in terms of network representation approach (macroscopic – mesoscopic – microscopic), processing technique of event (synchronous – asynchronous) and statistical

assumptions (deterministic – stochastic) are explained. Then, several commercial and academic models are described so as to show their applicability in railway contexts.
The last part of chapter 3 concerns an overview of state-of-knowledge in train traffic management in the case of disturbances or disruptions to the ordinary service.

Chapter 4 instead shows the mathematical formulations of the proposed procedure. Each model is therefore described focusing on the relations among them. In particular, the second part of this chapter is dedicated to the analysis of the architecture of the On-Platform Model. The resolution of the fixed point problem resulting from the evaluation of dwell times is also dealt with showing the feasibility of two algorithms.

Chapter 5 presents the framework of the OPM 1.0 tool for the dynamic assignment of travel demand to the rail service as well as DwTE 1.0, whose aim is the evaluation of dwell times as flow dependent. In addition, the operation mode of both tools is also provided showing the format of input and output files.

In chapter 6, several applications of the whole procedure on the Line 1 of Naples metro system shows the feasibility and the benefits of this approach particularly in the case of high density contexts. The adoption of the proposed model under stochastic assumptions is also presented so as to assess the stability of intervention strategies previously evaluated by means of a deterministic approach.
Further improvements to the model are also introduced revealing the possibility to consider a more complex objective function, increase the computational efficiency (macro – micro combination) and enlarge the feasibility of the approach also to conventional rail lines.

Finally, chapter 7 summarises all activities carried out in this work and reports conclusions and possible research prospects.

CHAPTER 2: A GENERAL OVERVIEW OF THE RAILWAY SYSTEM

Railway networks are very complex systems composed of several elements (i.e. infrastructure, rolling stock, signalling system) and characterised by the interactions of different subjects (Infrastructure Managers, Rail Operators, Customers). This chapter gives a general overview of the railway system and provides essential information for the comprehension and the development of the system of models discussed in this thesis.

First of all, a description of the main European legislation concerning public transport systems and railways is presented. In this way, it is shown how the railway sector has been completely reorganised in the last twenty years and why railway undertakings are more and more interested in offering high levels of service quality.

The second and the third part are dedicated to a brief analysis of the infrastructure and the rolling stock, including the description of train motion equations and their resolution. In the fourth part of this chapter, train space theory and signalling systems are summarised while the last paragraph deals with timetabling and rail capacity allocation.

2.1 Railway legislation in Europe.

The European Union has the ambitious strategy of creating a single, efficient and competitive market for rail throughout Europe. To achieve this target, several laws have been proposed, which aim at opening rail markets, promoting competition, tackling barriers to market entry and fostering interoperability.

In particular, this paragraph deals with the successive steps made by the European Commission to improve efficiency, attractiveness and productivity of railway networks taking into account also their importance within the public transportation system. Indeed, each law is briefly described underlining the contents and its main effects to the Member States. Finally, a short analysis of the Italian railway legislation is carried out.

2.1.1 Regulation (EEC) nr 1191/69 of the Council of 26 June 1969 (modified by Regulation (EEC) nr 1893/91 of 20 June 1991)

This regulation is very important since it marks the beginning of a new conception of public transport service, including railways. First of all, it defines the '*public service obligations*' as the obligations which the transport undertaking, if it were considering its own commercial interests, would not or would not assume to the same extent or under the same conditions. They can be divided in the obligation to operate, the obligation to carry and tariff obligations. The first one is the obligation imposed to transport undertaking to ensure the provision of a transport service satisfying fixed standards of continuity, regularity and capacity. Obligation to carry means that every transport undertaking has to accept and carry passengers or goods at specified rates and subject to specified conditions. Tariff obligation is instead the obligation to apply, in particular for certain categories of passengers or goods, fixed rates or approved by any public authority which are contrary to the commercial interests of the undertaking.
According to the regulation, Member States can decide if terminate or keep, entirely or in a part, a public service obligation. In this case the competent authority, having regard to public interests, has to sign a 'service contract' with the transport undertaking whose aim is offering a sufficient transport service to the community. In particular, a public service contract must contain:

- the nature of the service to be provided together with the standards of continuity, regularity, capacity and quality;
- the price of the service covered by the contract;
- the rules concerning amendment and modification of the contract, in particular to take account of unforeseeable events;
- the period of validity of the contract;
- the penalties in the event of failure to comply with the contract.

Basically, this regulation has been vital to reduce the huge cost for both financing and managing public transportation services. Moreover, thanks to the introduction of the service contract, public authorities can impose rules to transport undertakings and force them to keep high quality standards.

2.1.2 Council directive of 29 July 1991 on the development of the Community's railways

The Directive 91/440/ECC concerns specifically the railway sector. It deals with the necessity of promoting a single railway market all over the Community. Moreover, the directive wants to find solutions to increase the efficiency of railways. Basically, it proposes four actions:

- ensuring the management independence of railway undertakings from the State;
- separating the management of railway operations and infrastructure from the provision of railway transport services. In particular, the separation of accounts is compulsory while the organisational or institutional separation can be optional;
- improving the financial structure of undertakings;
- ensuring access to the networks of Member States for international grouping of railway undertakings and for railway undertakings engaged in the international combined transport of goods.

Thanks to these proposals, the Directive has given foundation to the liberalisation of the railway market, increasing competitiveness and sound financial management.

2.1.3 Council Directive 95/18/EC of 19 June 1995 on the licensing of railway undertakings.

The Directive 95/18/EC is nowadays repealed. However, it is worth being mentioned since it introduces the concept of license. Indeed, Railway Undertakings (RU), in order to provide the service, needs an authorisation (i.e. a license) issued by the competent authority of each Member State. To obtain this license, the RU has to demonstrate its good reputation, financial fitness, professional competence and coverage for its civil liability.

2.1.4 The First Railway Package

The First Railway Package adopted in 2001, is composed of several directives (2001/12/EC, 2001/13/EC, 2001/14/EC and 2001/16/EC) whose main task is that of making existing legislation more effective. In particular, it enables rail operators to have access to the trans-European network on a non-discriminatory basis and it fosters the completely opening of the rail freight market. To this purpose, the First Rail Package deals with two key factors such as the charging for the use of infrastructure and the allocation of railway infrastructure capacity. Regarding the former, the directive lays down charging principles: charges must be paid to the infrastructure managers and used to fund their business. In particular, the charge for the use of railway infrastructure is equal to the cost directly incurred as a result of operating trains. Moreover, the infrastructure charge may include a sum reflecting the scarcity of capacity and may be adjusted to take account of the cost of the environmental impact of operating the trains.

The allocation of infrastructure capacity is therefore granted by the Infrastructure Manager (IM) concerned, which is responsible for allocating the available capacity. The rights and the obligations of the IM and of the authorised RU are laid down in a contract. In particular, IM and RU may conclude a framework agreement which may not preclude use of the infrastructure by other railway undertakings and may be amended. The agreement will not specify a train path in detail but should meet the commercial needs of the authorised applicant. In principle, the framework agreement covers a period of five years, renewable for a period equal to this original duration. However, for services using specialised infrastructure, the framework agreement may be for a period of 15 years, which may be extended only in exceptional cases. To ensure close collaborations among the IM of all the Member States, the directive provides, inter alia, for the establishment of an organisation to coordinate, at international level, the allocation of capacity on different networks. This could include the establishment of international train paths. Obviously, IM must make any possible effort to meet all requests of capacity without supporting any RU. In case there is no possibility to meet all the requests, the IM has to declare the section in question congested. Afterwards, it is necessary to carry out a capacity analysis to determine the restrictions on capacity and propose alternatives. Within six months of the completion of a capacity analysis, the IM must produce a capacity enhancement plan.

All information related to the nature of the infrastructure and the conditions for accessing it, to the charging principles and to the criteria for capacity allocation is included into a network statement which has to be published by the IM.

2.1.5 The Second Railway Package

The Second Railway Package has introduced new legislation in terms of safety and interoperability of the European railways. First of all, by means of Directive 49/2004/EC, the following points are discussed:

- the setting up, in each Member State, of an authority responsible for supervising safety. This authority must be independent from RUs and IM, applicants for certificates and procurement entities;
- the mutual recognition of safety certificates delivered in the Member States;
- the establishment of Common Safety Indicators (CSIs) in order to assess that the system complies with the Common Safety Targets (CSTs) and facilitate the monitoring of railway safety performance;
- the definition of common rules for safety investigations.

In particular, one of the more interesting innovations is the introduction of a safety certification for RU in order to be granted access to the railway infrastructure. This safety certificate may cover the whole railway network of a Member State or only a defined part thereof. The validity period of the safety certification is not unlimited but it has to be renewed at intervals not exceeding five years. In addition to safety requirements, licensed RU must comply with national requirements, compatible with European law and applied in a non-discriminatory manner, relating to health, safety and social conditions, including legal provisions relating to driving time, and the rights of workers and consumers. Another important safety requirement is related to training and certification of the staff, in particular of train drivers. The training covers operating rules, the signalling system, the knowledge of routes and emergency procedures.

Moreover, the Second Railway Package, given the difficulties encountered by Member States in formulating common solutions for safety and rail interoperability, has introduced the European Railway Agency (ERA) whose main objectives are:

- increase the safety of the European railway system;
- improve the level of interoperability of the European railway system;
- contribute towards establishing a European certification system of vehicle maintenance workshops;
- contribute towards setting up a uniform training and recognition system for train drivers.

In particular, the Agency is responsible for creating and updating the Technical Specifications for Interoperability (TSIs) which ensures the development of the interoperability among the different Member States.

Finally, by means of Directive 2004/51/EC, the Second Railway Package completes the process of opening the rail market in the case of freight service. Therefore, RUs have been granted access on equitable conditions to the entire rail network of the Member States for the purpose of operating all types of rail freight services since 1 January 2007.

2.1.6 The Third Railway Package

The Third Railway Package deals with the development of Community Railways (Directive 58/2007), the certification of train drivers operating locomotives and trains (Directive 59/2007) and the establishment of new rights for rail passengers (Regulation 1371/2007).

The main target of Directive 58/2007 is to support the liberalisation of the international passenger rail market introducing new rules for the allocation of railway infrastructure capacity and the levying of charges for the use of railway infrastructure. As a consequence, RUs have to be granted the right of access to the infrastructure in all the Member States for the purpose of operating an international passenger service.

The Directive 59/2007 has introduced a certification system for locomotive and train drivers on the European Union rail network excluding drivers belonging to metros, tram and other light rail systems. In particular, all train drivers must have the necessary fitness and qualifications to drive trains and hold the following documents:

- a licence identifying the driver and the authority issuing the certificate and stating the duration of its validity. The licence will be the property of the driver and will be issued, on application, to drivers meeting the minimum requirements as regards medical and psychological fitness, basic education and general professional skills;
- a harmonised complementary certificate as evidence that the holder has received additional training under the railway undertaking's safety management system. The certificate should state the specific requirements of the authorised service (rolling stock and infrastructure) for each driver and its validity will therefore be restricted.

Finally, Regulation 1371/2007 strengthens passengers' rights regarding transport contract, information, tickets and service quality.

2.1.7 A bottleneck in the process of opening the market: the regulation 1370/2007

Before introducing the recent proposals for the definition of the Fourth Railway Package, it is worth analysing the contents of the Regulation 1370/2007 which has limited, in a certain way, the process of opening the rail market.

In particular, the regulation concerns the awarding process and the general rules for the definition of public service contracts. The competent authority (i.e. any public authority or group of public authorities in one or more Member

States which can intervene in public passenger transport in a given geographical area) is obliged to conclude a public service contract with the operator to which it grants an exclusive right and/or compensation in exchange for discharging public service obligations. Regarding passenger transport services by rail, the duration of the public service contract cannot exceed 15 years.

However, there is no obligation to instigate competitive procedures for the awarding process of the public service contract. In fact, public authorities may provide public transport services themselves or assign them to an internal operator over which they have control comparable to that over their own services. This means that, the tendering process is not mandatory and for this reason, new rail operators do not manage to enter the market which is controlled by state-owned RU. As can be seen in the following paragraph, one of the main proposals of the Fourth Railway Package is related to the modification of the regulation 1370/2007.

2.1.8 The Proposals for the Fourth Railway Package

The Fourth Railway Package proposals were announced by the European Commission at the end of January 2013 and can be grouped into three main pillars:

- to create better structures and governance for infrastructure managers;
- to open domestic passenger markets;
- to establish consistent approvals procedures for rail interoperability and safety.

Regarding the first task, as explained in previous directives, there was a requirement to have a degree of functional separation between infrastructure and service operation activities. However, the European Commission observed that the requirement often resulted in discrimination between incumbents and new entrants. To solve this problem, the Commission proposes a much greater distinction between IM and service operators, either through complete institutional separation (namely two completely separated groups) or in a vertical integrated company with a holding structure. The aim is to guarantee the necessary legal, financial and operational separation and ensure that the liberalisation of the rail sector is not hindered by discriminatory behaviours of the IMs.

The second target of the proposals concerns the opening of domestic passenger services. Indeed, by December 2019, the Commission wants European railway companies to be granted the access to infrastructure to provide all services (included the domestic passenger service) in all the Member States. Until now, this access was granted only for freight operations and international passenger services, although international passenger operators were allowed to transport passengers between stations within a single Member State as long as they were on the international route (the so-called ‘cabotage’). However, the opening of the domestic passenger markets would be subjected to the provision that the access granted must not compromise the ‘economic equilibrium’ of a public-service contract. In addition, contrary to regulation 1370/07, the Commission suggests that public-service railway contracts must be subjected to mandatory tendering. This will apply to all new public service contracts from December 2019 and all existing public service contracts from the end of 2022. However, the obligation will be linked to a value threshold, below which there can be direct awards if the cost of a competitive tender would exceed the expected savings of public funds.

The last point concerns the interoperability and safety of the European railways. To achieve rail liberalisation, it is necessary to have interoperability between railway infrastructure, rolling stock and signalling systems of different Member States. The European Railway Agency (ERA) was established to facilitate this target, producing common technical standards and safety indicators and targets. However, its role has largely been one of making recommendations. The new proposal is to enhance ERA’s role. As new powers, the agency will issue vehicle authorisations for placing on the market and safety certification for railway companies; it will also control national safety authorities and supervise national rules. In this way, it is believed to reduce by 20% the cost and length of the rolling-stock authorisation procedure and the time needed for new companies to enter the market.

2.1.9 European Legislation on service quality.

Previous paragraphs showed the process of opening the market, promoting liberalisation and improving the efficiency, effectiveness and quality of the railways. In particular, the adoption of public service contracts and tendering situations give public authorities the possibility to specify the various criteria which regulate the public transport service. The European Committee for Standardization (CEN) introduced recommendations and contents of agreement regarding quality with the document EN13816, which establishes the guidelines for allocation of responsibilities between authorities and competitors of the tender. In this way, the bidder knows precisely the requirements in terms of level of quality he is obliged to satisfy during the concession period. Furthermore, according to EN13816, the tender document has to include viable, manageable and measurable quality parameters. However, the railway system (as the other public transport systems) is extremely complex since quality has different characteristics depending on the vision of the subject considered. Therefore, the EN13816 specifies different quality perspectives and above all, the interactions among them. In particular, quality can be viewed in four different manners, namely:

- Service quality sought, which is the quality required by customers. It can be expressed as the number of weighted quality criteria assessed by qualitative analysis;
- Service quality targeted. It is the quality that the service providers decide to achieve. Generally, it is influenced by the customers' level of quality sought, external or internal pressures, budgetary and technical constraints as well as competitors' performance.
- Service quality delivered, which is the level of quality actually achieved by service operators. It is not just a technical evaluation of the service

(e.g. punctuality indicators) but it is measured from passengers' viewpoint.

- Service quality perceived, namely the user's perception of the quality delivered which depends on his/her personal experience of the service in question.

Interactions among these quality standpoints are extremely important for the purposes of describing the level of service quality. The difference between quality sought and quality targeted, for instance, gives indications to service providers about how they have to direct their efforts to satisfy customer purposes. The difference between quality targeted and quality delivered is an indicator of the capacity of service operators to achieve their targets. The gap between delivered and perceived quality is, by contrast, a measure of the customer's degree of experience with the service and his/her knowledge of the quality offered by the service provider. Finally, the difference between quality sought and quality perceived is a measure of customer satisfaction. All these interactions can be easily represented by the 'quality loop' shown in Figure 2.1

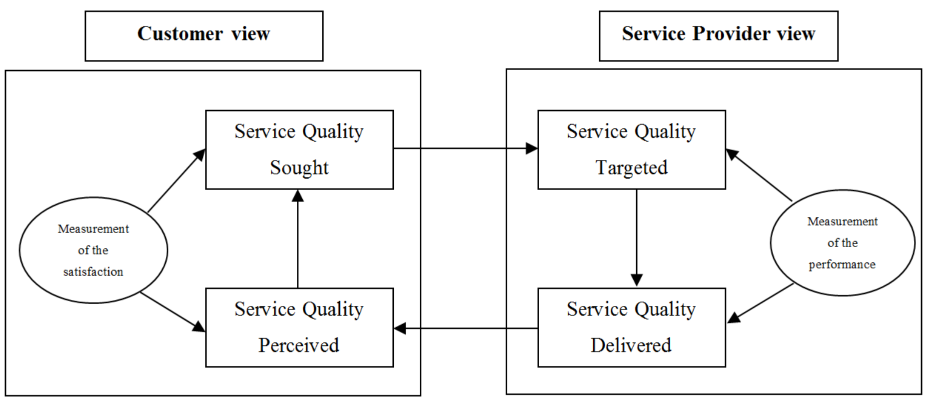


Figure 2.1 Representation of the quality loop

The EN13816 proposes to adopt these principles in the management of the service quality providing several indicators which represent customers' view of the service offered. In order to 1give an idea, some of these indicators with their explanation are reported in the following:

- Availability: the extent of the service offered in terms of geography, time, frequency and transport mode;
- Accessibility: access to the Public Passenger Transport (PPT) system including interface with other transport modes;
- Information: systematic provision of knowledge about a PPT system to assist the planning and execution of journeys;
- Customer care: service elements introduced to effect the closest practicable match between the standard service and the requirements of any individual customer;
- Comfort: service elements introduced for the purpose of making PPT journeys relaxing and pleasurable;
- Security: sense of personal protection experienced by customers derived from the actual measures implemented and from activity designed to ensure that customers are aware of those measures.

The application in real contexts of the quality cycle demonstrated its positive effects for the improvement of the attractiveness of the public transportation service. In Cascetta et al. (2013b), for instance, on the basis of EN13816, it was showed that even the aesthetic quality of the stations plays a role on user perception and mobility choices. In particular, the authors proposed a methodology for planning public transport services taking into account service providers, users' reaction, demand flows, planning activities, system monitoring and integration of standardised quality indicators. The application on a new metro line in Campania region (Italy) highlighted the increase of customer satisfaction and attractiveness of the service due to the beauty of the new stations, the new rolling stock, the new information system as well as the new integrated fare structure.

Another example of the application of the quality cycle is presented in Barabino et al. (2013). The authors proposed their methodology for applying the EN13816 quality cycle to any public transport system. In order to show its effectiveness, the method was then tested on a Light Rail Transport service so as to evaluate targeted, delivered, sought and perceived quality and their related gaps.

2.1.10 An overview of the Italian Legislation.

As part of the Member States, Italy adopted all indications provided by the European Union. In particular, by means of national law 'D.L. 422/97', the Italian government introduced the federalism in the management of the public transportation systems so as to pursue the principles of efficiency, effectiveness and service quality. Moreover, the law proposed the mechanism of competition for the market for the awarding process of public transport services within each region of the country. This means that tendering situations should be mandatory and the service must be regulated by a service contract so that all the requirements in terms of rolling stock, service typology, frequency, quality level and time of the concession period could be established precisely. However, due to a series of postponements and to the European Regulation 1370/07 (see paragraph 2.1.7), in the majority of the cases tendering situations have not been performed yet and regional train services are completely controlled by the old state incumbent.

As regards the service quality, Italian government anticipated the EN13816 legislation with the Ministerial Decree of 30 December 1998, which stipulates the general scheme for drawing up the '*Mobility Charter*' in which the delivered and perceived quality have to be explained through indexes measuring safety and service availability, cleanliness of trains and stations, degree of crowding and so on. In particular, much importance is attached to the activity of monitoring which ensures respect of the contract agreements. For this reason, during the period of concession, public authorities have to ascertain that the service quality is consistent with what is stated in the Mobility Charter.

As far as the adoption of the Railway Package is concerned, several national laws largely changed the structure of the railway system. National Railways have been reorganised since 2000: the infrastructure manager RFI (*Rete Ferroviaria Italiana*) has a completely separated management from the

national rail operator (*Trenitalia*), although both belong to the same group, whose holding firm is ‘*Ferrovie dello Stato*’.
The Rail Freight sector was the first to be completely liberalised in accordance with the European Legislation. Nowadays, several rail operators coming even from other countries (the Second Rail Package has indeed allowed the cabotage in the freight service) compete in the market while the national incumbent (i.e. ‘Trenitalia Cargo’) wants to quit the service considering it non-profitable.
Regarding High Speed Lines, Italy has been the first country to completely open the market: two rail operators, Trenitalia and NTV (Nuovi Treni Veloci) compete to offer high speed service after liberalisation in 2011. As a consequence, service quality largely increased and above all, ticket prices are very low compared to those of the other countries in Europe.

2.2 Railway infrastructure.

Railway infrastructure is composed of several elements such as tracks, stations, signalling system and, on electrified lines, the catenary or third rail system with power supply.
Tracks are the roadways of the rail system and they are also what mainly distinguishes this mean of transport from the other non-fixed guidance systems. A track is composed of the rails, the ties or sleepers, the tie plates between rail and ties and the ballast. Figure 2.2 shows the principal elements of an electrified double track line.
The sub-structure is equally as important as the track since it ensures a safe and comfortable ride for the trains. It consists mainly of three parts: the formation, the sub-ballast and the ballast. The formation is the ground upon which the track is laid. Like roadways, it can be the natural ground level or it can be an embankment or cutting. It is extremely important that this part of the formation is made of the right material and properly compacted so as to carry the loads of passing trains.
The upper surface of the formation has a little slope to enable a sufficient drainage of the track bed. The track is supported on ballast which is usually

made up of ‘granite’ stones. Further, this material has to be rough in shape in order to improve the locking of stones. Ballast is provided to give support, load transfer and drainage to the track. It is separated from the formation by a layer of sand which is laid over some sort of geotechnical screen to prevent water seepage to the lower part.

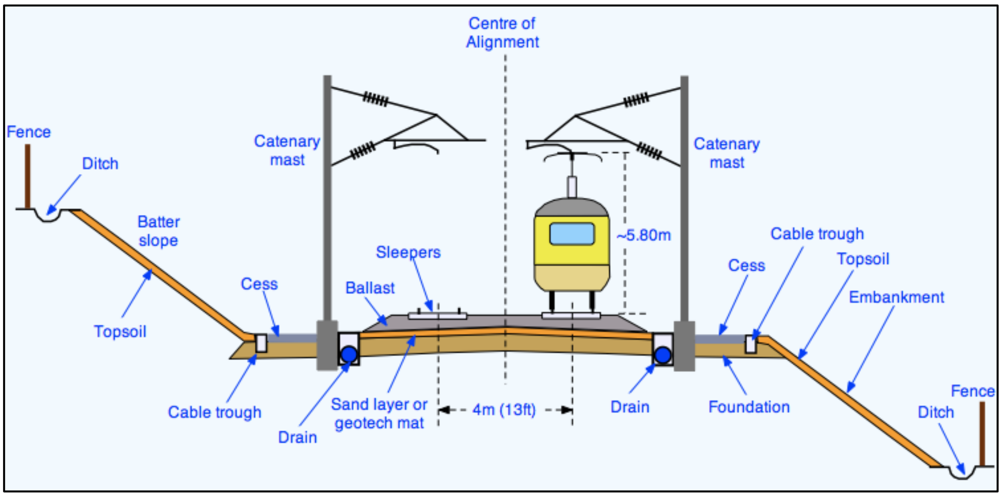


Figure 2.2 Main elements of the railway infrastructure

Some modern track forms, called slab track or non-ballast track, have a concrete base and are used in particular locations such as tunnels or bridges where a rigid structure is required (Figure 2.3).

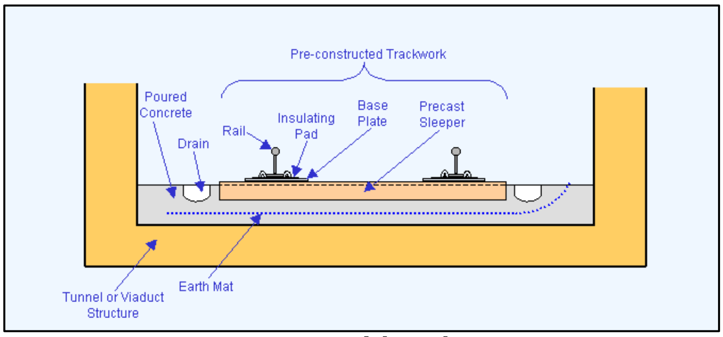


Figure 2.3 Slab track.

The earth mat is a steel mesh which is vital to try to keep stray return currents from connecting to utilities like pipes or other steel devices. In some cases, the sleepers lay on rubber pad so as to reduce vibrations (Hong Kong Mass Transit System is an example).

The usual form of a track consists of two steel rails, secured on ties (or sleepers) which keep the rail at the correct distance apart and give the possibility to support the weight of trains. The sleepers can be wooden (just old installations), concrete (most used) and steel (just in the case of light used tracks). Usually they are installed keeping a gap of 0.60 metres (Esveld, 2001).

The track gauge is the distance between the inner sides of the rail heads of a track. The most common track gauge is 1435 mm, known also as '*standard gauge*' since it is used on about 2/3 of all railway lines in the world. However, there are several kinds of track gauges used around the world (see Table 2.1) which results in great problems of interoperability especially within the European Union.

Broad gauge (India/BART):	1676 mm
Broad gauge (Spain):	1674 mm
Broad gauge (Portugal):	1665 mm
Broad gauge (Ireland):	1600 mm
Broad gauge (Finland):	1524 mm
Broad gauge (former USSR):	1520 mm
Standard gauge:	1435 mm
Narrow gauge (Cape gauge):	1067 mm
Narrow gauge (meter gauge):	1000 mm
Narrow gauge (US narrow):	914 mm

Table 2.1 List of gauges around the world

The standard form of a rail is the so-called 'flat bottom' (Figure 2.4). It has a wide base (called foot) and a narrower top (or head), whose dimensions largely vary from country to country. The rail normally rests on a cast steel plate which is screwed or bolted to the sleepers. The rail is then attached to the plate by a system of clips or clamps.

Figure 2.4 Standard form of the rail.

Rails are produced in fixed lengths and need to be joined to make a continuous surface on which trains can run. The traditional method was to bolt them by means of perforated steel plates (often called fishplates) which produce the jointed tracks.

Although the small gaps left between rails are useful to allow thermic expansion of the rails, when a train passes over a jointed track, it makes a particular noise and above all, produces vibrations and forces which are not desirable for high speed trains. For this reason, nowadays, continuous welded rails are constructed. In this case, the rails are welded together into long lengths which can be up to several kilometres. This solution provides a better ride, reduce wear, decrease damage to trains and eliminate the noise associated with joints. Expansion is minimised by installing and securing the rails in tension. However, special joints are provided in particular places to allow for temperature changes (Figure 2.5).

Figure 2.5 Expansion joints in welded rails

As can be seen, additional rails are bolted in the centre of the track to prevent the sleepers from being shifted by rail expansion.

On curve sections, careful calculations are required so as to avoid that trains can derail. In particular, the outer rail must be at a higher level than the inner rail. This is known as cant or super-elevation which is used to compensate for lateral forces generated by the train as it passes through the curve. However, the level of cant has to be limited to a maximum value (e.g. for the Italian Infrastructure Manager the maximum value is 0,16 metres) since on the line there are trains travelling the curve at different speeds but there can be also the occasion that trains stop on the curve and are subjected to an excess of centripetal acceleration. Indeed, it is worth dealing with the stability of the vehicle on the curve which gives the possibility to calculate the speed limit corresponding to a certain curve radius ρ. In particular, Figure 2.6 shows all forces acting on a vehicle running on a curved rail section with a gauge s and super-elevation α. The horizontal equilibrium between the centrifugal force F_c and weight force P can be calculated as:

$$F_c^{'} = \frac{m \cdot v^2 \cdot \cos\alpha}{\rho} \tag{2.1}$$

$$P_h = m \cdot g \cdot \sin\alpha \tag{2.2}$$

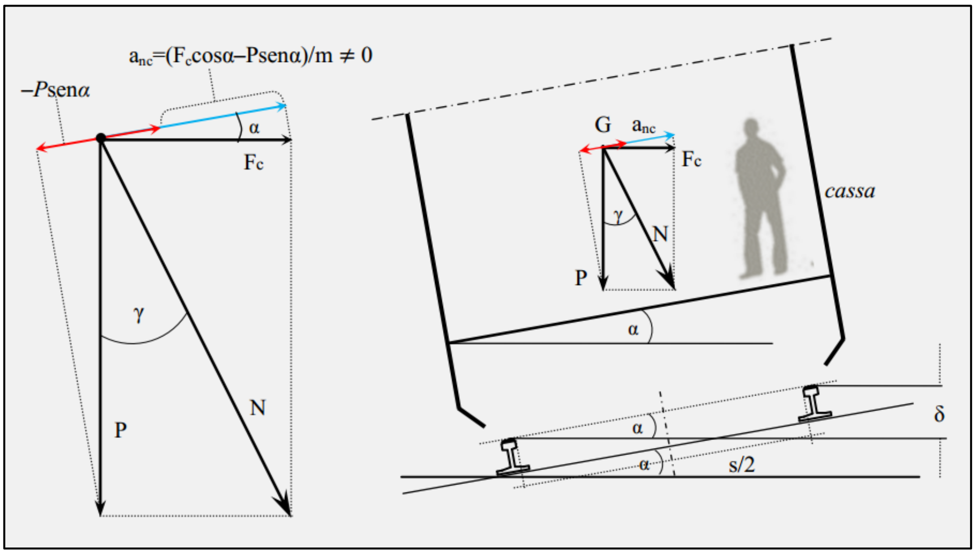


Figure 2.6 Forces acting on a vehicle while running on a curve.

Where m [kg] is the mass of the vehicle, g [m/s^2] is the gravitational acceleration, v [m/s] the velocity of the vehicle and ρ is the radius of the curve. The non-compensated centrifugal force F_{nc}, both the vehicle and passengers are subjected to, is therefore expressed by the following equation:

$$F_{nc} = F_c^{'} - P_h = \frac{m \cdot v^2 \cdot \cos\alpha}{\rho} - m \cdot g \cdot \sin\alpha \tag{2.3}$$

Since α is very small, $\cos\alpha \approx 1$ and the non-compensated acceleration a_{nc} becomes:

$$a_{nc} = F_{nc} / m = \frac{v^2}{\rho} - g \cdot \frac{h}{s} \quad (2.4)$$

In order to guarantee a good comfort level for passengers, it is necessary to limit this acceleration to values included within the interval 0,6 and 1,0 m/s^2.
From equation (2.4), the maximum allowable velocity $v_{\lim}$ corresponding to a curve with radius ρ, can be easily evaluated:

$$v_{\lim} = \sqrt{a_{nc} + g \cdot h / s} \cdot \sqrt{\rho} \quad (2.5)$$

Obviously, this speed largely varies according to the train category and the value of the non-compensated acceleration considered. For example, in the case of metropolitan system, it is possible to tolerate higher values of a_{nc}, due to the short travels.
As regards operational purposes, tracks can be divided into main tracks and sidings. The formers are the tracks used for train movements while the latters are all the other tracks used for shunting movements. Main tracks used for passing and overtaking trains are called loops.
A railway turnout (or point) is a mechanical device consisting of an assembly of rails, movable points and frog, which allows a train to be driven from one track to another one.
The points can be operated manually or by point machine. In case of a small angle, movable frog (known as '*swing nose frog'*) are provided so as to eliminate the common gap which causes noise and waste. Moreover, to prevent trains from unattended movements, particular devices called '*derails*' are installed on converging tracks.
A crossing, by contrast, is an assembly of rails which allows two tracks to cross at grade. Crossings with a large angle are constructed rigidly with double frogs (see Figure 2.7) while in case of small angle intersection, a crossing can have movable points instead of frogs.

Figure 2.7 Rigid frog system (left) and movable frog system (right).

Other important infrastructure elements are junctions, crossovers, ladders and wyes. A junction is the point where a line is joined to another one by means of turnouts. A crossover is instead an arrangement of corresponding turnouts which provides connection between two parallel tracks. A ladder is a particular track with a series of turnouts allowing the access to any of several parallel tracks. Finally, a wye is system composed of three or two turnouts and one high angle crossing forming a triangle of tracks. Figure 2.8, shows all the above mentioned elements.

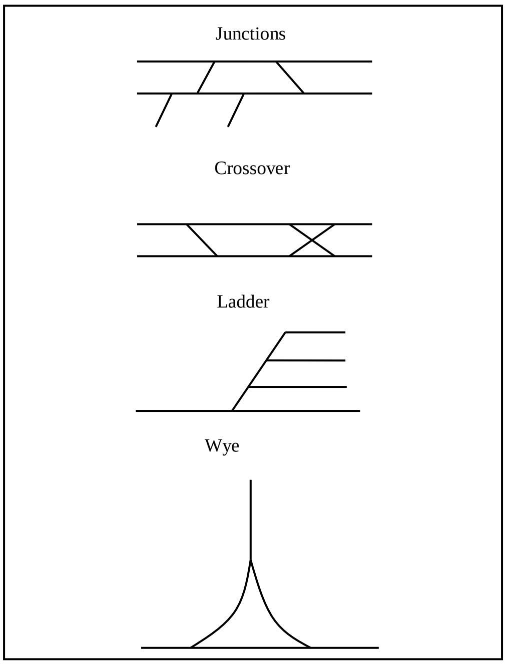


Figure 2.8 Junctions, Crossover, Ladders and Wyes (source: Pachl, 2009).

Another fundamental element of the rail infrastructure is the station, which can be described as the place where train-stopping operations are performed (loading/unloading passengers or goods, on-board staff exchange, etc.). Since the station is the first point of contact between passengers and the railway system, it should be regarded as the shop window of the service provided. Further, especially in the last decades, great importance was given to station design inasmuch as it enhances the attractiveness of the rail service (Cascetta et al., 2013b).

Generally stations can be classified in terminal stations, ordinary stations and stop stations according to their layout, the number of tracks and platforms and their position within the network.

Terminal stations are located at the end of a railway line. For this reason, trains arriving to this kind of stations have to end their journeys and reverse out to start a new run. Usually, the layout of the station allows passengers to reach every platform without crossing tracks. Moreover, in some cases, the railway lines continue for a short distance until a depot where trains stand, are assembled or maintained.

The ordinary stations, by contrast, differ from the previous ones since they are not located at the end of a railway line. However, they can be composed of main tracks and loops as well as several platforms to permit stopping operations. Sometimes, even sidings are provided to store or assemble rolling stock.

Stop stations are simply stops along the railway line to let passengers board/alight the train. Usually, these stations do not have any track rather than the main tracks and therefore, do not need junctions or interlocking systems.

2.3 Rolling Stock and dynamics of train movements.

Rolling stock can be divided into the following categories (Esveld, 2001):

- Passenger and freight stock;
- Hauled and powered stock;

- Electric and diesel stock (in particular the latter can be diesel-electric or diesel-hydraulic).

Diesel locomotives are self-sufficient units that combine a prime mover, traction motors, fuel tank, and operator controls to pull or push passenger cars over routes without additional wayside infrastructure to supply power, except for fuel filling stations within yards.

Electric locomotives do not carry an internal prime mover and instead rely on energy supplied by an off-car electrified traction power supply and distribution system. For a given horsepower, an electric locomotive, is considerably lighter than a diesel locomotive. Further, it is possible to achieve much higher overall horsepower in a similarly sized electric locomotive than in a diesel locomotive. The net effect is a benefit of higher overall train acceleration, speed and system capacity.

To better understand the dynamics of train movements, it is necessary to give first a brief description of the wheel/rail interface.

The basic unit of a rail vehicle is the wheelset. It consists of two wheels fixed on a common axle so as to keep a constant distance between each other and to rotate with a common angular velocity. Flanges are also provided, which, contrary to common belief, should not touch the rails but they are safety features to prevent wheels from derailing. Rail wheels sit on the rails without guidance except for the shape of the tyre in relation to the rail head. The wheel tyre is coned while the rail head is slightly curved and is set at an inward angle.

Figure 2.9 shows the wheel/rail interface on straight track where, as said before, the flanges do not normally touch the rails.

On curved track, the outer wheel has a greater distance to travel than the inner wheel. In this case, the wheelset moves sideways so that the larger tyre radius on the inner edge of the wheel (near the flange) is used on the outer rail of the curve (see Figure 2.10).

The inner wheel, by contrast, uses the outer edge which has a smaller diameter in order to reduce the travelled distance.

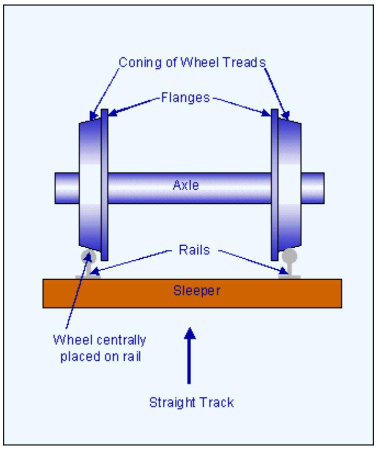


Figure 2.9 Wheel/rail interface on a straight track.

During the passage on the curve, the flange of the outer wheel could touch the rail if the movement of the train is not in exact symmetry with the movement of the track. Obviously this event causes wear and increases resistances.

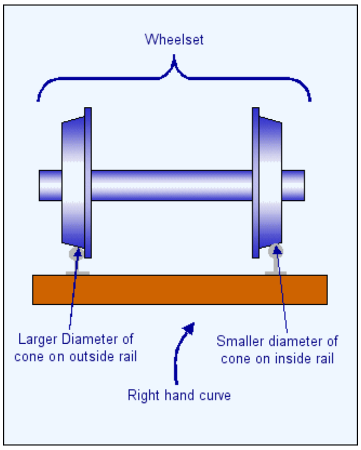


Figure 2.10 Wheel/rail interface on a right hand curve.

Two wheel-sets are mounted in a bogie which usually has rigid frames as shown in Figure 2.11. For this reason, when approaching a curve, it is required a lot of force to allow the change of direction. To overcome some of the mechanical problems of the rigid wheelset mounted in a rigid bogie frame, some modern typologies allow a form of radial movement in the wheel-set as shown in Figure 2.12.

In this case, the force wearing the tyres and the flanges are reduced as well as the stress on the bogie frame.

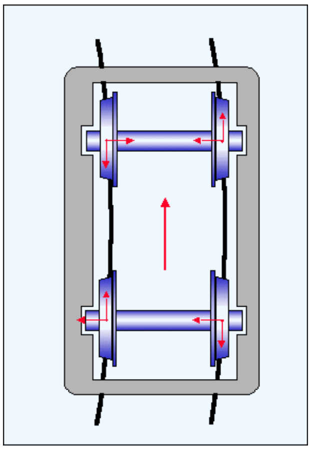


Figure 2.11 Two wheel-sets included in a bogie.

The movement of a train on a given route can be analysed looking at four components (Hansen and Pachl, 2008):

- tractive effort F_T;
- traction unit resistance F_{Rt};
- rail vehicle or wagon resistance F_{Rw};
- line resistances F_{Rl};

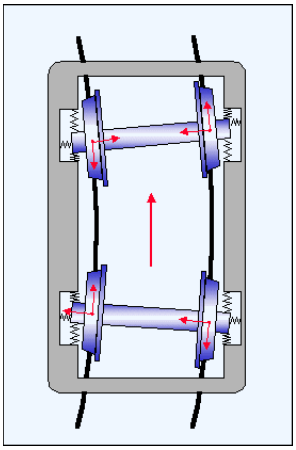


Figure 2.12 Example of modern bogie allowing radial movements in the wheel-set

2.3.1 Tractive effort F_T

To move the train, the locomotive or the power equipment of the multiple unit generates an effort which is called induced tractive effort F_T. Not the whole amount of this effort can be used due to the following reasons:

- there are some losses caused by the internal power transmission which consumes between 2% and 3% of the effort;
- the effort has to be limited to a maximum value to prevent the power equipment from overheating;
- the wheels will spin if the effort exceeds the maximum adhesion between rail and wheels.

The latter is described by the adhesion coefficient μ and the wheel load F_L on the driven wheels. In particular, the adhesion value μ can be calculated by the Curtius and Kniffer formula (Curtius and Kniffler, 1950):

$$\mu(v) = 7.5 / (3.6v + 44) + 0.161 \tag{2.6}$$

The tractive effort at wheel rim F_{Tr} is described by a curve similar to the one shown in Figure 2.13 as example.

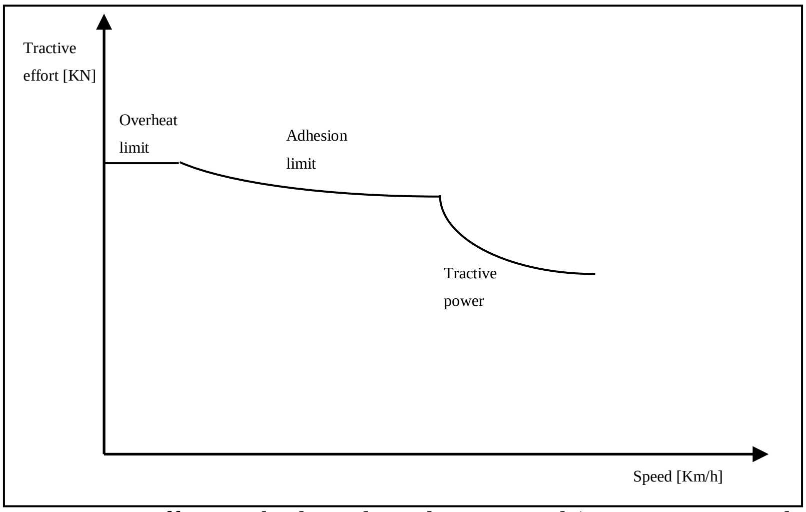


Figure 2.13 Tractive effort at wheel rim dependent on speed (source: Hansen and Pachl, 2008)

To use this diagram in running time estimations, the characteristic curve can be approximated by means of some hyperbolic or parabolic formulas, each of which is defined for a determined speed interval v_k to v_{k+1}:

$$F_{Tr}(v) = c_{0,k} + c_{1,k} \cdot v + c_{2,k} \cdot v^2, \; v_k < v < v_{k+1} \tag{2.7}$$

$$F_{Tr}(v) = c_{h,k} / v, \; v_k < v < v_{k+1} \tag{2.8}$$

The coefficients $c_{0,k}$, $c_{1,k}$, $c_{2,k}$ and $c_{h,k}$, as well as the limits v_k are the input parameters to be known for the running time estimation.

2.3.2 Vehicle resistances.

Some of the power of the engine is consumed by the locomotive and the wagons because of air resistance, rolling resistance caused by wheel rims, axle boxes and adhesion. Part of these resistances is constant while the other part has a linear dependency on the velocity v.

Generally, these resistances are described by parabolas whose coefficients r_i are function of the train characteristics and the wind speed:

$$F_R(v) = r_0 + r_1 \cdot v + r_2 \cdot v^2 \tag{2.9}$$

On the basis of this formula, railways companies have calculated lists of approximated formulae to evaluate the resistances of their own types of train. Indeed, the traction unit resistance F_{Rt}, given the parameters a_0, a_1, a_2 or a_{2r}, has the following analytical formulation:

$$F_{Rt}(v) = g \cdot m_T \cdot (a_0 + a_1 \cdot v) + a_2 \cdot v^2 + a_{2r} \cdot v^2 \tag{2.10}$$

where:

- m_T [kg] is the mass of the traction;
- v [m/s] is the speed of the vehicle;
- v_r [m/s] is the relative speed between air and vehicle, usually assumed as $v + 4.17$ m/s.

The vehicle resistance for passenger vehicle train can be also described according to formula (2.9), considering the following parameters:

- the mass of the traction unit m_w [kg];
- a factor c_b concerning the number of axles which can be assume as 0.0025 for vehicles with 4 axles, 0.004 for those with 3 axles and 0.007

for those with 2 axles;

- the number of vehicles n_w;
- a value A_f [m^2] which represents the cross-sectional area of the vehicles weighted with their aerodynamic behaviour (normally assumed as 1.45).

By means of the Sauthoff formula, the relationship between the passenger vehicle resistance and the speed has been experimentally determined (see Sauthoff, 1932 for more details):

$$F_{Rwp}(v) = 1000 m_w \cdot g \cdot (1.9 + c_b \cdot 3.6v) + 0.0471 \cdot (n_w + 2.7) \cdot A_f \cdot (3.6 v_r)^2 \quad (2.11)$$

For freight train, the Strahl formula (Strahl, 1913) is instead used:

$$F_{Rwf}(v) = 1000 m_w \cdot g \cdot \left(c_a + (0.007 + c_m) \cdot (3.6v)^2 / 100\right) \quad (2.12)$$

where:

- m_w [kg] is the mass of the wagons;
- c_a is the coefficient for axles adhesion which is equal to 1.4 in the case of roller bearings and 2.0 for older plain-bearing axle-boxes;
- c_m stands for air resistance depending on the kind of wagons. It can be assumed 0.05 for mixed trains, 0.032 for full train loads of coal or ore, 0.04 for closed wagons and 0.1 for empty open wagons.

2.3.3 Line resistances F_{Rl}.

Concerning the line resistances, the main influence is the gradient of the line which can be expressed as the following:

$$F = m \cdot g \cdot \sin\alpha \quad (2.13)$$

Indeed, since the gradients of railways are very slight, the sinσ can be approximate with tanσ which is measured as *n* in per thousand [‰].

Considering the complete mass m of the train (locomotives plus all the coaches) in kg, the gradient line resistance F_{Rlg} becomes:

$$F_{Rlg} = g \cdot 1000m \cdot n \tag{2.14}$$

Line resistances are also caused by sharp curves and can be described as:

$$F_{Rlc} = g \cdot 1000m \cdot 700 / r \tag{2.15}$$

Where r [m] is the radius of the curve. It is worth nothing that this kind of resistances can be neglected in the case of curves with a radius higher than 700 metres.

Finally, the influence of air resistance as a function of cross section and speed in tunnels has to be considered, especially when trains meet each other. However, there is no formula of general acceptance that can be used for the running time estimation.

2.3.4 Running time estimation.

The basic equation of dynamics (i.e. Newton's formula) is the basis of the calculation of train motion:

$$F = m \cdot a \tag{2.16}$$

Where F [N] is the tractive effort of the engine, m [kg] is the mass of the train and a [m/s^2] is the acceleration of the train.

In order for a train to accelerate, the tractive effort at wheel rim $F_{Tr}(v)$ must exceed the sum of all the above mentioned resistances, that is $R = F_{Rt}(v) + F_{Rw}(v) + F_{Rlg} + F_{Rlc}$. Indeed, to calculate it, the following differential equation has to be solved:

$$F_{Tr}(v) - R = f_p \cdot m \cdot \frac{dv}{dt} \tag{2.17}$$

where f_p is a mass factor, which takes into account the consumption of the effort due to rotating masses. For each part composing the train, this factor has to be evaluated. As regards the traction unit, the value f_{pT} is given with the engine data (usually ≈1.09) while for passenger vehicles and freight wagons f_{pW} ≈1.06. For the whole train, it comes out to:

$$f_p = \left(f_{pT} \cdot m_T + f_{pW} \cdot m_W \right) / \left(m_T + m_W \right) \tag{2.18}$$

Equation (2.17) is not solvable by means of an analytical form but it is necessary to adopt numerical methods. The Euler's method (Butcher, 1987) is one of these approaches which works by calculating the change in a variable from a given starting point. In particular, it estimates each functional value using the preceding functional value (start value at each step), the preceding derivative of the function and a fixed time step:

$$v(t) = v(t - \Delta t) + \Delta t \cdot \frac{dv}{dt}(t - \Delta t); \; v(t_0) = v_0 \tag{2.19}$$

As can be seen by Figure 2.14, the method consists in approximating the real curve with a broken line. Indeed, the smaller is the time step, the smaller is the error committed.

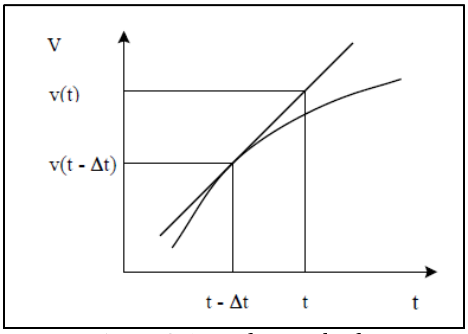


Figure 2.14 Euler method.

2.4 Train Spacing and Signalling system

In street traffic, the separation of vehicles is governed by the relative breaking distance. This means that, if one vehicle breaks, the following one will notice the brake backlight and will start breaking too. Therefore, the two vehicles can be separated by a distance which is equal to the difference of the braking distances of the vehicles plus an additional distance which depends on the reaction time of the driver of the following car.

In a steel wheel on steel rail system, the coefficient of adhesion is on average eight times less than that in highway traffic (Hansen and Pachl, 2008). As a consequence, the braking force that can be transmitted from a train to the track is also eight times less than the one transmitted by highway vehicles to the street surface. For this reason, train separation by the sight of the driver is only possible in restricted area where speeds are very low (usually no more than 30 km/h). Generally, this is possible for shunting movements and non-regular movements. For regular train movements, by contrast, train separation procedures are required which work independently from the range of view of the driver.
Basically, there are three theoretical principles of train separation (Pachl, 2009):

- Train separation in relative braking distance;
- Train separation in absolute braking distance;
- Train separation in fixed block distance.

2.4.1 Train separation in relative braking distance

According to train separation in relative breaking distance principles, the braking distance between two following trains is equal to the difference between the breaking distances of the trains plus an additional safety distance.

$$d = d_{brake,2} - d_{brake,1} + S \tag{2.20}$$

Where d is the distance between trains; $d_{brake,1}$ is the braking distance of train 1; $d_{brake,2}$ is the breaking distance of train 2; S is the safety distance.

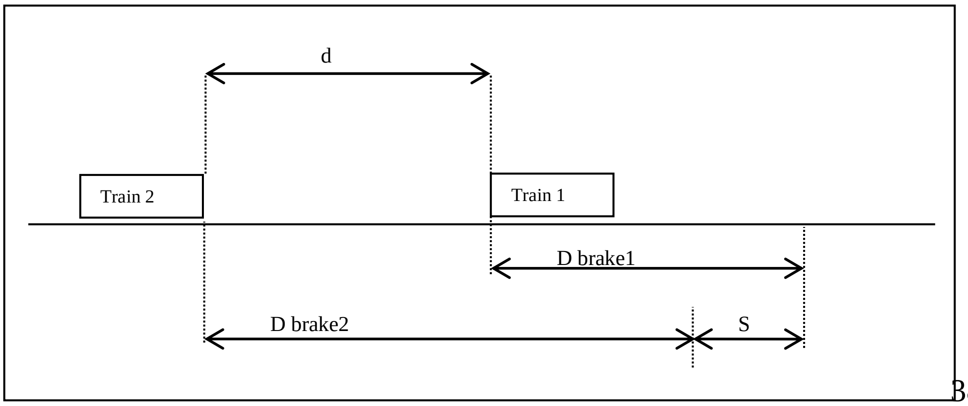


Figure 2.15 Train separation in a relative braking distance (source: Pachl, 2009)

This kind of approach is just theoretical and cannot be realistically adopted in railway operations because of some essential problems. First of all, in case of an accident of the first train, the second train has no possibility to stop and it will collide with the first train. Furthermore, when running through interlockings, it is not possible to move points between two trains. When points are to be moved between two trains, the second train has to have full braking distance to the points until the points are locked in the new position.

2.4.2 Train separation in absolute braking distance

Spacing trains in absolute braking distance means that the distance between two trains is equal to the braking distance of the second train plus a safety distance:

$$d = d_{brake,2} + S \qquad (2.21)$$

Where d is the distance between trains; $d_{brake,2}$ is the breaking distance of train 2; S is the safety distance.

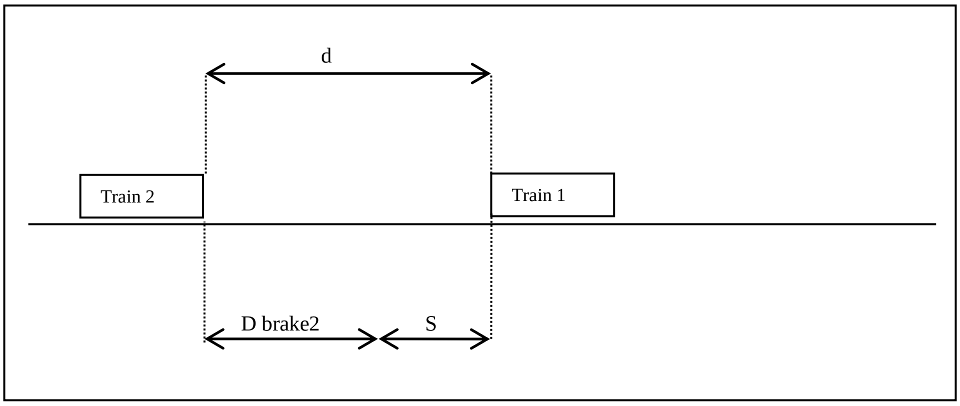


Figure 2.16 Train separation in absolute braking distance (source: Pachl, 2009)

This kind of train separation, which is also known as '*moving block*', is going to be implemented in real rail service in the future. In fact, until now, there are still researches on safe technologies for the end train location.

2.4.3 Train separation in fixed block distance.

In the case of train separation in fixed block distances, the track is divided into block sections. A block can be occupied exclusively by one train. Indeed, the distance between two following trains is equal to the maximum breaking distance plus the length of the block section plus an additional safety distance:

$$d = d_{break,max} + l_{block} + S \tag{2.22}$$

Where d is the distance between trains; $d_{brake,max}$ is the maximum breaking distance; l_{block} is the length of the block section; S is the safety distance.

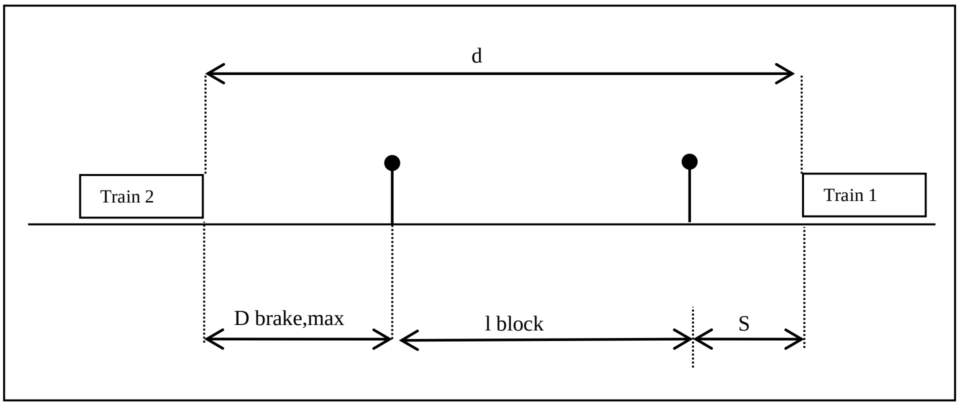


Figure 2.17 Train separation in fixed block distance (source: Pachl, 2009)

In rail lines with lineside signals, the block sections are limited by block signals. However, thanks to the development of high speed lines, there is an increasing use of cab signalling systems, since lineside signals cannot be watched safely.

This is the most common principle of train separation worldwide and for this reason, it is worth examining in depth signalled block operations.

2.4.4 Signalled fixed block operation.

As described in the previous paragraph, a fixed block system is composed by fixed block sections which are protected by signals (lineside or cab signals).

To clear the signal and so to give a train the permission to enter a block section, the following conditions must have been fulfilled:

- the train ahead must have cleared the block section;
- the train ahead must have cleared the overlap behind the next signal (on lines where the overlaps are used);
- the train ahead must be protected from following train movements by a stop signal;
- the train is protected against opposite movement.

Figure 2.18 and 2.19 show the difference between lines without or with block overlaps. In case block overlaps are not required, the control length of a signal is equal to the block section. In the other system by contrast, the control length of the signal is longer than the block section so as to provide additional safety in case the driver fails to brake before a stop signal. The difference is called '*overlap*' since in that area the control length of a signal overlaps with the control length of the next signal.

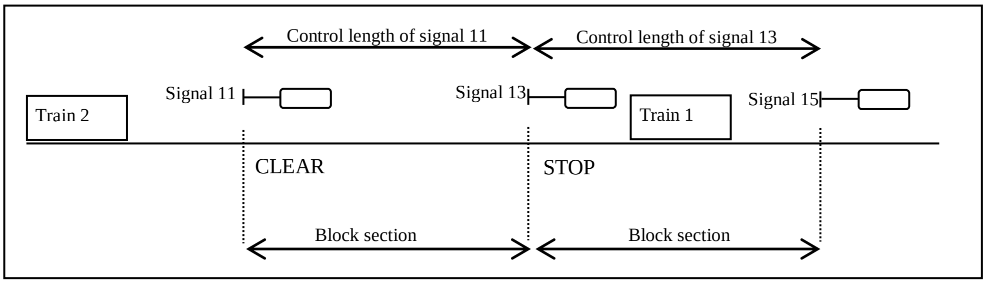


Figure 2.18 Control length of signals in fixed block territory without block overlaps (source: Pachl, 2009)

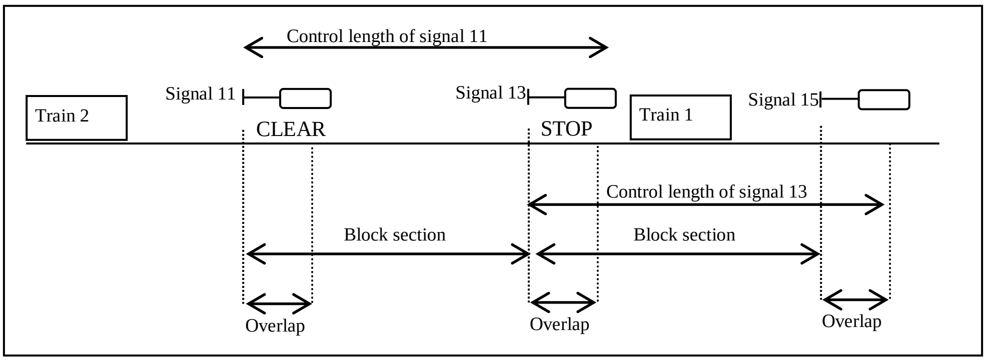


Figure 2.19 Control length of signals in fixed block territory without block overlaps (source: Pachl, 2009)

Therefore, a signal may not be cleared until the full control length is clear. This means that the clearing point behind a signal corresponds to the end of the control length of the signal in rear.

The basic rules of fixed block operations are very important since they are necessary to understand the blocking time theory which influence the headway between two trains.

2.4.5 The blocking time model.

The '*blocking time*' is the time interval in which a block section is allocated just to a train and therefore, it is blocked for other trains. This time strongly influences the minimum headway, namely the minimum time interval between two following trains. In fact, it lasts from the moment a train receives the permission to enter a block section (by clearing a signal) to the moment it is possible to issue a movement authority to another train to enter the same block section. Usually, the blocking time of a track element is much longer than the time the train actually occupies the track element. In a line with lineside signals, the blocking time can be divided into the following intervals:

- The time for clearing the signal;
- The '*signal watching time*', that is the time the driver needed to view the clear aspect at the signal which gives the approach indication to the signal at the entrance of the block section (this can be the block signal in rear or a separate distant signal)
- The '*approach time*' between the signal that provides the approach indication and the main signal at the entrance of the block section;
- The time between the block signal;
- The '*clearing time*', i.e. the time to clear the block section and, in case it is necessary, the eventual overlap with the full length of the train;
- The '*release time*' to unlock the block system.

In a territory with cab signalling, the intervals are quite similar but the approach time is now the time the train runs through the braking distance which is signalled by the cab signal system. Linking the blocking times of all block sections which are passed by a train into a time-over-distance diagram provides the so-called '*blocking time stairway*' (Figure 2.21).

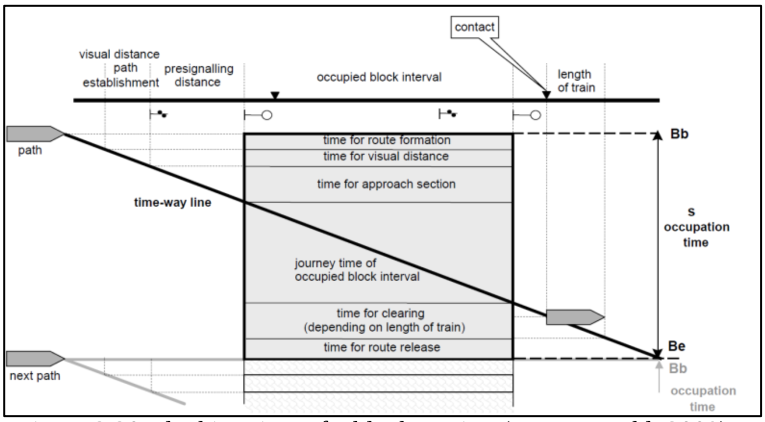


Figure 2.20 Blocking time of a block section (source: Pachl, 2009)

This diagram is quite important because it represents the operational use of a line by a train and it gives the possibility to determine the minimum headway between two trains.

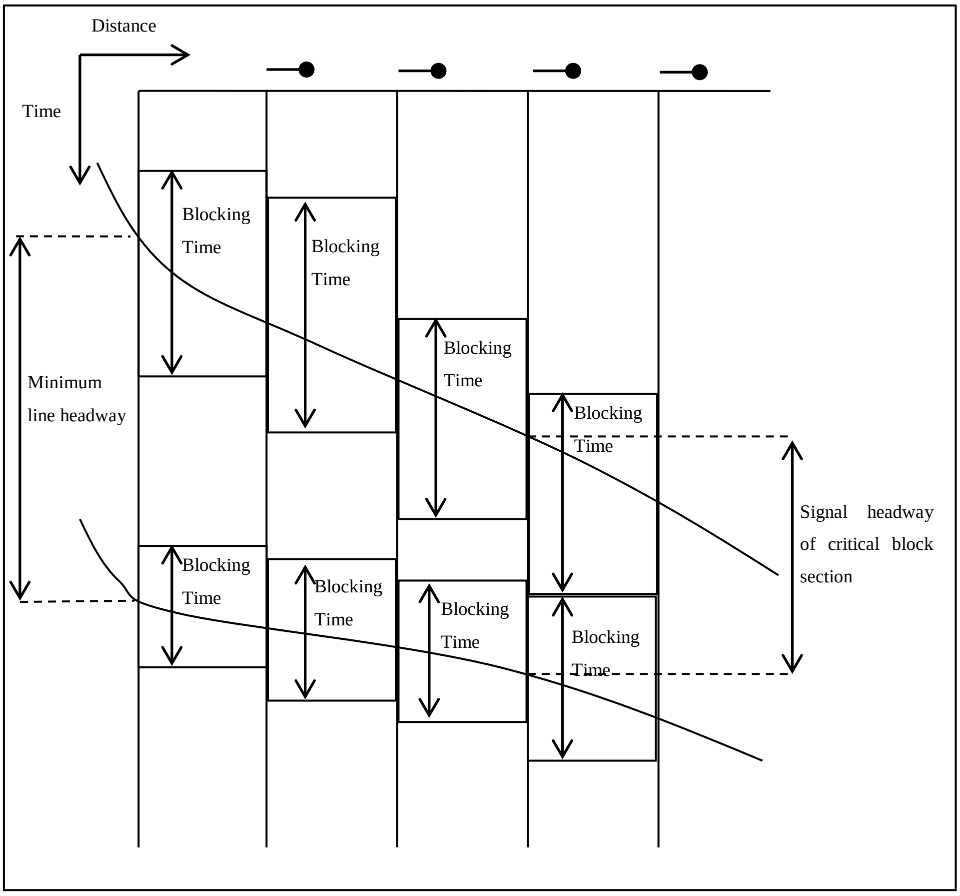


Figure 2.21 Blocking time stairways and Signal Headway for a simple railway line (source: Pachl, 2009)

In fact, while the blocking times directly defines the '*signal headway*' (i.e. the minimum time lag between two consecutive trains considering only one block section), the blocking time stairways define the '*line headway*' which is the minimum headway considering not only one block section but the whole blocking time stairways of the line. In this way the blocking time stairways of two following trains touch each other in at least one block section, whose name is the '*critical block section*'.

Hansen and Pachl (2008) showed how to calculate the minimum headway of two trains. To this purpose, the train paths have to be put one over the other with the same departure time which results in overlapping blocking times in the block sections. Indeed, for a certain block section, the blocking time overlap represents the amount of time the train path has to be postponed so as to eliminate the blocking time conflict in this block section (see Figure 2.22). By calculating all blocking time overlaps for all block sections, it is possible to know the maximum value a train has to be postponed to eliminate all conflicts between the blocking time stairways. This time represents the minimum line headway between the two analysed trains. Indeed, in analytical term, the minimum line headway $t_{h,ij}$ for train *j* following train *i* can be calculated as:

$$t_{h,ij} = max\{t_{be,1(k)} - t_{bb,2(k)}\} \; for \; k = 1....n_b \tag{2.23}$$

where $t_{be,1(k)}$ is the end of the blocking time of train I in block section k, $t_{bb,i(k)}$ is the begin of blocking time of train I in block section *k* and n_b is the number of block sections.

However, the time distance between two consecutive trains is larger than the minimum line headway determined by formula (2.23). In fact, this procedure is deterministic and does not take into account any kind of delay. Therefore, this value of time is generally increased by a certain buffer time to compensate for small delays and to provide a more robust timetable.

On lines operated with a moving block system, it is still possible to determine the blocking time diagram which is transformed in a continuous time channel.

(Figure 2.23).

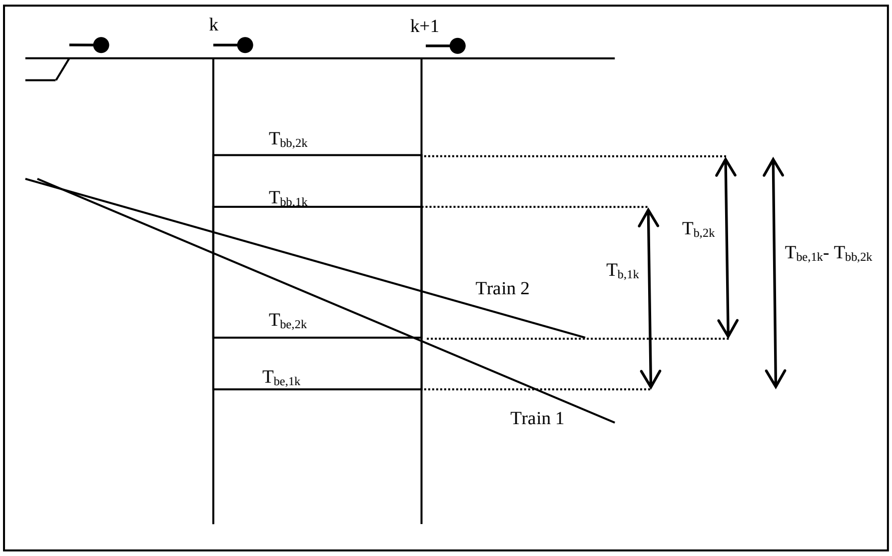


Figure 2.22 Principle of calculating the blocking time overlapof a single block section (source: Hansen and Pachl, 2008)

In fact, in this case, all components of the blocking time can be reproduced except from the running time between the block signals, since the length of the block sections is reduced to zero.

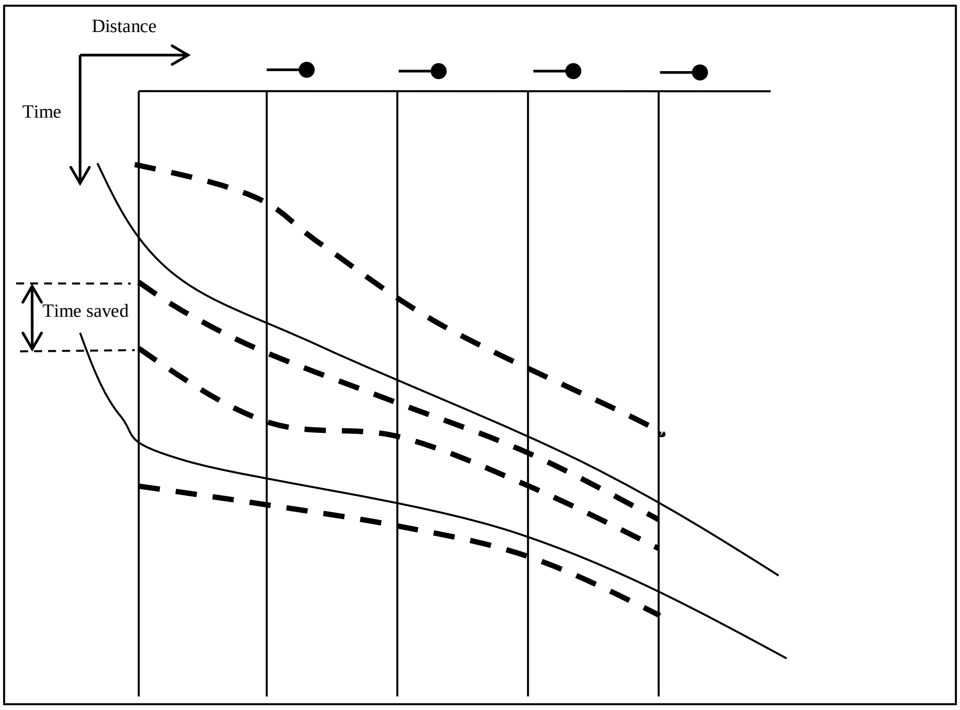


Figure 2.23 Blocking time of Moving Block (source: Pachl, 2009)

2.4.6 Principles of signalling systems.

There are several kinds of signalling systems spread all over the world and this has always constituted one of the main problems of interoperability of the rail systems. This paragraph wants to give an overview of signalling systems without focusing on a particular type.

First of all, as regards the classification of the signal aspects, it is possible to divide the signalling systems into two groups:

- Speed signalling
- Route signalling

The former indicates the speed which cannot be exceeded by a train while the latter provides information about the route over which the train is being sent. In route signalling territory, the driver must know the speed limit of the route the train has to run over. This kind of system can be largely found on British and American railways (see Pachl, 2009), while most modern systems follow the speed signalling principle.

Generally, the speed or route information is integrated with the block signal aspects through the combination of lights. Anyway, some modern systems in Europe have supplementary indicators and the block signal itself gives indication about the occupation of the following block section. In this case, the block signal aspects are generally three and divided as follows:

- ‘red’, stop;
- ‘yellow’, approach (i.e. it is necessary to stop at next signal);
- ‘green’, clear.

Regarding the way to provide the approach indication, it is possible to distinguish between two kinds of signalling:

- One-block signalling;
- Multiple block signalling.

In one-block signalling system, the block signal gives just information about the block section protected by it and therefore, there are no approach indications for the next signal. For this reason, every block signal must have a distant signal which gives the required approach information and it is placed at the braking distance before the block signal. In case of short block sections, the distant signal is placed at the rear block signal mounting them one above the other on the same pole.

In a multiple block signalling the block signal informs the driver about the status of two or more following block sections. One of the most common is the two block signalling in which the approach information is provided by the aspect of the rear signal without using any distant signal. This system is also called '*three aspects signalling*' since a block signal can show three different aspects. However, in combination with a progressive speed signalling, this system can have more than three aspects. On lines with short block sections for instance, a train approaching a stop signal can be progressively slowed down by speed indications. This progressive speed signalling system, although requires more than three signal indications (mostly four or even more), gives approach indication just for the next block signal.

Some railways use a three-block signalling in which a block signal informs about the status of three block sections ahead. One example is the British system, where the following four aspects are provided:

- red – stop
- yellow – approach (caution)
- double yellow – advanced approach (preliminary caution)
- green – clear.

The '*advanced approach*' prepares the driver to stop at the second signal. Anyway, this four aspects signalling can be useful just on lines where block sections are not much longer than the braking distance. In fact, in case of longer block sections, this kind of system would reduce the capacity of the line increasing the signal headway of the following train.

2.4.7 Block systems.

Block systems consist of signalling appliances or operating procedures which ensure a safe train separation on lines provided by fixed block sections. Basically, block systems can be divided into manual and automatic block systems.

Manual block systems need a local operator who is responsible to check the clearance of the block sections before giving a train the permission to enter it. More in detail, the operators have to check the train integrity by watching the rear end train markers, then operate the block signal manually and transmit the block information by telecommunications (usually by telephone, that is why this system is called telephone block). Furthermore, all train movements and train messages have to be recorded in a hand-written train record with the following information:

a) at departure station
   - the train description,
   - the time the train has been accepted by the receiving station,
   - the departure time,
   - the time the clearance message station has been received from the next block station.

b) at an intermediate block station
   - the train description,
   - the time the train has departed from the departure station,
   - the time the train has passed the block station, the time the clearance message has been sent to the block station in rear,
   - the time the clearance message station has been received from the next block station.

c) at the receiving station
   - the train description,
   - the time the train has departed from the departure station,
   - the arrival time,

- the time the clearance message has been sent to the block station in rear.

To reduce the probability of human errors during telephone communication, some railways use also special lamps or movable signs to indicate the state of the line. Other systems removed completely the risk of accidents introducing more sophisticated technologies. In a controlled manual blocked system for instance, the signals are still operated manually but controlled by continuous track circuits which require the cooperation of both operators of two adjacent block stations. Indeed, a signal cannot be cleared if the block section is still occupied or when one operator has opened the signal for an opposite movement. Another advantage of this system is the fact that it is not necessary to check the integrity of the train by watching the rear end train markers.

The interlocked manual block system is another example of block system more common in European railway lines (Pachl, 2009) where block sections are interlocked by means of a block apparatus. In particular, when a train has entered a block section the signal is blocked in a stop position and the operator cannot clear it until he has received (through electric line) the clearance information from the operator of the next signal. Furthermore, after the signal has been reset to a stop position, it is automatically locked by a rotation locked device so as to secure it in case the electric block instrument fails. The rotation lock can be only released after the block instruments has worked properly, that is the train has electrically ‘blocked in’. After the train has cleared the block section and it is protected by a stop signal, the operator can ‘block out’ the train by a release button that will unlock the signal at the block station in rear. This button is also electrically connected to a short track circuit which prevents from clearing the signal unless the train has passed it. However, it is still necessary to check the train integrity by watching the rear end train markers.

Automatic block systems instead, check the clearance of the block sections by means of track clear detection devices and therefore, the signals work automatically. For this reason, there is no need to have local operators checking the train integrity. To better explain these kinds of systems, it is necessary to introduce the block logics to protect train movement in the same direction. In particular, the two principles are the following:

- closed block
- open block

In a block system following the closed block logic, in normal position the line remains always in a block state. It is only cleared when a train has to enter the block section. In an open block system by contrast, the section is only blocked when it is occupied by a train. As soon as the train has left the section, it is cleared and it remains in this state which is the normal position.

In many automatic block systems, the operations are based on an open block principle, although some railways use approach-controlled automatic block signals which follow the closed block logic.

Basically, track clear detection systems are composed of track circuits or axle counters.

A track circuit is an electrical circuit a rail section is a part of. It has a source of current at one end and a detection device at the other. Sections are usually divided by insulated rail joints. When a section is occupied by a train, the axles produce a short circuit between the two rails and the device at the beginning of the section does not receive any current detecting the section as occupied. The detection device is usually constituted of a track relay which is in a 'picked up' position when the section is clear and dropped when the section is occupied (Figure 2.24).

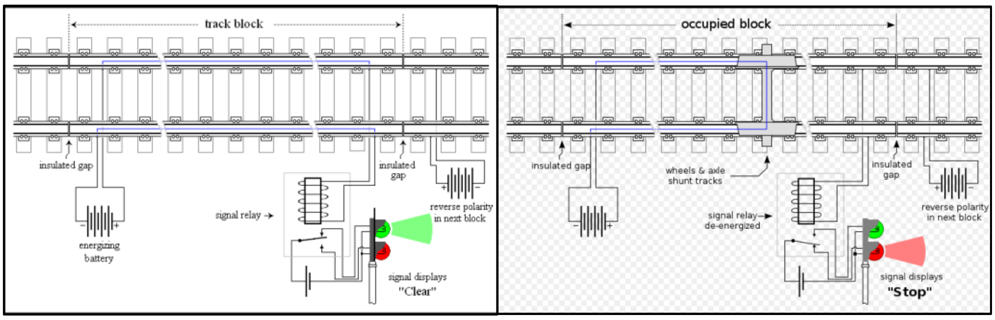


Figure 2.24 Track circuits

As regards the characteristics of relays, it is possible to distinguish between DC (Direct Current) and AC (Alternate Current) track circuits. In particular, the latter have polyphase relays which work following the same principles of a polyphase motor. In fact, these relays must be fed by two phases of the same frequency, one coming from the track circuit and the other from a local source. If a wrong frequency is received from the track, the relay will not move and the track circuit cannot be wrongly cleared by foreign currents of different frequencies.

Some railways adopt jointless track circuits which work with an audio frequency AC track current. The working length of such track circuit would limit itself due to the inductive and capacitive track characteristics and therefore, for a safe continuous track clear detection, it is necessary that track circuits overlap each other. A fixed limit is required though. For this reason, at the boundary between two track sections, S-shaped rail connectors shortcut the rails allowing the adjacent track to overlap exactly by the length of the connectors

Jointless track circuits can be used just for short block section and can be found within city railway lines.

Other railways adopt automatic block system based on coded track circuits. In this case, the track current is overlaid by a code track which contains signal information. As a result, it is possible to:

- provide information for cab signalling
- improve the safety of the track circuit
- transmit signal information from the signal at the exit of the block section to the signal at the entrance of the block section.

In a cab signal territory, the train receives continuously information about the indication of the next signal or about the state of the next block section. Coding the track circuit allows reaching this target. In fact, the code can be received by trainborne devices and transformed to cab signal indication which are displayed at the driver’s desk.

Coded track circuits are also important for improving the safety since a track section is only detected when a correct code is transmitted. In this way, wrong currents would never clear the track circuit.
Finally, coded track circuits provide information about the state of the next signal without the need of additional communication lines for the transmission of signal indications along the track.

An axle counter, by contrast, is a system composed of counting points at both ends of a section and a counter connected to them. The train occupancy of a section is detected by comparing the number of axles entering the section with the number of axles leaving it. Obviously, to clear the signal at the entrance of the section, the parity of numbers is necessary. Moreover, counting points usually consist of double contacts to detect the direction of movement.

According to the automatic block system logic, a signal must only be cleared when the entire control length of the signal is also clear and a train ahead is protected by a stop signal. However, this system could not be safe enough in case a signal fails in automatic resetting to stop position after the passage of a train. For this reason, some railways add more protections to increase the safety of the system. For example, one solution is the adoption of additional rail contacts which will reset a signal to stop position independently from the track circuits. Another possibility is to make two adjacent signals work jointly. Thus, a signal will be only cleared if at the same moment the next signal is in stop position. In case of a fault, one signal would protect two block sections. A more sophisticated automatic block logic is the one following a principle very similar to the interlocked manual block system. In particular, after the train has entered the block section the signal is set to stop and locked in stop position. To unlock it, the train must both have cleared the section and passed the next signal.

In addition to the previous principles related to protection of train movements in the same direction, it is important to consider two kinds of block working:

- Absolute block working;

- Permissive block working

In an absolute block system, a train can never enter a block section which is still occupied by another train. In case of failure of the block system, if a train has to enter a block section showing a stop signal, the driver needs a special moving authority from the operator (written or verbal orders or call-on signals). In this case, the operations can continue following practices similar to manual block without any technical protection and the clearance of a block section must be checked by watching the rear end train markers.

On lines with permissive block working by contrast, a train may pass a stop signal proceeding with a low speed. This is the most typical automatic block system not only because the previous system notably decreases the capacity of a line, but also because the number of local operators to protect train movements has been largely reduced. In a territory with permissive signalling it is extremely important to reset a signal to its stop position with a very high reliability after a train has entered a block section. Otherwise it would authorise a following train that is running under permissive rules to enter the block section without limitations. For this reason, in block system where this high reliability cannot be guaranteed, trains running under a permissive rules have to ignore the first clear aspect at the next signal, entering the block section with a low speed. Only when a train passes a second signal in a clear position it can start accelerating again.

2.4.8 Overlaps in an automatic block territory.

To prevent collision in case a train is standing immediately beyond a signal, some railway lines use a ‘safe braking distance’ called overlap since this distance overlaps part of the following block section.

In automatic block lines where overlaps are used, there are three different principles to provide this protection. The first one consists of placing the signals in a distance before the track detection limit that equals the overlap.

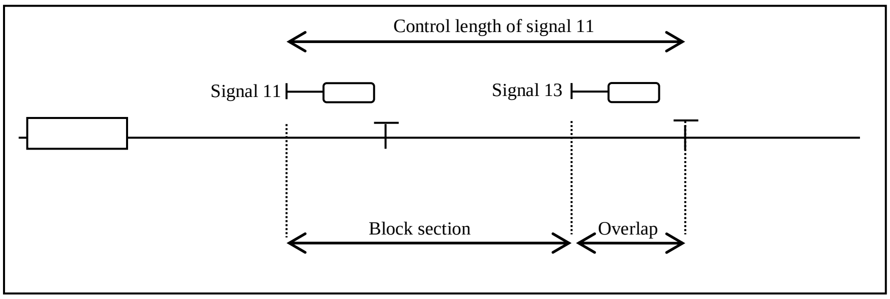


Figure 2.25 Placing the signals in a distance before the track detection limit that equals the overlap (source: Pachl, 2009).

This principle has the disadvantage that the signal are quite late to reset to stop position and if a locomotive stands directly behind the signal would not be protected by a stop indication. On lines with permissive working this situation is dangerous and could be solved by placing additional rail contacts which reset the signal to stop when the train is with the overlap distance.

The second solution is the adoption of separated track sections with an own track clear detection for the overlap.

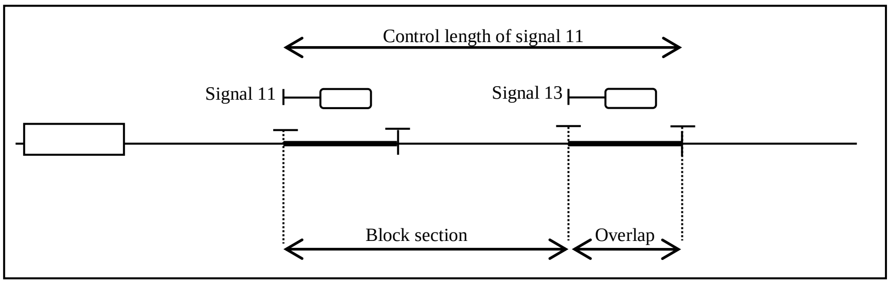


Figure 2.26 Separate track sections with an own track clear detection for the overlap (source: Pachl, 2009)

In this case the signals are reset to stop position immediately and with high reliability.

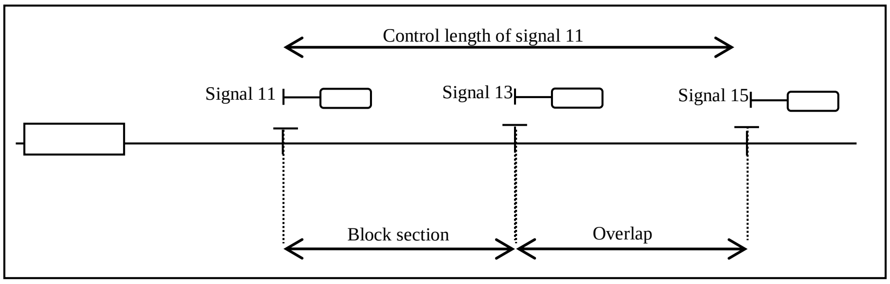


Figure 2.27 Block section used as overlap (source: Pachl, 2009)

The last principle is suitable in the case of lines with short block section such us electric city railways or undergrounds. Indeed, an entire block section could be used for the signal in rear as overlaps.

2.4.9 Train protection.

As already explained in the previous paragraph, human being is the weakest element in the railway safety. Trackside signals and interlockings are sometimes not sufficient to provide high safety levels and therefore, train protection systems are necessary to guard against driver errors or to check his attentiveness.

Before explaining the different Automatic Train Protection (ATP) systems, it is worth focussing on the functions they can provide, mainly divided in cab signalling functions, supervision functions and intervention functions (Theeg and Vlasenko, 2009).

Cab signalling functions can be classified into the following groups:

- Non-selective warning signals (mainly audible): when a train passes a particular position (a distant signal for example), a warning tone sounds to direct driver’s attention to the trackside signals, independently from their aspects. This function, which is still applied in old train protection systems, does not need information connection between trackside signal and protection system.
- Selective warning signals (mainly audible): The audible signal is applied just in case there is a restriction for the driver, for example a ‘Caution’ or a ‘Speed Restriction Warning’, which requires the start of a braking process.
- Visual repetition of trackside signals: this function provides the aspect of a trackside signal in advance in the cab while the train is running within two signals or close to a trackside signal. The cab signal is indeed visible in any weather conditions and gives information to the driver earlier than the track signal. However, it does not give more

information and the driver is still responsible for estimating the braking requirements.

- Continuous static speed information: in this function, information about the permitted speed under consideration of all restrictions is displayed together with the indications of trackside signals. For this reason, in several systems, this kind of cab signalling replaces trackside signals. However, the driver is still responsible for the estimation of the braking curve. Some modern systems provide static speed profiles for each track element instead of imposing one speed for the whole section between two track side signals.
- Dynamic speed information: based on the static speed information, braking patterns are calculated on the train and/or in the trackside equipment. The guidance speed is indeed displayed continuously and the driver must not exceed it in order to comply with the next target speed. Further, information concerning the distance to the next braking target has to be provided either for each track section or standardised by the uniform length of the sections (this case is suitable only for lines with uniform traffic).

The supervision functions are instead divided into the following groups:

- Check on driver ability: at regular intervals, the driver has to use an alertness device to guard against falling asleep or similarly, independently from the trackside signal aspect. This interval can be time- or distance-measured. Usually, this system is known as '*dead-man's handle*'. In some systems, this device needs to be handle only if the driver does not undertake any operation during a defined time interval.
- Check on driver attentiveness: in certain situations (e.g. after passing a signal restriction) the driver has to push a special button to acknowledge his attentiveness. In this way, it is more difficult that he

can fail to perceive the signal although many cases have occurred in which the driver pushed the button habitually without braking.

- Train stop function: as soon as a train passes a red signal, this function provides an immediate emergency stop. In case of permissive driving, driving on written instruction or on an auxiliary signal, special procedures have to be undertaken. For example, by using additional override handles in the driver's cab which allow passing the signal at very low speed. However, it is worth nothing that except in cases with very low speed, high braking performance or long overlap, this function is not sufficient to stop the train safely within the overlap.
- Braking supervision: since the above mentioned functions are not sufficient to stop the train before the point of conflict, modern systems adopt braking supervision. In particular, when a train has to brake for a red signal or to comply with a speed restriction, the braking process is supervised at certain points or continuously. Indeed, different methods are used among the systems and are shown in Figure 2.28.

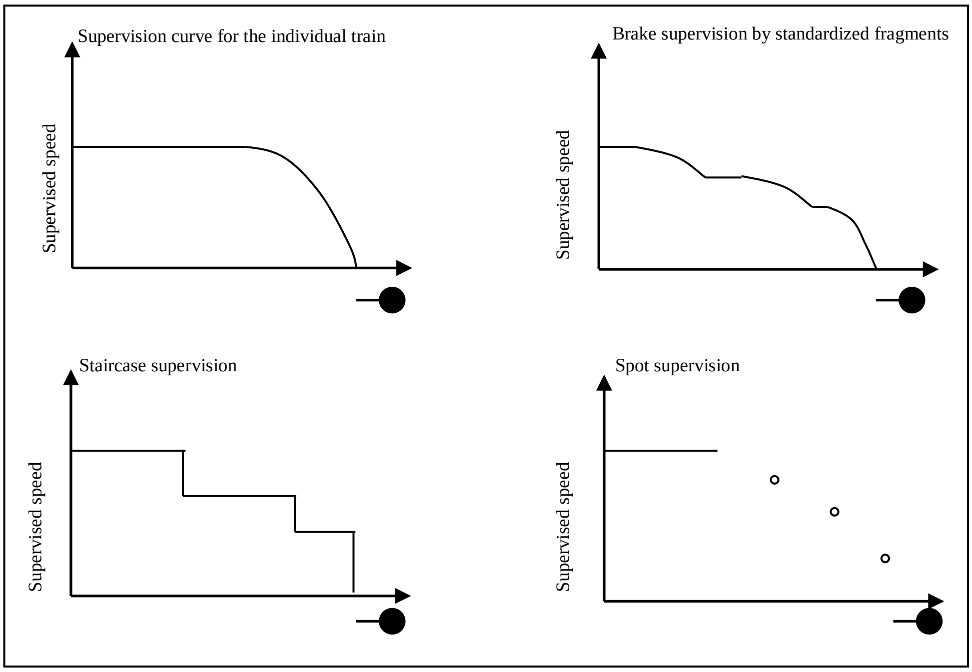


Figure 2.28 Form of brake supervision curves in train protection systems (Theeg and Vlasenko, 2009)

Supervision curve for the individual train means that the brake supervision pattern is calculated individually for the train and the track layout. Indeed, one supervision curve is used for the whole braking process before a stop signal. The second method provides a stock of standardised fragments of brake patterns depending on speed level, proximity to the stop signal and/or train category.

The staircase supervision is the one usually applied in the case of coded track circuits. Indeed, the supervision function has the shape of a staircase and the same data input is valid during the whole length of the track circuit.

In the fourth method, the speed is instead checked in form of multiple spots. As can be seen, when approaching a stop signal, the supervision speed decreases from one checkpoint to the next one.

In some systems, driver can exit manually from the braking supervision if the signal aspect in advance has been upgraded and this information cannot be transmitted automatically by the train protection system.

- Compliance with speed limit: together with the supervision of the braking process, many systems provide checking of speed restrictions. This can be for example the maximum speed of the line, a local restriction, or a limit imposed on the vehicle.

Intervention functions start working when supervision functions detect a problem in the behaviour of the vehicle. There are different levels of interventions. The weakest is to warn the driver of the problem (e.g. by an audible warning sound) so as to demand correction. The next step usually applied by some railways is to switch the traction off automatically. The following one is the service brake intervention while the strongest intervention function is the activation of the emergency brake. These measures can be applied consecutively according to different tolerance margins or it is possible to adopt just one of them (mostly the emergency brake). Anyway, after passing a red signal, the consequence in all systems, which have the supervision function, is an immediate emergency stop.

From the operational aspect, the forms of transmission adopted can be intermittent (spot and interrupted linear transmission) or continuous. Basically, in the former case transmission is only possible at selected locations where apposite trackside equipment is installed (transponders or balises). The latter by contrast, requires a continuous data link between track and train.

According to their functions and type of transmission, the ATP systems can be classified into five groups (Theeg and Vlasenko, 2009):

1. Systems with intermittent transmission and without braking supervision;
2. Systems with intermittent transmission at low data volume and with braking supervision;
3. Systems with continuous transmission of signal aspects by coded track circuits with (3a) or without (3b) braking supervision;
4. Systems with intermittent transmission at high volume data and dynamic speed supervision;
5. Systems with continuous transmission at high volume data and dynamic speed supervision.

functions / transmission	Attentiveness check, trainstop function and other without brake supervision	With brake supervision in different form, but without dynamic speed profile	Dynamic speed profile
Intermittent	Group 1	Group 2	Group 4
Continuous	(3a) Group 3	(3b)	Group 5

Figure 2.29 Classification of train protection systems (Theeg and Vlasenko, 2009)

In the following, a short description of each group is proposed.

Group 1: Systems with intermittent transmission and without braking supervision

This kind of systems provides mainly two supervision functions, namely an attentiveness check at the signals with a restricted aspect and/or a train stop function. The gain in terms of safety is really limited and is not sufficient for modern requirements. In fact, there are cases in which, even after the correct acknowledgement action by the driver, the stop signal has been violated (Kondo, 1980). Further, the train stop function without braking supervision requires overlaps which are as long as the braking distance of the train which results in a great reduction of capacity.

Examples of this system are the mechanical train-stop (Berlin S-Bahn), the French Crocodile, the British AWS and the Swiss Signum.

Group 2: Systems with intermittent transmission at low data volume and without braking supervision.

In these systems, a dynamic speed profile for the braking process is provided in addition to the attentiveness check and train stop functions. In most cases, resonant circuits are used for data transmission which can be switched effective or ineffective, or can be switched between different active statuses according to the signal aspects. The major problem of this kind of system is that the ineffective status (i.e. 'permitting status') cannot be distinguished from the absence of a trackside transmitter. This means that it is a non-fail-safe behaviour system which is not suitable for cab-signalling system and have to work in background until the train is driven correctly. Examples of this system are the German Indusi and the Japanese ATS-P.

Group 3: Systems with continuous transmission of signal aspects by coded track circuits.

The third category is characterised by the transmission to the train of the trackside signal aspect ahead through the rails. This system is applied in several European countries, also in Italy (whose BACC system is described in the following paragraph). The required track circuits are used for track clear detection and for the transmission of block information. Further, the signal aspect ahead is repeated in the cab and the supervision function can vary from simple acknowledgement checks to braking supervision with standardised fragments. The main advantage of this system respect to group 1 and 2, is that it can be designed fail safe. Therefore, malfunction of the equipment causes a more restrictive indication in the cab. Moreover, the driver continuously receives updated information in each position of the way which prevents from forgetting signal aspects and enables immediate reaction of the system in case the signals aspect changes. However, an important disadvantage is that, although the length of the track is standardised or additional transmitters for length information are provided, calculation of an adjusted breaking curve is not possible. For this reason, some systems use both track circuits and intermittent transmission systems.

The Italian BACC system for conventional and high speed traffic lines.

The BACC system is applied on Italian conventional lines and lines with an increased speed up to 200 km/h and consists of a cab signalling system with four different aspects. In particular, the track circuit current is modulated with different frequencies which correspond to different signal aspects. Since track circuits have almost equal length (i.e. 1350 metres), it is indeed possible to calculate and supervise the braking curve. Further, a stop is announced 2700 metres before the stop position. For speeds up to 200 km/h (as for example the so-called '*direttissima Roma-Firenze*' line), since it is necessary to assure a higher braking distance, the system has been upgraded with additional frequencies which increase the number of codes to 9. This has been done combining the codes of another carrier frequency (178Hz) with the previous codes based on 50 Hz frequency. The system is downward compatible which means that the high speed trains can run without problems on conventional lines and conventional trains can do the same on high speed lines.

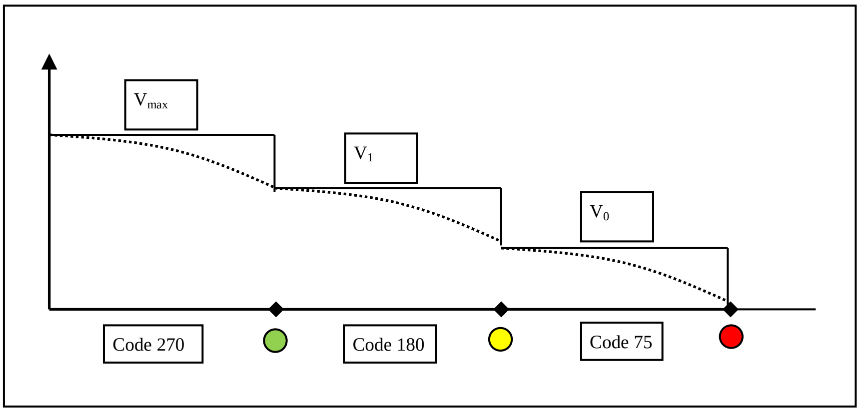


Figure 2.30 Simple scheme showing the Italian BACC signalling systems based on coded track circuits.

Group 4: Systems with intermittent transmission at high volume data and dynamic speed supervision.

Group 4 concerns the modern systems for intermittent transmission characterised by a fail-safe behaviour and the possibility to supervise the complete dynamic speed profile. According to these principles, many systems have been developed in several countries which differ in data coding, amount of detailed information and antennas. For this reason, they are incompatible with each other.

The main trackside transmission media are:

- transponder balises, which work without power supply using the energy sent from the vehicle;
- inductive loops with limited extension, usually powered from trackside;
- locally limited radio transmission devices.

Further, according to data contents, there are static or switchable transmission media. The first ones can transmit only static information such as speed and gradient, which are independent from trackside input (i.e. signal aspects). The second ones contain also information depending on the status of the trackside, especially signal aspects. However, the majority of the systems belonging to this group adopts transponder balises which store static line data and dynamically communicate to the train the aspect of the signals.

Some European examples are the Ebicab (Scandinavia, Portugal, Bulgaria), the ATB-NG (Netherlands), the TBL (Belgium), the ZUB (Switzerland, Denmark), the KVB (France) and the ETCS level 1 (international standard) which will be dealt with in detail in the following.

Group 5: Systems with continuous transmission at high volume data and dynamic speed supervision.

The main characteristic of the systems belonging to group 5 is the continuous or quasi-continuous data link between track and train. In particular, the technical transmission media applied are the following:

- Codec track circuits, whose main examples are the digital ATC in Japan or the TVM 430 which is applied on French and Belgian high speed lines;
- Cable loops, used mainly on German high speed lines;
- Radio transmission, which is the one implemented within the ETCS level 2/3 (described in detail in the following).

To allow the continuous communication, information flow is centralised in most cases, using a line-side control centre. An important criterion to distinguish the systems is whether they are used as the only signal system on the respective lines (in this case system-inherent fall-back levels are provided or driving on sight is the only one considered for degraded operation) or if they are used mixed with trackside signals with the possibility to have shorter block sections.

On lines for mixed traffic, the assignment of functions to the interlocking system or the train control system is defined as follows:

- The interlocking functions including track clear detection, which are necessary for all movements on the line, are assigned to the interlocking system;
- The cab signalling and the train protection functions are assigned to the train control system. In some cases, additional auxiliary functions for

interlocking can also be carried out in the train control system, like for example the detection of the halt of the train for route release.

After this assignment of functions, route information has to be transmitted from interlocking system to the trackside control centre. For other functions, such as sending information about the halt of the train to the interlocking system, a bidirectional data connection is necessary, otherwise a unidirectional connection suffices.

2.4.10 The European Train Control System (ETCS).

The previous paragraphs have shown that there is a large variety of train protection systems within Europe which is a great obstacle for the international interoperability of the rail system. For this reason, the European Train Control System (ETCS) has been developed since the early 1990s so as to promote a unified system for the continent. The main subjects involved in this process are the European Commission, the UIC (International Union of Railways) and the Unisig (Consortium of the seven largest European signalling manufactures). The ETCS is currently implemented on several railways networks in Europe and in some other countries outside the continent such as Taiwan, China, South Korea, Saudi Arabia, Turkey, India, Australia and Mexico (Winter et al., 2009). Problems in the introduction process are relevant though. In particular, these consist mainly in the high investments and the migration from the old national system to the ETCS requiring double equipment of lines and/or vehicles.

Basically, five different levels have been identified for ETCS systems: level 0, 1, 2, 3 and STM.

2.4.10.1 ETCS level 0.

The term '*ETCS level 0*' describes the situation where a vehicle which is equipped with ETCS moves in an unequipped area. In this case, the train is driven looking at the trackside signals and the supervision functions are limited to the supervision of a constant speed, which is equal to the minimum between the maximum train speed and a general nationally defined speed limit for level 0.

2.4.10.2 ETCS level 1.

Level 1 is a cab signalling system which can be superimposed on the existing signalling system, leaving the previous fixed signals in place. In particular, transponder balises called '*Eurobalise*' transmit movement authorities and profile data to the train which is not individually known when it passes above them. As other train protection systems belonging to group 4, balises can be fix data or switchable. In this second case, balises pick up signal aspects from trackside signals by means of Lineside Electronic Units (LEU) and transmit them to the vehicle together with movement authority and route data at fixed points. According to this information, the on-board computer continuously estimates the maximum speed and the braking curve. Because of the spot transmission of the data, a train has to pass over the balises to obtain a new movement authority. By installing additional Eurobalises or a Euroloop (i.e. cable loops in the rail) between the distant signal and the main signal, the new proceed aspect is transmitted continuously.

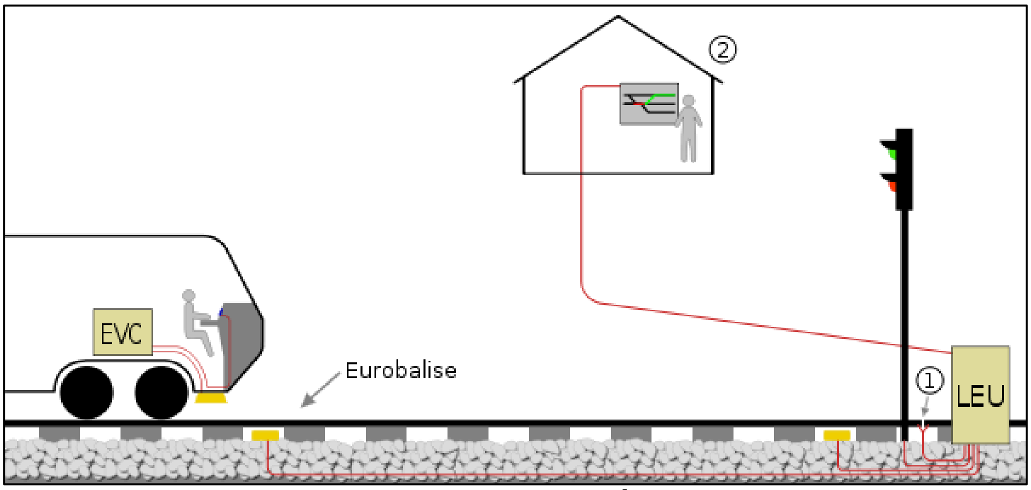


Figure 2.31 ETCS Level 1.

2.4.10.3 ETCS level 2.

ETCS level 2 is a digital radio-based system. Both movement authority and other signal aspects are displayed in the cab and therefore it would be possible to dispense with trackside signalling. However, track circuits are still necessary for the train integrity supervision while trackside signals are used during degrade operations. All trains automatically communicate their position and direction of travel to the Radio Block Center (RBC) at regular intervals. This is the central trackside unit and it is responsible for a longer section of the line, stores static data and obtains dynamic data like signal and point positions from the interlocking stations in the area. Moreover, contrary to level 1, all trains are individually known in the RBC. The movement authority is transmitted via GSM-R (European standard for radio communication) together with speed information and route data. Eurobalises are used as passive positioning beacons since the trains determines its positions via sensors (transducers, accelerometer, radar). In particular, the positioning balises are fixed reference points to correct distance measurement errors.

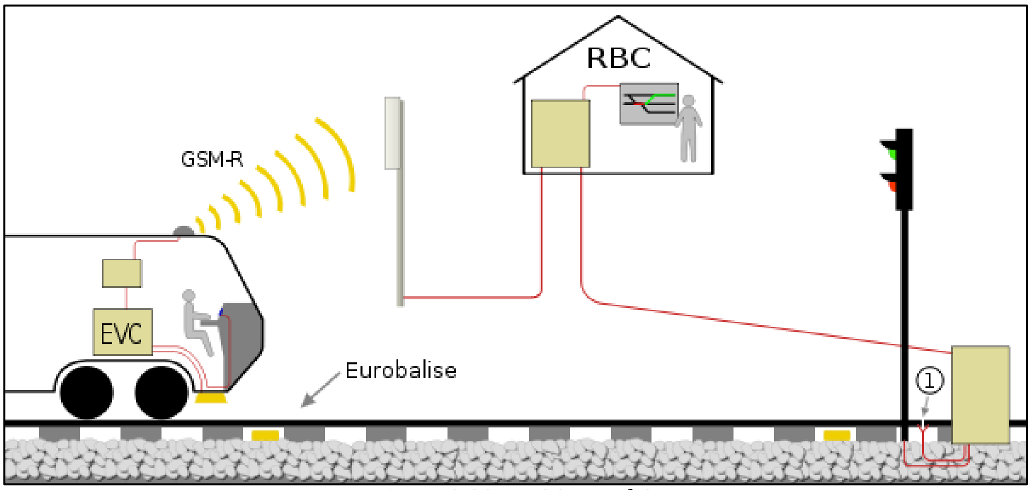


Figure 2.32 ETCS Level 2.

ETCS level 2 has been introduced on several lines, especially high speed lines (e.g. Italian high speed line).

2.4.10.4 ETCS level 3.

ETCS level 3 is similar to level 2 but provides the implementation of full radio-based train spacing and therefore, fixed trackside signalling devices are no longer required. Indeed, trains estimates their positions themselves by means of positioning balises and sensors but are also capable of determining train integrity on-board with a high level of reliability. Thus, 'Moving block' principle is applied, which means that fixed block sections are removed and train spacing is based just on the absolute braking distance. Since solutions for reliable train integrity supervision are highly complex, level 3 is still under development.

2.4.10.5 Level STM.

Level STM stands for Specific Transmission Model and it is designed for situations where a train, which is equipped with ETCS devices, moves on a line without ETCS, but with a national train protection system. This level has been introduced to facilitate the migration to ETCS systems, In particular, an additional module, the STM, is added to the on-board equipment to translate between the respective national system and ETCS.

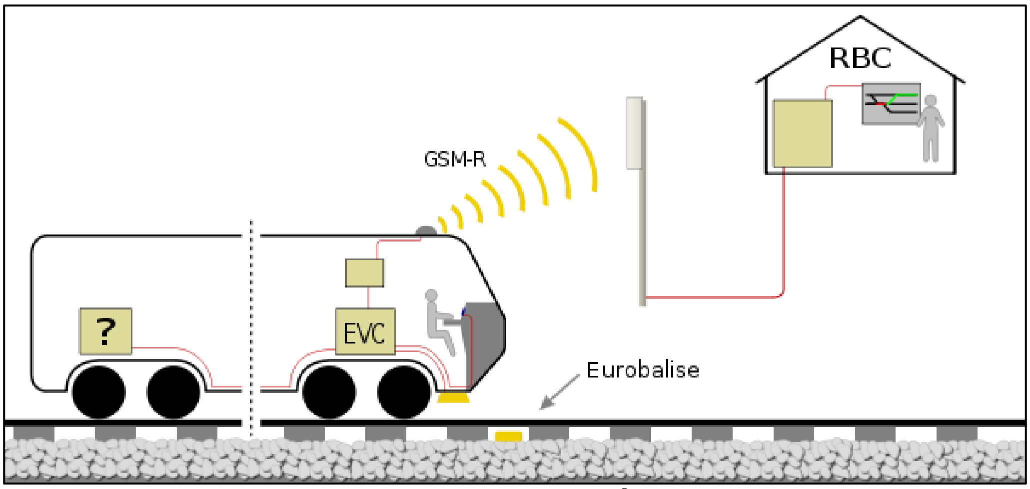


Figure 2.33 ETCS Level 3.

2.4.11 Automatic Train Operation.

The Automatic Train Protection systems can be defined as the set of devices used to help automate operations of trains. In fact, with complete dynamic speed profile present on the train, train operations could be easily automated. The main hindrance to the complete introduction of these systems is the lack of ability to react to unpredicted events such as obstacles on the track. Therefore, they can be applied only on lines where there is the completely continuous detection or protection from external objects (by means of barriers for

example). Indeed, such investment is practicable in some cases of metropolitan areas due to the limited extension of the network and the high density of traffic which makes the expense economically reasonable.
However, different steps of automated operations can be distinguished:

- Manual driving without any automation. In this case the driver is completely responsible of the train driving.
- Manual driving with technical supervision. At this level, there is a train protection system which supervises the driver and enforces safety in case of driver's errors.
- Partially automatic operation: in this case, some operations are assigned to the driver and others to automatic system. In the ATC system provided on Japanese High Speed lines, the driver is responsible for acceleration and platform stopping while the automatic system for safety related braking processes.
- Automatic driving with human supervision. This system allows the automatic train driving but the driver watches the track and intervenes in case of danger. However, it is worth underlining that this ATO is not so often used because of some psychological reasons. In fact, a driver whose only task is that of supervising the process, would not be able to act properly in case of danger due to the lack of attentiveness and driving practise. To overcome this problem, some rail lines (like the Victoria line of London Underground) have introduced positive tasks for the driver.
- Full automation. In this case no driver is present on the train since all operations are completely automated. However, a person who is normally in charge of other tasks can take control if necessary. Examples of fully automatic systems are some metropolitan lines of Paris, Lille, London, Vancouver, and Copenaghen.

2.4.12 Principles of interlocking machines.

The interlocking machines are an arrangement of signal apparatus which prevents conflicting movements within an area with crossings and junctions.

Basically, an interlocking system can be divided into three main functions (Theeg and Vlasenko, 2009):

- The '*operation*' *control level*' which includes the interface to the signaller and may include different non-vital functions of automatic operation control such as automatic train routeing;
- The '*interlocking level*' which is constituted of the vital functions to interlock signals, routes, movable track elements, block application with each other;
- The '*element control level*' includes instead functions of commanding, power and information transmission to and from the field elements such as signals, movable track elements, track sections, level crossing etc.

Over the years, the interlocking had a technical progression according to the following categories which are briefly described:

- Human or manual, without technical support. Actually, this interlocking cannot be considered a real one since no technical locks are provided. Indeed, the signaller or the shunting staff is totally responsible for checking the preconditions for clearing signals, switching movable track elements and for transmitting information to the field elements by walking between them. Nowadays, this method has been widely replaced by more technical and advanced solutions;
- Mechanical interlocking. In this system, the signaller operates mechanical levers which are interlocked with each other. Power and information transmission to the field elements is by wires and rods.
- Electric (relay) interlocking. This is the case where signaller operates buttons. Indeed, the interlocking functions are in relay technology and the field elements are operated and controlled electrically.

- Electronic interlocking. In this interlocking system, all functions are performed and control by hardware and software. Indeed, the interlocking logic is defined in programmed software.

The abovementioned kinds of interlocking are pure forms. However, several hybrid forms still exist since the shift from mechanical to electrical, as well as from electrical to electronic, took place in several steps.

Table 2.2 shows the classification of interlocking systems considering the function levels they provide.

	Operational level	**Interlocking level**	**Element control level**
Human interlocking		Human (to be reminded by the signaller)	Human (walking between the elements)
Mechanical interlocking	Mechanical levers	Mechanical (lever frame)	Mechanical (wires or rods)
Electric relay interlocking	Electric buttons and illuminations	Electric (relays)	Electric
Electronic interlocking	Electronic (monitor, mouse or tablet, keyboard)	Electronic (software/hardware)	Electronic

Table 2.2 Basic interlocking technologies and technical application of the functions (Theeg and Vlasenko, 2009).

2.5 Timetable and capacity allocation.

The last paragraph of this chapter concerns the basic principles of timetabling and capacity allocation. Basically, all the topics dealt with in the previous paragraphs related to legislation, signalling system and train space separation are necessary to the comprehension of the following issue.

Timetabling is a very complex designing and planning process through which all train operations are considered in order to offer a stable and robust service which satisfies travel demand requirements. For this reason, timetable prescribes working conditions of the whole railway network by planning arrival and departure time of trains, dwell times at stations, train headways and connections between runs. Moreover, timetable dictates also operations which are not directly linked to the ordinary service, such as shunting movements on siding tracks for composing or maintaining rail vehicles.

The process of designing a timetable starts from the evaluation of the scheduled train running time so as to understand the time a train needs to complete its path (Hansen and Pachl, 2008). In particular, it consists of the following components:

- Pure running time between scheduled stops
- Dwell time at scheduled stops
- Recovery time
- Scheduled waiting time.

The pure running time is the shortest running time possible and can be calculated by a running time estimation as shown in paragraph 2.3.4. Obviously, in order to give a train the possibility to recover from small delays, some recovery time must be added. There are two kinds of recovery time:

- Regular recovery time
- Special recovery time.

The regular recovery time is basically added to every train path as a percentage of the pure running time. In European railways, this percentage is generally a value between 3-7% and there are two different approaches to consider it: some railways prefer to spread recovery time over the train path while others prefer to concentrate it at the end of the run and at large intermediate terminals (Hansen and Pachl, 2008).

The special recovery time by contrast, is used in the case of maintenance and construction works or sections with temporarily bad track conditions. It is not added as a percentage of the running time but as a fixed supplement on the section concerned.

Scheduled waiting time is considered for scheduling reasons such as to synchronise schedules of different passenger lines at changing points, or to

wait for a scheduled passing or overtaking. Generally, this supplement is added to the dwell time at scheduled stops but sometimes also to the running time.
To design a timetable, the estimation of the train path is not sufficient. Indeed, as already explained, the signalling system strongly influences the operational use of a line by a train and overall, the interactions between consecutive trains. To this purpose, the application of the '*blocking time model*' and the construction of the '*blocking time stairways*' (see paragraph 2.4.5 for more details) is necessary to identify critical network sections, such as bottlenecks or singular sections when certain kind of train conflicts can arise and to plan the headway between runs and the required buffer time for reducing transfer of delays from one train to the next.

After the timetabling process, the evaluation of the consumed capacity of the line has to be performed. As shown by the UIC Code 406, capacity is based on the relations between the following parameters:

- The number of trains. In fact, the more trains are, the less capacity is left for traffic quality;
- The average speed. The braking distance increases proportionally more than the average speed;
- The stability. As already explained, in order to avoid the propagation of minor delays, margins and buffers have to be added to the running time of trains and between paths;
- The heterogeneity. The more are the differences between the train running times, the more capacity will be consumed.

The relation between these parameters are shown in the so-called '*capacity balance*' in Figure 2.34 (UIC Code 406, 2004). As can be seen, a chord links the points on the axes, expressing the value for each parameter and the length of the chord corresponds to the capacity. The capacity utilisation is then defined by the positions of the chord on the four axes.

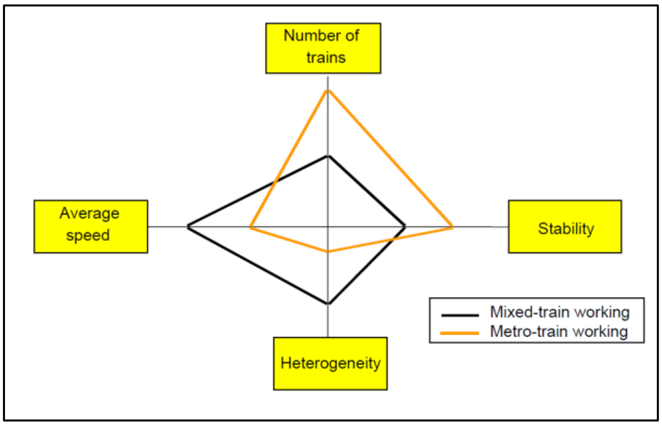


Figure 2.34 Capacity balance (source: UIC 406 Code)

Obviously, likewise to service quality (see paragraph 2.1.9) even capacity can be viewed differently according to the subject considered. Indeed, while from a market point of view, capacity demands are oriented to satisfy peak values, infrastructure planning is interested in a definition of capacity which guarantees a profitable utilisation of the infrastructure. From a timetable standpoint by contrast, capacity considerations are necessary to define train paths trying to fulfil travel demand needs on a given infrastructure. Finally, from an operational point of view, capacity evolves continuously and depends on current infrastructure availability, delays, diversion and number of additional trains. For this reason a unique definition of capacity is not possible. However, a general accepted definition proposed by UIC Code 406 is the following:

'*the total number of possible paths in a defined time window, considering the actual path mix or known developments respectively and the Infrastructure Manager's own assumptions*'.

The consumed capacity of a line can be evaluated by the compression method (UIC Code 406, 2004). The procedure consists of virtually moving the blocking time stairways together, up to the minimum theoretical headway without adding any buffer time and without changing the trains' order. During this process, neither the timetable running times nor the given overtakings, crossings or stopping times may be changed. Figure 2.35 highlights the results of the procedure.

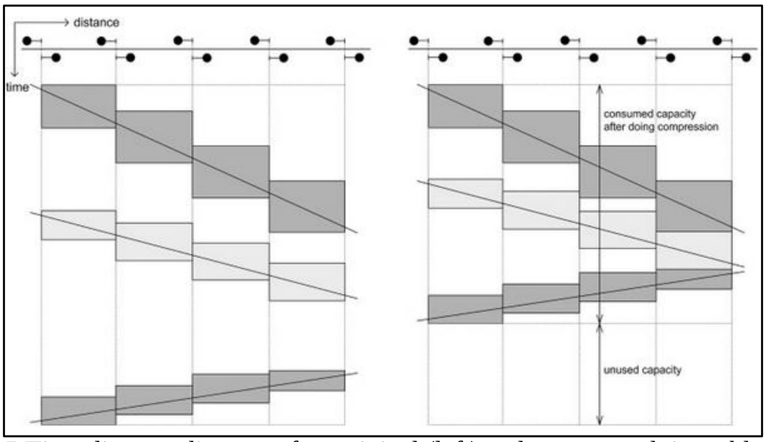


Figure 2.35 Time distance diagram of an original (left) and compressed timetable (right) (source: Lindner, 2011)

After this first step, capacity consumption is measured considering the infrastructure occupation in a defined time position, to which buffer times are added for timetable stabilisation and, where necessary, maintenance requirements (see Figure 2.36).

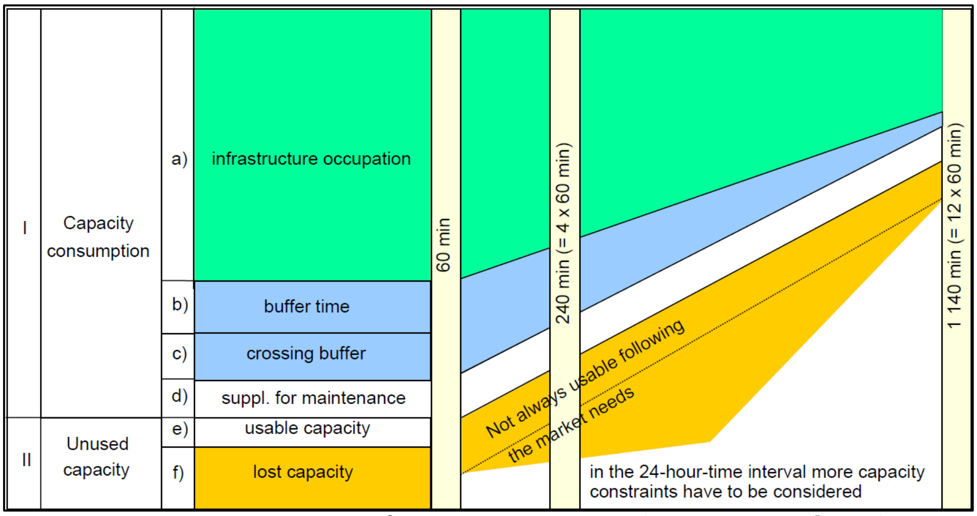


Figure 2.36 Determination of capacity consumption (source: UIC code 406)

In particular, the formula for the determination of the capacity consumption is:

$$k = A + B + C + D \tag{2.24}$$

$$K = k \cdot 100 / U \tag{2.25}$$

where:

- k is the total consumption time [min];
- A is the infrastructure consumption [min];
- B is the buffer time [min];
- C is the supplement for single track lines [min];
- D is the supplements for maintenance [min];
- K is the capacity consumption [%];
- U is the hosen time window [min].

The difference between the chosen time window and the capacity consumption is the '*unused capacity*'. This amount may be divided in '*usable capacity*' and '*lost capacity*'. The former is the case in which additional train paths can still be added, while the latter is the opposite situation.
The capacity consumption value has to be compared to standard indexes. If the infrastructure occupation is higher than a typical value, the analysed line section shall be called congested and no more additional train paths may be added to the timetable. If the infrastructure consumption is lower than a typical value by contrast, further analysis must be carried out so as to include additional train paths. UIC Code 406 proposed some recommendations about the typical values to adopt as capacity consumption limits (see Table 2.3). However, these values have to be considered as guideline and not fixed since each line section has its own characteristics and constraints.

Type of line	Peak hour	Daily period
Dedicated suburban passenger traffic	85%	70%
Dedicated high-speed line	75%	60%
Mixed-traffic lines	75%	60%

Table 2.3 Guidelines for capacity limits (source: UIC code 406).

In conclusion, this chapter has provided the basic knowledge of railway systems in terms of legislation, technology and operation. Indeed, working on railways requires firstly an in-depth analysis of all elements which are part of this complex system.

In particular, the first section of this chapter has focused on the regulation of the last twenty years which has completely liberalised the sector. As a result, nowadays rail undertakings pay serious attention to service quality in order to increase the attractiveness of the service. This is the reason why, the decision support system presented in this thesis focuses on customers' needs instead of analysing just operational aspects of the service.

The following paragraphs have instead provided a description of the infrastructure, rolling stock and signalling systems as well as timetabling and capacity allocation. These are key elements for representing railway networks within a microscopic simulation model.

CHAPTER 3: RAIL SIMULATION MODELS AND THEIR APPLICATIONS IN REAL CONTEXTS.

In all engineering fields, the use of simulation assumed great importance for supporting planning and management phases of any kind of process. In particular, in the last decade, thanks to the development of powerful software and hardware, simulation of railway systems has become an important day-by-day tool for several applications such as timetabling, dispatching operations, and rescheduling solutions. Indeed, simulation approaches give the possibility to evaluate the interactions of railway operations, to estimate system performance as well as effects of implemented strategies prior to their application in real contexts. Therefore, benefits are numerous providing the saving of a lot of money and above all, avoiding disturbance during real railway operations.

According to the target of the analysis, different methods of simulation are adopted. In particular, due to dissimilar theories, railway simulation models can be classified in:

- Macroscopic/mesoscopic/microscopic, in agreement with the network scale;
- Synchronous/asynchronous depending on the processing technique;
- Deterministic/stochastic for the assumptions on the analytical approach.

This chapter analyses all kinds of simulation models providing a brief description of the main examples around the world. In addition, the state of art of rescheduling procedures after the occurrence of disruptions or disturbances to the service is presented so as to evaluate different approaches adopted for managing the rail system during this complex phase.

3.1 Infrastructure modelling and graph theory.

Before analysing the different simulation models, it is worth introducing the basic concepts of infrastructure modelling.

Generally, railway infrastructure is modelled using models derived from Graph theory (Hauptmann, 2000; Radtke and Watson, 2007). In fact, thanks to their flexibility, graph models can represent even the most complex network through an efficient mathematical model.

A graph is defined as an ordered pair of sets:

- a set of nodes (called 'vertex' in mathematical terms) which are a representation of an arbitrary location in a railway network;
- a set of links (called also 'edge') which are the connections between two nodes.

Therefore, a graph can be represented as follows (Hansen and Pachl, 2008):

$$G := (V, E, c) \tag{3.1}$$

Where *V* (i.e., Vertex) is the set of nodes, *E* (i.e. Edges) is the set of links and *c* is the weighted function:

$$c(e) \geq 0 \ \forall e \in E . \tag{3.2}$$

Moreover, the graph can be defined as:

- '*directed*', if two adjacent nodes are linked with at least one link and the direction is indicated as an arrow;
- '*simple*', is the graph does not contain parallel links or loops;
- '*connected*', if for any two nodes of the graph, links exist connecting the node.

Figure 3.1 shows a real track layout and its representation with a directed (the direction of the link is indicated by an arrow) and simple (there are no loops or parallel links) graph.

Obviously, links and nodes can contain attributes which describe the railway infrastructure: for example, typical node attributes are geographical information (kilometre marking, position coordinates, names and so on) or infrastructure elements (signal, points, timing points etc.).

In particular, two different approaches are possible:

- Link-oriented models where each link contains all relevant information such as speed, gradient, radius, direction, electrification;
- Node-oriented models where special nodes indicate the changes of an attribute. This means that links do not contain any information while each node has positional coordinate. The length of a link can be evaluated by the difference of the position of the start and the end node.

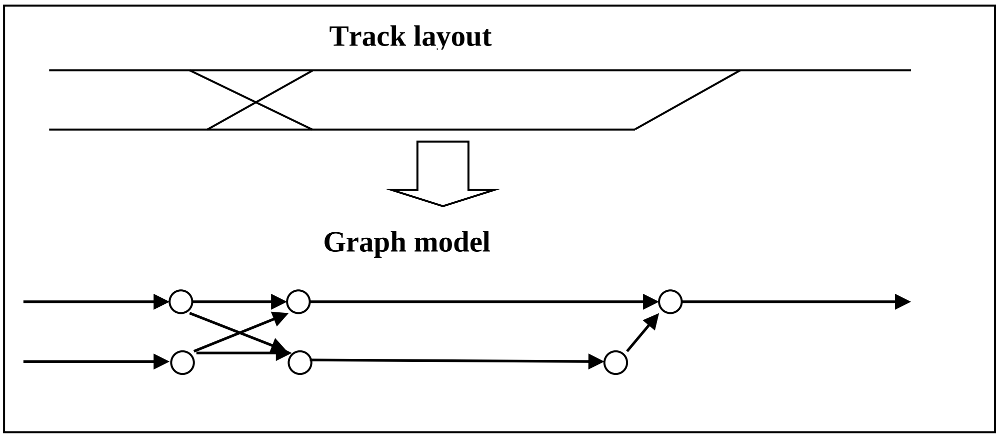


Figure 3.1 Graph model of a simple track layout.

The main disadvantage of a link-oriented model could be the redundant assignment of all railway infrastructure attributes to each single link, which causes great waste of storage capacity and data-handling problems. However, thanks to modern infrastructure editors and powerful PCs, these negative aspects are nowadays solved.

Concerning the node-oriented models, an advantage of their use is the redundancy of free data storage. For example, to introduce a speed change, it is necessary to set a node of the type ‘speed node’ at the desired position. On the other hand, these models require complicated algorithms to calculate all the attributes on a certain link of the network.

An alternative approach to node modelling is the so-called colon-graph or double vertexes graph which allows to represent properly train movements. In fact, each node can be crossed if and only if both vertexes of the node are crossed. This is a valid representation especially in the case of points. Looking at the example in Figure 3.2, it is obvious that only path from A via B to C or via D to E are feasible while a path D-B-C is excluded since node B is not completely crossed.

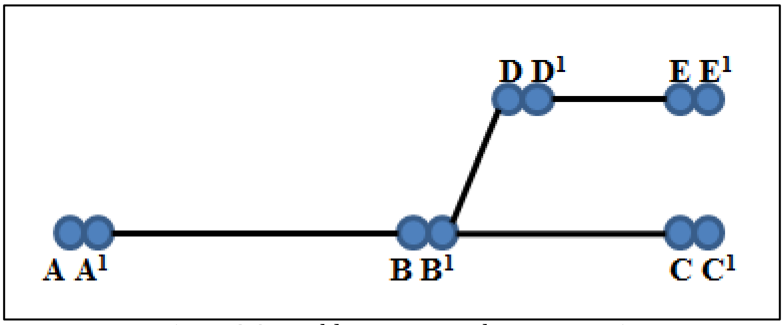


Figure 3.2 Double-vertex graph representation

However, the level of infrastructure detail strongly influences the characteristic of the graph. For this reason, it is necessary to distinguish among microscopic, macroscopic and mesoscopic models. The former are characterised by a graph which, depending on the purpose, contains the highest level of details on nodes and links. In the case of macroscopic models by contrast, infrastructure information is aggregated. Finally, mesoscopic model graphs are syntheses of both microscopic and macroscopic approaches. To give an idea about the reduction of infrastructure information passing from a microscopic to a macroscopic model, an example (Figure 3.3) is provided. As can be seen, a node in a macroscopic model represents a station or a junction, while the microscopic one contains all information of the considered rail line.
Anyway, the following paragraphs will give more details about these models.

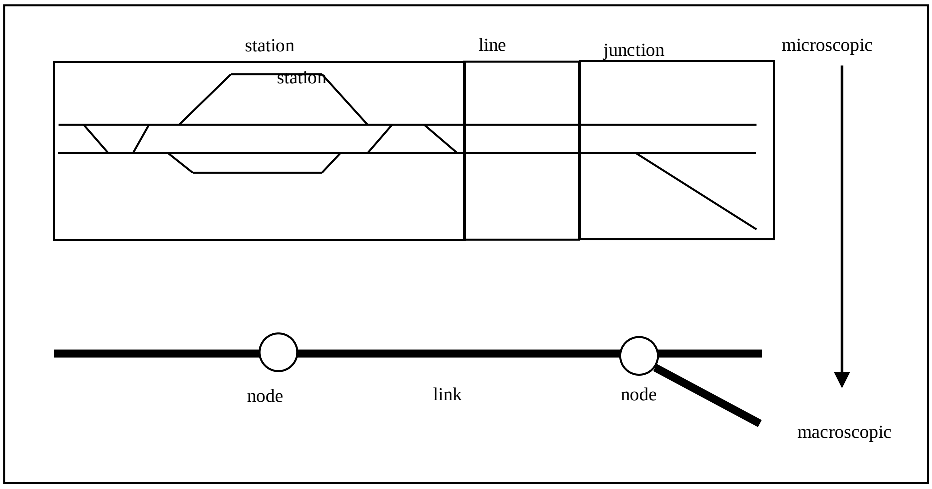


Figure 3.3 Differences between macroscopic and microscopic models (source: Hansen and Pachl, 2008).

3.2 Macroscopic simulation models.

Macroscopic models are generally adopted in long term planning tasks or routing problems. As said in the previous paragraph, the infrastructure representation contains less nodes and links than the microscopic one since it is sufficient to have a more abstract view of the network. Furthermore, the macroscopic data can be easily entered manually from various sources; generally, a microscopic database of the infrastructure is preferable, but even public sources (i.e. internet) can be adopted.

In particular, the nodes need some information related to geographical attributes (coordinates, names etc.) or to their typology (station, shunting yard, junction etc.). Macroscopic links by contrast, contain details on the length, type of line (high speed, passenger, freight or mixed line), number of tracks, train availability, average running time and average capacity (according to UIC 406, see paragraph 2.5).

The main applications of macroscopic network models are the strategic evaluation of different infrastructure scenarios or the search for train paths without time restrictions. In particular, the latter, as shown in Sewcyk et al. (2007), is a two-steps process. The first search on a macroscopic model considers just the sequence of stations for the train trip and neglects all individual tracks on lines or stations. Therefore, only some fundamental criteria are considered such as axle load, electrification or operational rules (i.e. stopping patterns for passenger trains or preferred routes for high speed trains). After finding the initial solution by means of the Dijkstra algorithm, the process has to be completed on a microscopic network so as to consider all constraints and evaluate the feasibility of the presumed train path.

However, macroscopic models cannot be used for the evaluation of running time or conflict detection between trains since they do not contain information about the exact maximum speed and track gradient or restrictions due to the signalling system. Therefore, it is not possible to evaluate a correct running time, but only an estimation based on average values of velocity. On the other hand, the ease of computation, not only does it give the possibility to obtain the output in short time, but it generally enables a more complete and complex analysis of the network. That is why, usually, macroscopic models are adopted for simulating the rail system taking into account also travel demand.

In the literature, different macroscopic models have been developed either for consulting purposes or for academic research. Three of the main examples are described in the following paragraphs. Two of these, namely the '*Nemo*' and the '*Simone*' models, concern specifically the macroscopic simulation of railway networks while the third, the '*TransCAD*' model, is a more generic transportation planning software.

3.2.1 Network Evaluation Model (NEMO).

The Network Evaluation Model (NEMO) is a strategic network planning and traffic evaluation model developed by the Institute for Transportation, Railway Construction and Operation (IVE) at the University of Hanover (for more details see Sewcyk and Kettner, 2001). In particular, this macroscopic simulation tool is extremely useful for supporting railway companies in the planning optimisation of any activity concerning infrastructure and railway operations. In fact, it gives the possibility to evaluate the effects of changes in the existing network providing information about the benefits of new transport services. Since the variations of transport offers (such as changes to the network's structure or to operation programs) strongly influence passengers' mode choice, the model considers the dynamic interaction between offer and travel demand. Furthermore, the model evaluates both costs and earnings that result from the provided traffic services. The calculations' results are of great importance for the operators as well as for the infrastructure managers due to the fact that benefits of investments can be estimated in advance.

The NEMO model is composed of the following different modules (Figure 3.4):

- Infrastructure module;
- Traffic module for passenger trains;
- Traffic module for freight train;
- Evaluation module.

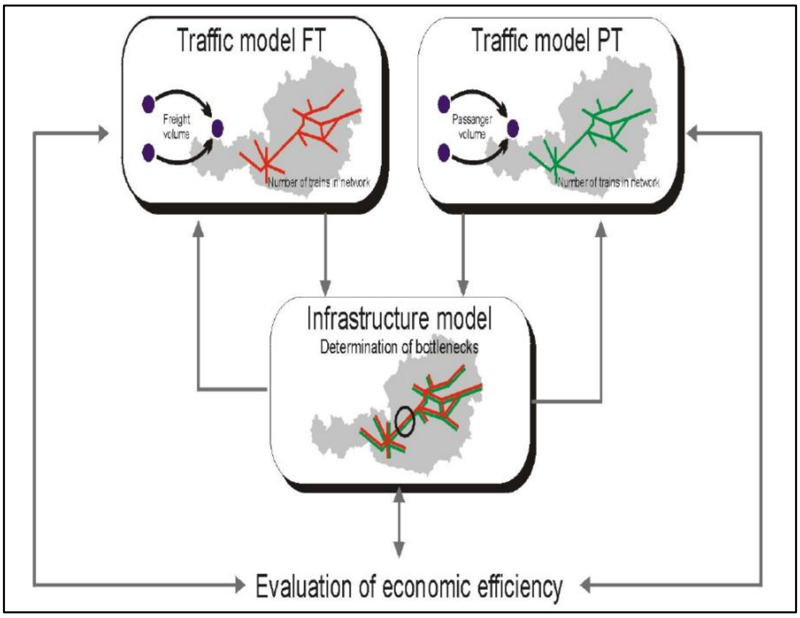


Figure 3.4 Framework of the 'NEMO' model (source: Sewcyk and Kettner, 2001)

The infrastructure module represents both the given or the planned railway network as a link-oriented graph. As all macroscopic models, nodes correspond to all entry points of the passenger and freight traffic (i.e. stations) or junctions, whereas links represent the tracks between consecutive nodes (Figure 3.5).

The traffic module for passenger trains computes the balance between traffic offered and demanded and the whole process is based on the search of this equilibrium. As input data, the restrictions given by the infrastructure and the current or planned offer for passenger traffic are required. Basically, the module works according to the following steps:

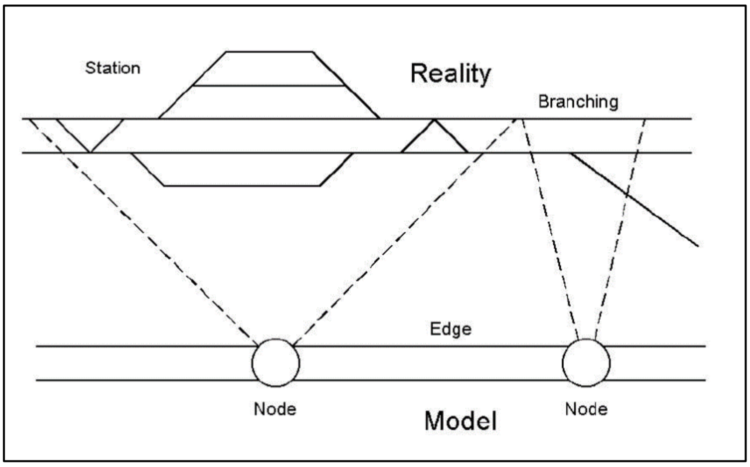


Figure 3.5 Macroscopic representation of the 'NEMO' model (source: Sewcyk and Kettner, 2001)

- Evaluation of the traffic volume: in a first step, the whole traffic with origin or destination from/to a region is determined, without considering the means of transport;
- Evaluation of traffic relations: the whole traffic within a region is divided considering the different destination regions and it is assigned to the network using the individual entry points;
- Route search: the passenger quantity in between the entry points is assigned to the model trains. Since network links contain attributes about costs and travel times, the route search can be found minimising these quantities;
- Evaluation of train composition: by using the outputs of the route search, the number of model trains on each network link can be determined. Furthermore, in this phase the capacity of trains can also be taken into account;
- Offer comparison: the output of the train composition is the train offer which is compared to the original one.

This process is repeated cyclically until an equilibrium is reached (Figure 3.6).

The traffic volume is evaluated before the first simulation and it is considered fixed. If the new offer presents changes, these changes will influence the travel demand. This effect is evaluated by the modal split model which determines the relation between railway and road traffic considering travel times and costs of the different means of transport.

The traffic module for freight trains works like the previous module based on the freight traffic’s quantity in between different regions. Therefore, information about the infrastructure and the production system within the freight traffic on rail is necessary.

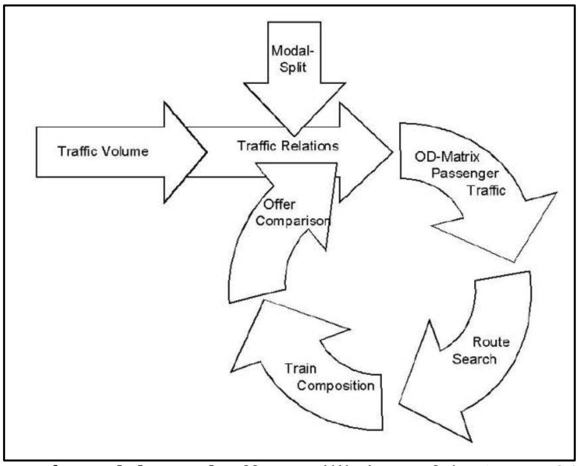


Figure 3.6 Sequence of travel demand-offer equilibrium of the ‘NEMO’ model (source: Sewcyk and Kettner, 2001)

In this case, the process is slightly different from the previous one and can be summarised as follows:

- Evaluation of traffic volume: for each analysed region, the originating freight carried by train is computed. These amounts are assigned to the entry points within a region by means of an origin and destination matrix;
- Production systems of freight traffic: the whole amount of freight is assigned to the production systems of freight traffic. When the quantity between two entry points exceeds a certain amount, block trains which serve directly the two points are created. The remaining transport volume is worked off considering single wagon traffic and the wagons are rearranged in shunting yards;
- Route search: the wagon volume is then assigned to the network taking into account that all production systems have their own specific model

trains. In the single wagon traffic, the condition that wagons have to go through a chain of shunting yards is added. In this way, the module computes the fastest and cheapest route within the network;

- Train composition: for each link, the number of necessary model trains is evaluated. Therefore, the resulting load of the infrastructure by the freight traffic can be determined. In this case, the modal split with road traffic is not calculated but the transport quantity determined at the beginning can be reduced, if a defined transport time is exceeded. Furthermore, at each entry point, empty stock wagons are taken into account by determining the shortage or surplus of freight wagons. This imbalance is levelled by an optimised disposition of empty stock wagons.

The outputs of both traffic modules therefore define the passenger and the freight traffic. Those calculated train numbers are combined for the total infrastructure load considering different time slices. Within each time slice, the infrastructure elements (nodes and links) are examined so as to identify possible bottlenecks. In case the number of trains exceeds the capacity, appropriate measures have to be chosen by the operator in order to solve the problem. In particular, some solutions could be the adoption of different routes for those trains which are not bound to certain stops, the change of train speeds or the designing of new infrastructure elements.

Finally, the evaluation module on the basis of the computed train offer, the demand and the load of the infrastructure, calculates all arising costs and earnings. To this purpose, fixed cost for the infrastructure, model trains as well as earnings for a given transport services are predefined. In this way, the economic evaluation of the different computed scenarios can be carried out and all benefits of these strategic measures can be estimated in advance.

3.2.2 SImulation of Model NEtwork (SIMONE).

The '*SIMONE*' model is a macroscopic simulation tool for simulating and analysing complex and large scale train networks (Middelkoop and Bouwman, 2001). The software was developed by Incontrol (which is a worldwide simulation consultancy firm) and Railned (i.e. the capacity manager for the Dutch railway network) in order to support strategic planning decisions such as the possibility to develop new railway infrastructure or the allocation of network capacity to train operating companies. Moreover, Simone can be used for assessing the robustness of timetables, determining the stability of the network and analysing cause and effects of delays.

The Simone architecture is composed of several integrated applications which work seamlessly together (Figure 3.7):

- Incontrol Center
- Simulation Library
- Infra and Timetable Database Interface
- Automatic Model Generator
- Simulation Models
- Scenario Manager
- Output Generator
- Output Analyser and Manager

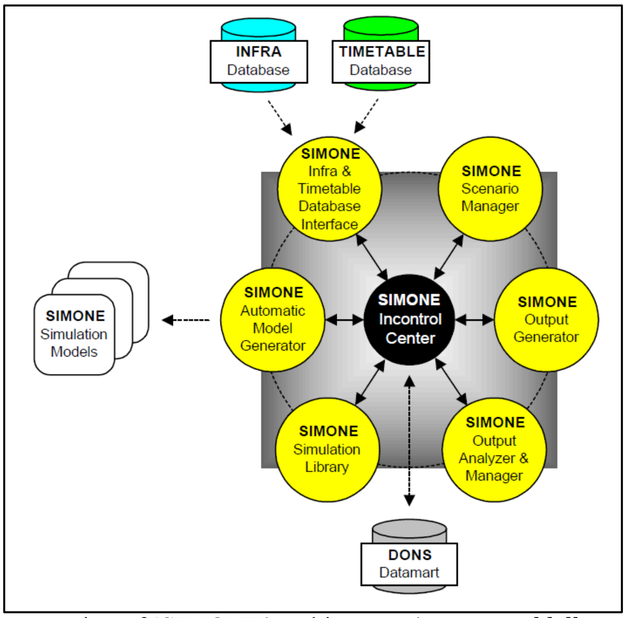


Figure 3.7 Representation of ‘SIMONE’ architecture (source Middelkoop and Bouwman, 2001)

The *Incontrol Center* is the core of Simone simulation environment from where everything is controlled. All necessary information for the simulation is stored in the Oracle database of Railned.

The *Simulation Library* is a collection of six simulation modules (namely the Simulation setting module, the Network setting module, the Statistic module, the Timetable module, the Station module and the Connection module) which are necessary to construct a Simulation Model. In particular the Station Module and the Connection Module are used for representing the infrastructure network. The other modules allow the addition of extra inputs such as different types of capacity constraints, possible conflicts between trains at stations, the bidirectional use of tracks and so on. The timetable module contains information about running times, dwell times, slack and the track allocation for the connecting tracks.

The *Infra and Timetable interface* gives the possibility to generate simulation models based on timetable and infrastructure information contained in a database. In particular, Simone is designed for being interfaced with the DONS database, which is the database of the Dutch railway network. Once all input data concerning the railway network infrastructure and the traffic demand are provided, the module automatically generates a cyclic timetable.

The *Automatic Model Generator* can generate simulation models without user intervention. In particular, after specifying the model settings, information is automatically extracted from the database and is used to construct a simulation model.

The *Simulation Model* contains a graphical representation of the network. During the simulation, the software shows running trains through different colours according to their typology (e.g. Intercity, freight or high speed trains) and to their delay.

The *Output Generator* provides statistical analyses of both infrastructure elements and trains with different level of aggregation. This means that outputs can range from the whole model containing all trains to the single station or the single train. Figures 3.8 and 3.9 provide two examples of SIMONE outputs.

Finally, the *Output Analyser and Manager* is a tool for the analysis of the simulation model. Basically, it provides the possibility to compare the performance of several scenarios and evaluate the effects of different planning strategies on the network.

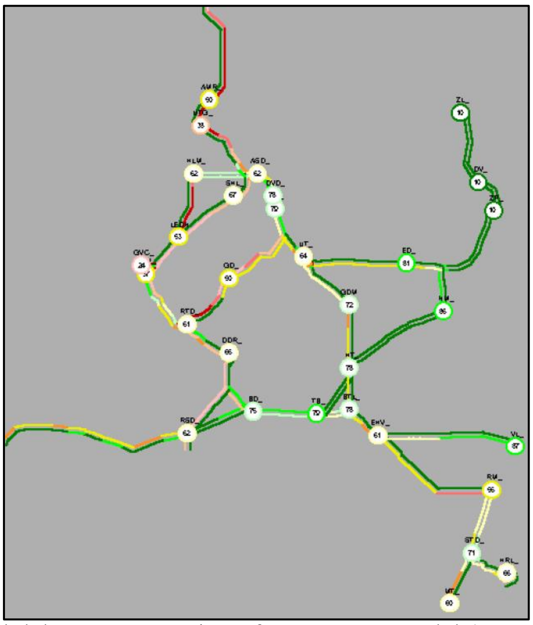


Figure 3.8Graphical delay representation of 'SIMONE' model (source: Middelkoop and Bouwman, 2001)

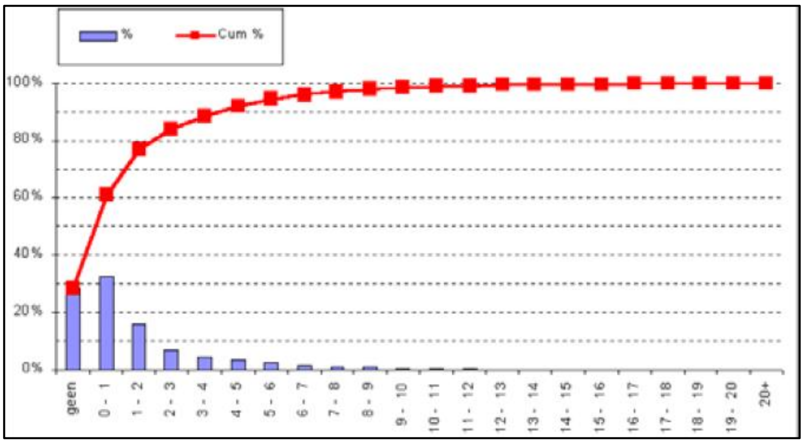


Figure 3.9 Punctuality of trains on a selected path of the network (source: Middelkoop and Bouwman, 2001)

3.2.3 The TransCAD model.

The 'TransCAD' model is the first GIS (Geographic Information System) designed specifically for the analysis of all transportation modes at a macroscopic level of detail. It combines GIS and transportation modelling capabilities in a single integrated platform in order to provide the most complete simulation of the transportation system.

Roads, public transport lines as well as rail lines are represented by means of weighted oriented graphs called ‘*networks*’ (Figure 3.10). These are linked to a database where several kinds of attributes can be stored like, for instance, link classifications and performance functions, intermodal or interline terminals, transfer points, and delay functions, transit access, egress, and walk transfer links etc. Networks are stored in a highly-efficient way, enabling TransCAD to solve routing problems very quickly.

Figure 3.10 Network representation in TransCad.

The program adopts matrices to hold data such as distance, travel times, and origin-destination flows that are essential for many transportation applications. In particular, it supports several travel demand modelling such as sketch planning methods, four-step demand models, activity models, and other advanced disaggregate modelling techniques. Furthermore, the integration with GIS functions provides demand forecasting in response to changes in regional development, demographics, and transportation supply. Traffic assignment can be performed through different advanced models, namely:

- Multi-modal toll road assignment
- Origin user equilibrium

- Path-based assignment
- Multi-point equilibrium assignment
- Combined distribution-assignment
- Assignment with traffic signals and HCM intersection delay
- Dynamic equilibrium traffic assignment.

As regards railway lines, they are considered as transit networks. In this case, transit assignment models are used to estimate the number of passengers that utilise segments in a transit network as a function of transit level of service. These models take as input a matrix of passenger flows between origins and destinations and a transit network, and produce link levels and aggregate ridership statistics. Methods are included that are sensitive to fares and park and ride access, as well as equilibrium assignments which take into account the capacity of transit service and the effect of ridership on crowding, comfort, and, optionally through dwell time effects, on travel time on the route. These methods distribute the flow between a particular origin and destination to multiple paths, based on their relative attractiveness. The transit assignment procedures produce a table of ridership at every stop along each route in the transit network. Optional outputs include critical link analysis, boarding and alighting counts, stop-to-stop flows, route-to-route transfers, and aggregate ridership counts.

3.3 Microscopic simulation models.

Microscopic models represent rail infrastructure through high detailed node-link model which combines track information such as speed, gradient or radius, with the signalling system (signals, block sections, release points) and some operational information like routes, alternative platforms and timing points. Every time there is a single change in one of the abovementioned attributes, a new node splitting an existing link and generating a new one is required. Information can be assigned to either the node or the link depending on the adopted approach.

The number of input data required by this kind of model is very large and it includes the followings:

- Length of a link;
- Gradient;
- Permissible speed;
- Speed indications and speed boards;
- Electrification;
- Radius;
- Signalling system;
- Overlaps;
- Release contact and clearance location;
- Track circuits;
- Stop boards;
- Blocks and routes;
- Interlocking techniques.

This accurate representation makes microscopic models essential for the exact estimation of running times, timetable construction and simulation and conflict detection and resolution.

Figure 3.11 shows a typical microscopic infrastructure model. As can be seen, signalling system is depicted in detail providing the sequence of block sections which can be composed of several links. More block sections can thus be grouped in route sections which specify the correct train path through the line or through each station. This simple approach can be used to transfer technical requirements or operational rules into the model. Indeed, only route sections which are feasible from the operational point of view can be created. For instance, a route section which is supposed to be used by electric locomotives must not be created in a model if some links are not electrified.

However, due to the large amount of information, microscopic models are inefficient from a computational point of view when simulating large-scale networks or when supporting analyses which need a consistent number of simulations (e.g. probabilistic analyses, black-box optimization problems). In these cases, they can be used only by means of complex off-line procedures which take long time to be completed.

Several microscopic simulation programs have been developed to fulfil both research and practical objectives. In the following paragraphs three models will be described: the OpenTrack model developed by the ETH-Zurich and the RailSys model realised at the Leibniz Universitat of Hannover, which are two examples of commercial software; the EGtrain model, which is a microscopic simulation tool addressed to research purposes.

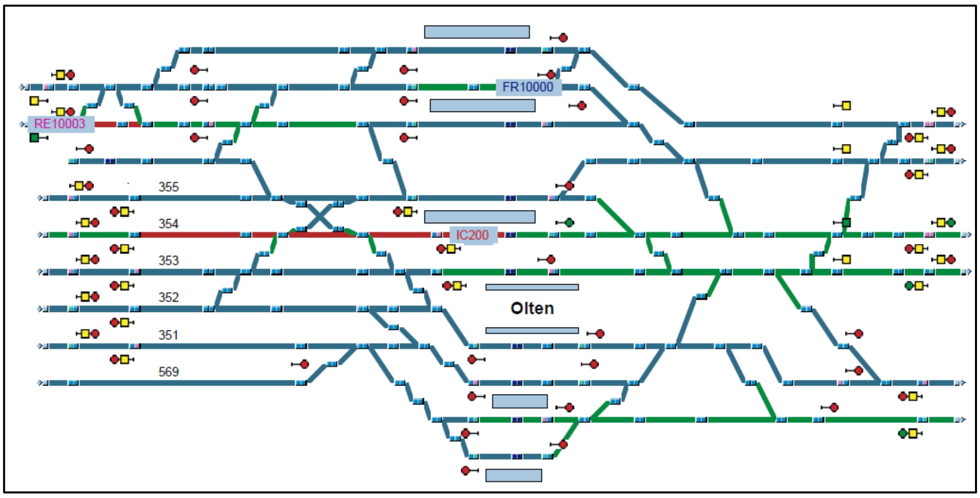


Figure 3.11 Example of a microscopic representation (source: www.opentrack.ch)

3.3.1 The RailSys model.

The Railsys software (Radke and Bendfeldt, 2001) was developed initially at the Leibniz Universitaet Hannover and then it is being further improved and distributed in cooperation with Rail Management Consultants. The software includes accurate tools for running time calculation, infrastructure mapping, timetable-construction and evaluation and planning of vehicle rosters. The complete architecture of the model is shown in Figure 3.12.

Rail lines can be modelled by means of an infrastructure editor which allows the user to model the whole network with the accuracy of one meter. The location of points, signals, stations, stopping points, speed indicators, platforms and tracks are considered and can be edited manually or imported with interfaces from other data sources.

The exact calculation of running time is performed by '*Dynamics*' which considers the traction force/speed diagram (see paragraph 2.3.1) of the locomotives as well as the weight and length of each vehicle. It includes an interactive tool to get an exact overview of the calculated train run and to modify all input data interactively. Additionally, energy saving driving strategies, computation of the location of signalling infrastructure or check of the alignment of the track can be undertaken.

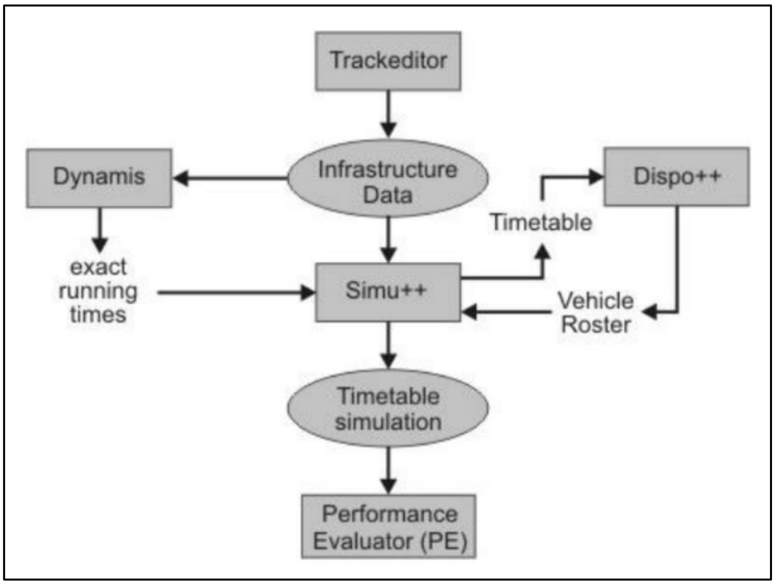


Figure 3.12 Framework of the 'Railsys' model (source: Radke and Bendfeldt, 2001).

The outputs of the running time calculations can be then transferred into '*Simu++*' so as to perform the timetable construction. The handling of large railway networks consisting of many lines and stations is supported by a modular set up of the railway network. Every line in a large-scale network can be planned and simulated independently. Any line can be joined to a network at any time during the planning process and can be separated again afterwards. For example, the network of a railway company can be separated into several local areas or sectors (i.e. suburban areas). Trains running on more than one line can be scheduled by supervising planners. Simu++ indicates inconsistencies such as unfeasible connection times and differing transit tracks in stations for these trains. Furthermore, Simu++ supports also the different signalling systems specified by the ETCS standards.

'*Dispo++*' provides the optimal allocation of locomotives in large networks according to given boundary conditions (e. g. reversing times, compulsory connections or track slot prices). In this way, it is possible to minimise the number of locomotives and to reduce the number and length of empty train runs considering also maintenance tasks.

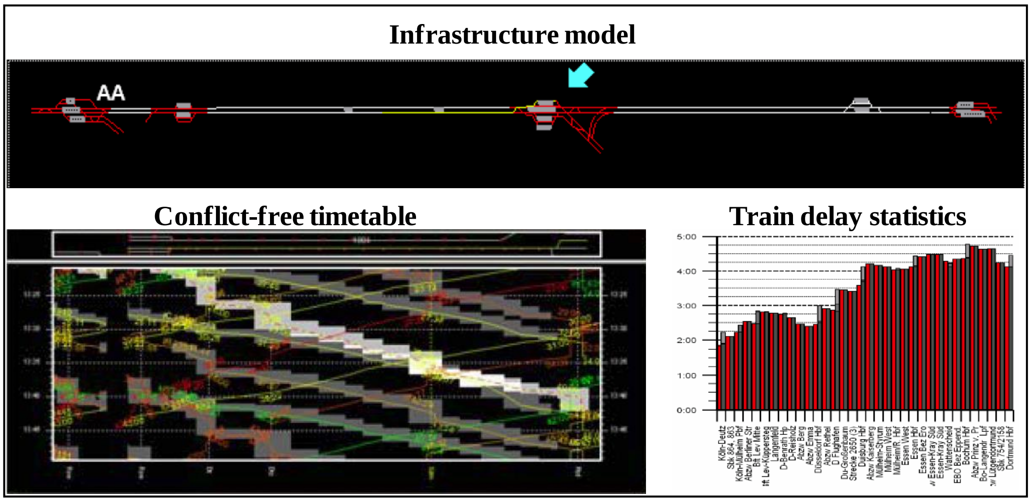


Figure 3.13 Simulation outputs of the 'Railsys' model (source: Radke and Bendfeldt, 2001).

Finally, the Performance Evaluator calculates the impact of infrastructure and/or timetable alternatives. This tool is intended to evaluate performances of the simulated operational program. The main bases for this task are the delays of trains during the simulation, which are statistically prepared and analysed. Results of the simulation can be displayed as performance of the whole network or only some lines and/or stations.

3.3.2 The OpenTrack model.

The 'OpenTrack' model is one of the most used commercial software in Europe. It was developed by Nash and Huerlimann at the ETH of Zurich (Huerlimann, 2001; Nash and Huerlimann, 2004).

This software is a complete and potential microscopic simulation model whose architecture is summarised in Figure 3.14.

Input data are divided into three modules: rolling stock, infrastructure and timetable. Rolling stock consists of locomotive and wagons which are combined to form trains. The user can enter all details concerning the tractive effort diagram, the length as well as the weight for single axle so as to reproduce them realistically.

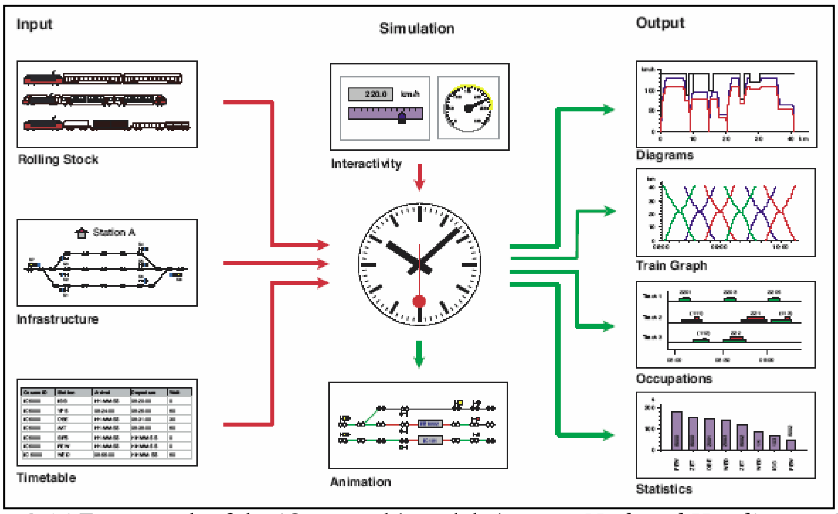


Figure 3.14 Framework of the 'Opentrack' model. (source: Nash and Huerlimann, 2004)

Infrastructure is represented by means of a double vertex graph which allows directional data to be more easily managed. Vertices and links can be used to assign track layout data like length, gradient, and maximum speed for different train categories. Other important data are related to the signalling system implemented, to the length of block sections through positioning their delimiting signals on graph nodes, and to the type of interlocking systems which regulate train movements within station or junction areas.

Timetable data consists of information on the movement of trains. They include desired arrival and departure times, connection information, minimum stop times, and stop information. Dwell times can be also set up specifying different values for each station and train categories. Moreover, according to the analysis targets, all these data can be considered as deterministic or stochastic variables, giving the possibilities to consider stochastic disturbances which affect real operations.

The simulation of train is carried out using a mixed discrete/continuous simulation process that calculates both the continuous numerical solution of the differential motion equations for the vehicles (trains) and the discrete processes of signal box states and delay distributions.

After the simulation process, a large variety of outputs can be provided (Figure 3.15) as, for example, train motion diagrams (speed-distance, speed-time, distance-time trajectories), occupation times of rail sections (in both numerical and graphical format), statistics, such as percentage of delayed trains at a certain station, overall train punctuality (fixing a certain delay threshold), energy consumption diagrams (electrical or mechanical power-time diagrams, electrical or mechanical energy-space diagrams).

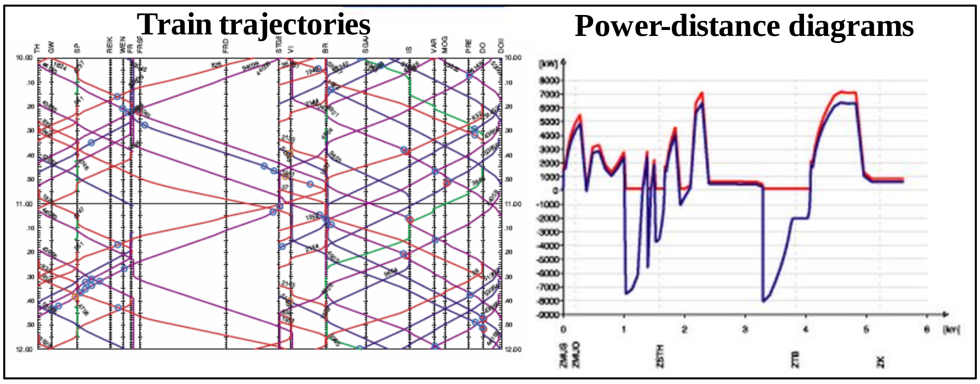


Figure 3.15 Simulation outputs of the ‘Opentrack’ model (source: Nash and Huerlimann, 2004).

3.3.3 The EGTrain model.

Commercial programs, as the ones presented in the previous paragraphs, are very useful for consulting purposes but have the great disadvantage of being sold as a black box. This means that the code which is behind the program is not known to the user and often it is not possible to interface it with other tools. This aspect could be a great drawback for research purposes inasmuch as it does not allow the development of new functions and the interactions between different models. For this reason, it is worth mentioning ‘EGTRAIN’ (Environment for the design and simulaTion of RAIlway Networks) developed in C++ language to overcome the applicability limits of commercial models (see Quaglietta, 2011; Quaglietta and Punzo, 2013).

This object-oriented microscopic simulation tool gathers input data in the following four modules:

- Infrastructure module. The railway network is modelled through a link-oriented graph where the links contain information about track attributes (speed limits, gradients, curvature radii) and the nodes include details about spatial coordinates of signals and stations.
- Rolling stock module. Rail vehicles can be represented considering both their physical and mechanical characteristics. The 'tractive effort-speed' curve of the traction unit, the maximum deceleration rate, the jerk value, as well as the train composition (number of wagons, masses of coaches and traction units, etc.) are some examples of input data of this module. Moreover, a further sub-module for the calculation of the mechanical energy consumption is included.
- Signalling system module. A specific module is addressed to the specification of the signalling system. In particular, different signalling system can be implemented such as the Italian BACC system, the ETCS level 1, and the ETCS level 2.

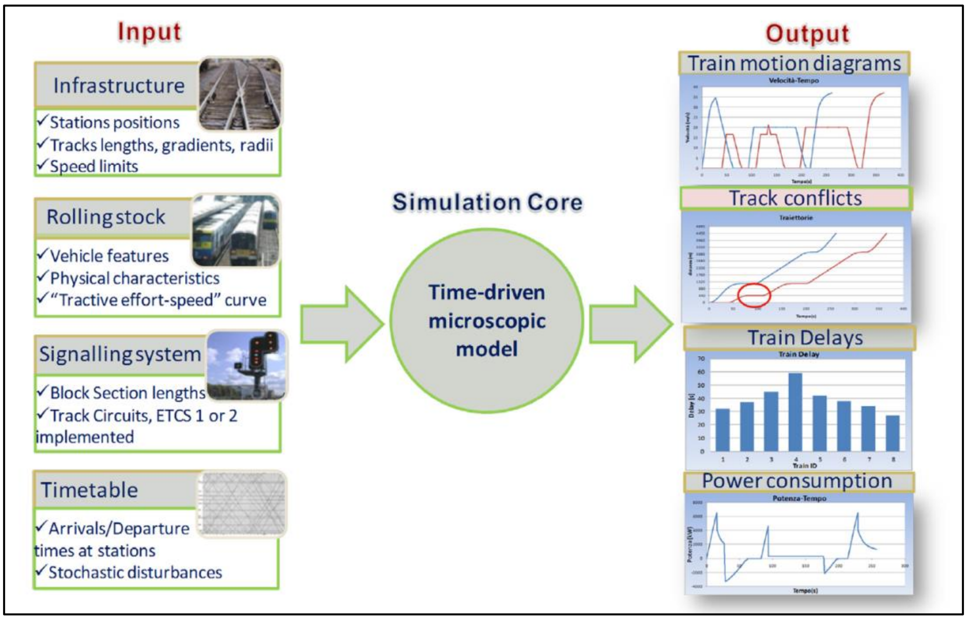


Figure 3.16 EGTrain architecture (source: Quaglietta, 2011).

- Timetable module. All data regarding departure/arrival times and/or minimum dwell times at stations, can be set up in this module. Moreover it is possible to introduce disturbances to ordinary train operations by imposing a deterministic or a stochastic delay to a specific train at a certain station. In the case of stochastic simulations, random variables are modelled by their probability density function (pdf), as well as the mean and the standard deviation of the pdf itself.

As all microscopic simulation models, train movements are simulated by performing a time integration of the Newton's motion formula. After this process, different output data can be obtained like for instance, train diagrams, track conflict and energy consumption diagram.

3.4 Combination of microscopic and macroscopic models.

Previous paragraphs have shown advantages and disadvantages of both macroscopic and microscopic models. In particular, as already explained, the macroscopic approach needs low input data and provides outputs in short computational times but it is not able to evaluate precisely the interactions between trains. The microscopic approach, by contrast, produces accurate results but needs a great number of input data and requires long computational time. Therefore, recently, many authors have developed innovative models which combine the two approaches taking advantage of the benefits of both.

Before dealing with these new procedures, it is worth analysing in depth the differences between the two infrastructure models. To this purpose, Figure 3.17 shows a small section of a railway line where the vertical axis indicates the speed limit over the section and the horizontal axis describes the length of the section. Basically, the microscopic model is composed of several different links which start at the exact position where an attribute changes value (for the sake of simplicity, in this case just the speed parameter is considered). The macroscopic model, by contrast, represents the same infrastructure using just one link. Therefore, there is a problem of assigning a speed limit which would be consonant with the different values of the microscopic model. Three different options are possible:

- To use the lowest microscopic speed limit;
- To use the highest microscopic speed limit;
- To use the average speed limit considering the proportionate length of the microscopic links.

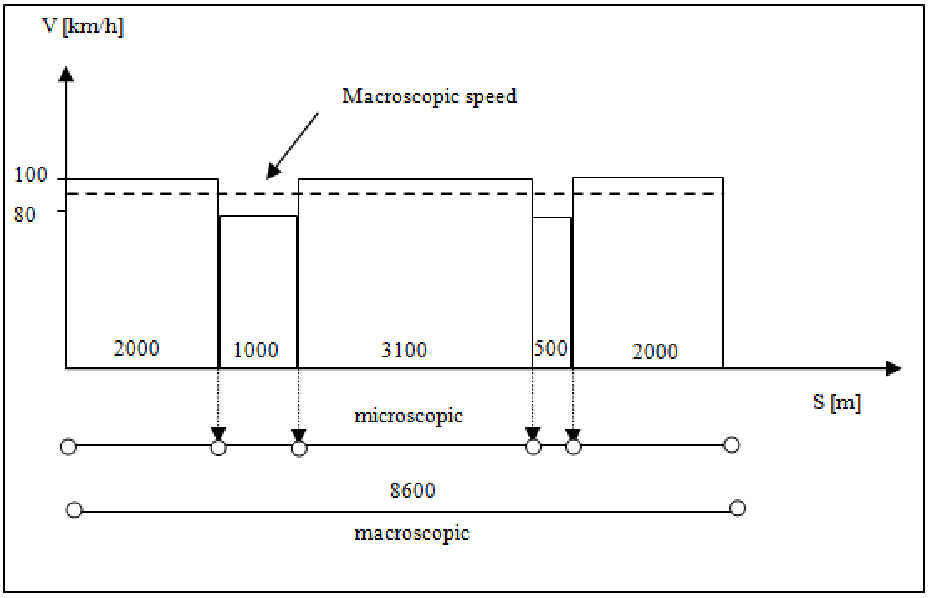


Figure 3.17 Relation between macroscopic and microscopic models (source: Hansen and Pachl, 2008)

A survey of running time estimations using the three options of macroscopic speed was carried out by Radtke, (2005). Results showed variations between +6% (too fast) and -20% (too slow). This is the reason why macroscopic models should not be used for running time calculations, timetable construction or conflict detection and resolution. However, they can be adopted for identifying possible initial solutions in short computational time which can be tested afterwards by a microscopic model. In this case, the problem of migrating from one infrastructure model to another one has to be solved. Starting from a microscopic model, since it contains far more links and nodes, it is evident that the macroscopic model can be derived in a straightforward manner reducing the number of information. This approach is called ‘*bottom-up*’ as indicated in Figure 3.18 (Hansen and Pachl, 2008).

The ‘*top-down*’ approach, by contrast, can be used for generating artificial microscopic infrastructure whose level of detail depends on the targets of the analysis.

In this context, Kettner et al. (2003) proposed to combine the NEMO model with the Railsys model in order to develop a procedure for the railway network evaluation simplifying data storage, data administration and especially the process of data acquisition. In particular, by means of a special interface, all infrastructure data stored within the microscopic software (i.e. Railsys) are transferred to the macroscopic one (i.e. NEMO) with a lower level of details (bottom-up approach).

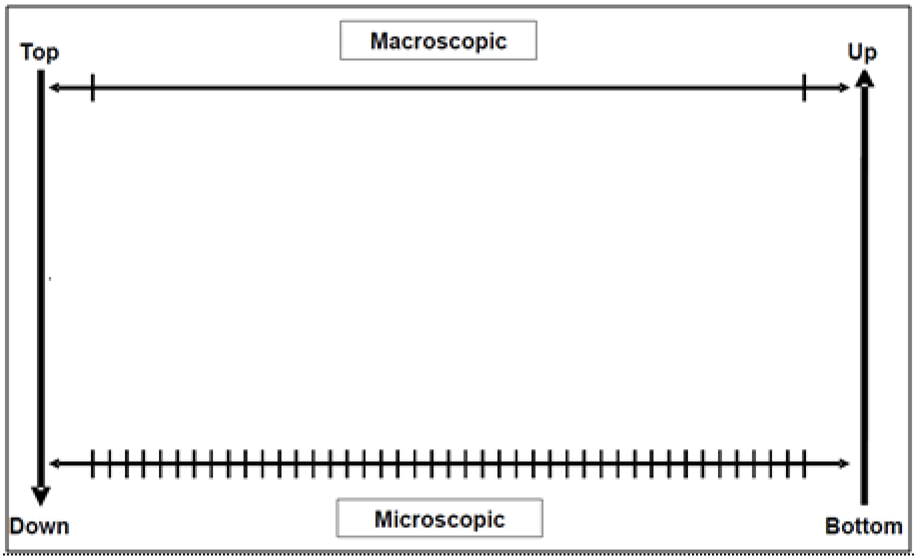


Figure 3.18 Migration approaches between infrastructure models (source: Hansen and Pachl, 2008).

Ones the macroscopic network is automatically created and all required attributes are set up, the NEMO model can run its simulation providing for instance, possible train paths according to the passenger and freight demand (see paragraph 3.2.1 for more details concerning the outputs of NEMO).The Railsys model can then evaluate these solutions so as to check the convergence to a conflict free timetable.

Schlecthe et al. (2011) described an algorithm for micro-macro transformation of railway networks. In particular, starting from a microscopic rail line (developed in OpenTrack), the procedure generates a new infrastructure model aggregating block sections to macroscopic tracks and station areas to macroscopic nodes. In order to take into account the interaction between different trains in particular points, like convergence, divergence and crossing routes, the algorithm generates 'pseudo-nodes'. This kind of nodes has to be considered as pseudo-stations where trains cannot stop or change direction. The following step consists of the evaluation of rounded values of running times based on microscopic simulation data in order to obtain aggregated values for the macroscopic model. After the transformation is completed, an optimisation process is performed for determining conflict-free track allocations. Finally, a comeback to the microscopic model ensures the feasibility of the solution obtained.

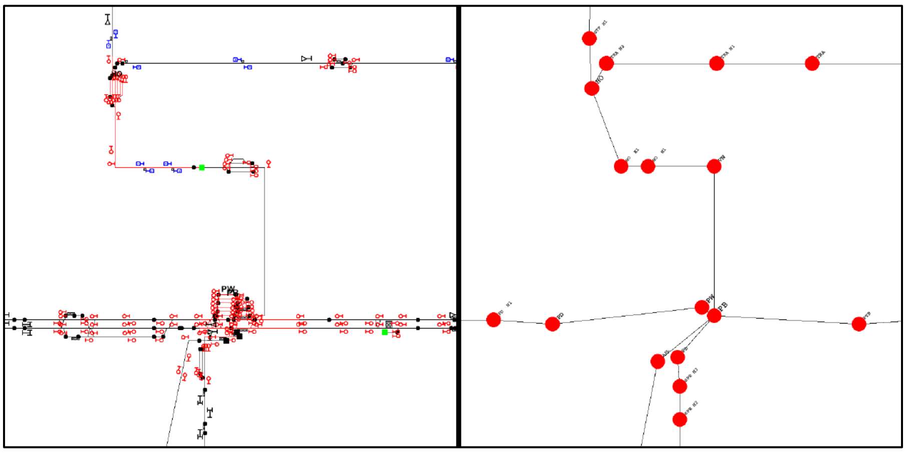


Figure 3.19 Transformation of microscopic into macroscopic network.

3.5 Mesoscopic simulation models.

Mesoscopic simulation models are placed in between macroscopic and microscopic models. Indeed, the rail network is modelled combining areas which are depicted on a microscopic level and areas represented on a macroscopic level. This the reason why this paragraph has been placed after the definition of both macro and micro models.

A great advantage of this procedure is the possibility to reduce the efforts of modelling complex infrastructure sections which are not relevant for the overall targets of the investigation. In the case of planning a wide network timetable, for instance, tracks have to be represented considering their all characteristics and the signalling system while stations can be described with less details neglecting shunting yards or vehicle depots. In this way, point elements of a macroscopic representation (e.g. stations), are better depicted since individual tracks and possible paths approaching or crossing the station can be modelled.

However, no specific commercial mesoscopic tools have been developed for simulating railway except from some software addressed to meet pure research requirements as the ones described in the following.

3.5.1 A mesoscopic model for simulating freight train operations.

Marinov and Viegas (2011) proposed a mesoscopic simulation modelling methodology for analysing and evaluating freight train operations in a rail network. In particular, the simulation is performed by means of Simul8, which is an event-based simulation computer package. The rail network is modelled as a queuing system where all components are interconnected and interact each other. More in-depth, the system is composed of Work Centres and Storage Areas. The former replicate the operating processes with freight trains (i.e. where a freight train is served by a component of the rail system) and are characterised with inbound traffic, service pattern and outbound traffic. The inbound traffic is the number of freight trains waiting for entering a Work Centre; the service pattern is reproduced through a particular distribution inasmuch as information is obtained by observations, real data collection and statistical analysis; finally, the outcome of the Work Centre is the outbound traffic which is routed to other Work Centres and Storage Areas. The Storage Areas are attributes describing the places where the freight trains are held while waiting to be processed by a given component of the rail network. The completion of the train service is replicated by other attributes called Work

Exit Point which are subordinated to Arrival Patterns. Work Flow Arrows by contrast, provide the connections between Work Centres and Storage Areas. As outputs, the simulation model computes the total number of freight trains processed by a given Work Centre, the number of freight trains in a given Storage Area, queuing (waiting) time per freight train on average for the period of the experiment, utilisation levels of the rail network subcomponents and utilisation rates of system resources. Figure 3.20 illustrates the transition from the microscopic representation of a rail marshalling yard to the mesoscopic model developed in Simul8.

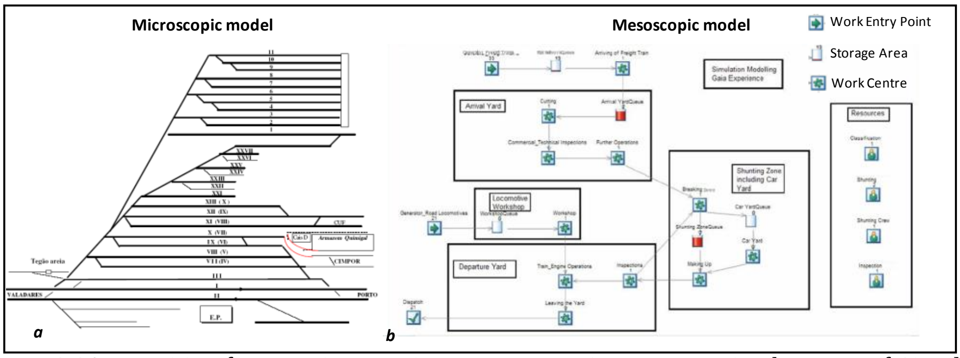


Figure 3.20 Transition from a microscopic to a mesoscopic representation by mans of Simul8 (source: Marino and Viegas, 2011).

3.5.2 A hybrid mesoscopic-microscopic railway simulation model.

As already explained, the limits of applicability of microscopic models concern the computational efficiency in the case of simulation of large-sized or complex railway networks, and probabilistic analyses aiming at the evaluation of effects of components breakdown. However, this problem could be solved through the implementation of a hybrid modelling methodology which ‘dynamically’ integrates microscopic and mesoscopic approaches as shown in Quaglietta (2011) and Quaglietta et al. (2011b). Basically, the procedure allows the simulation of large-scale railway networks at mesoscopic level, with the possibility to focus at a microscopic level on those sections where local dynamics need to be investigated at a higher detail. While the microscopic software adopted is the EGTrain tool (described in paragraph 3.3.3), the mesoscopic package is an event-based multi train simulation model which depicts the network as a graph where nodes represent block section joints and stations, while links are block sections. Train movement on such links is modelled as a sequence of activities whose durations is equal to the free-flow train travel times on the corresponding track sections. As input data, the model requires train operation attributes like free-flow train travel times for each block section and dwell times at stations as well as train headway. Moreover, information regarding failure rates and corresponding MTTR (Mean Time To Restore, i.e. the average time to restore ordinary conditions) for critical components (vehicles, signalling system, track equipment, etc.) are also necessary so as to run stochastic simulations. In this way, the model can be used to evaluate global impacts of failures on network operation, assessing reliability, availability, maintainability and performance levels of different track layouts and fall-back strategies. Indeed, as outputs, the model provides simulated train arrival/departure times to/from each block section and each station and a series of performance parameters such as reliability, availability, maintainability, and punctuality.

3.6 Synchronous and Asynchronous models.

In addition to the different scale of representation, simulation models can be divided in synchronous and asynchronous models.

Synchronous models simulate all events within the network simultaneously. For this reason, all trains are included in the network at the same time influencing each other and their status is updated continuously. Every single step of the simulation follows the real chronological progression of time and the system has to react immediately to any kind of situation.

The main target of synchronous simulation is the modelling of the interactions between the trains in the system. Basically, each considered train has a status which changes during the progress of time according to the results of the running time calculations, the signalling system and the rules and measures of the dispatching subsystem. This kind of simulation is also called ‘event-driven simulation’ since an event is the time-related occurrence where the status of a train changes and affect other trains within the network. Some examples of event are: the start of a train; passing a station, a distant signal or a main signal; start or finish of a station stop. If a train has to stop at a main signal because it is hampered by another rail convoy, it is stored in a special queue and no following event is created until the signal aspects gives the authorisation to proceed.

A lot of commercial synchronous simulation tools are available on the market. OpenTrack and Railsys belong to this typology. Other models with the same characteristics are: VISION and RAILPLAN developed in the United Kingdom, FALKO and TRANSIT distributed by Siemens and RAILSIM commercialised by Berkley Simulation Software in the USA.

Asynchronous simulation, by contrast, does not consider all trains at the same time but divides the simulation in more steps following a particular criterion which is related to the category the trains belong to. First, only the trains with high priority are included within the network. Afterwards, step by step, also the other classes of train with lower priority are simulated. Generally, long distance passenger trains have the highest priority while freight trains the lowest. The former are therefore influenced only by disturbances which happen to themselves, like longer dwell time and technical failures, and they are not influenced but low priority rail convoys. These ones instead, fit into the time windows which are left by high priority trains and experience more delays than they would in the reality.

This kind of simulation is especially suitable for timetable construction since it reproduces the process of timetable design. Moreover, if an asynchronous simulation model is employed to analyse railway operations, trains with high priority are always preferred and displace trains with lower priority.

It is evident that asynchronous simulation is more static than synchronous simulation. The fact that at every simulation step the state of every individual element (e.g. train, signals) of the simulated reality is known, enables the synchronous simulation to be more flexible than asynchronous simulation models. In fact, in this case the status of processed trains is not altered after they are inserted into the simulation timetable unless appropriate dispatching rules are implemented. For instance, if train delays are higher than a certain threshold they can be assigned a higher priority so as to come back to their schedule as soon as possible.

Few examples of asynchronous models are presented in the literature. It is worth mentioning BABSI (Gröger, 2002) and STRESI (Shultze, 1985), both developed by the RWTH of Aachen (Germany).

3.7 Deterministic and stochastic models.

The last classification of simulation models concerns the assumptions on the analysed variables which are the input data of the simulation. In particular, in a deterministic model, all events like departures, arrivals and running times are constant values established in the timetable. These kinds of model are mainly used for the designing or the primary evaluation of timetables when it is necessary to check if there are scheduled conflicts due to the overlap of blocking time. Another application could be the performance validation of system elements within the network like, for example, the signalling system.

Stochastic simulation models, by contrast, consider arrival and departure times, dwell times, or running times as random variables. They are mainly applied for evaluating the robustness of timetables against operation disturbances, testing network stability and checking the feasibility of possible operational strategies.

3.8 The adoption of simulation models for railway tasks.

The possibilities to adopt simulation models for railway tasks are numerous. Basically, their applications can be classified according to three different time horizons:

- Strategic planning
- Medium-term planning
- Short-term planning

Strategic or long-term planning starts about five to ten years before the opening of new infrastructure or the starting of new operation services. Therefore, information regarding train types and running times is not available during this phase. A macroscopic simulation with approximate data about infrastructure and rolling stock characteristics can be used to obtain a draft of operational program. Obviously, it is not possible to design a detailed timetable but, considering also travel demand, a survey on the feasibility of the project can be carried out.

The time interval of a medium-term planning is generally one year. In this case, detailed information is available and a microscopic simulation model can be set up. Exact running time calculations are indeed possible, as well as track occupations. The simulation supports the planner in order to obtain a conflict-free timetable. Furthermore, computations are performed without adding external perturbations to compute the delays resulting from the timetable itself.

Short-term planning involves two different applications of simulation. The first one concerns the capacity analysis of an existing timetable so as to add additional train paths. Microscopic simulation is in this case necessary. The other application is part of the daily operations and consists of finding rescheduling actions to bring back the system to the planned railway operations when disturbances or delays occur. Starting from the actual situation, the dispatcher can use microscopic simulation model to develop different scenarios looking for the best strategy. However, in order to intervene rapidly in real time, sometimes also macroscopic models can be adopted to achieve this task.

In the following paragraphs, taking into account the objectives of this thesis, further in-depth analyses about rescheduling simulation models and optimisation models are presented.

3.9 Dispatching and Rescheduling models.

During real time railway operation, disturbances or disruptions to the service can occur creating delays and conflicts between trains. Disturbances are generally considered as small perturbations influencing the system, while disruptions indicate large external incidents which can lead to the cancelation of train runs within the timetable. In both cases, dispatchers have to react appropriately so as to re-establish the ordinary conditions as rapidly as possible. Generally, rescheduling actions are still taken according to the experience of dispatchers as a result of rules which have been proven in practice. Possible control actions to reduce the propagations of delays include changing dwell times at scheduled stops and changing train speeds along lines or train orders at junctions, stations and passing points. Major modifications can be changing train routes or even cancelling train runs which often involves also rolling stock and crew rescheduling.

However, numerous models have been developed as real time dispatching support system to help dispatchers during their daily operations. Basically, these tools are composed of the following components (Hansen and Pachl, 2008):

- Conflict detection module, which determines potential conflicting train routes within a pre-established time horizon considering the current infrastructure status, timetable and rolling stock information, the position and speed of each running train;
- Conflict resolution module, which, according to the actual delays and the predicted conflicts, proposes the most suitable strategies to re-establish the ordinary service conditions.

Conflict detection modules are generally based on microscopic simulation tools. Indeed, in this process, it is necessary to evaluate a space-time diagram with all simulated train trajectories and this is possible only through a detailed description of trains and network.

Different approaches can be instead adopted for the conflict resolution module. The most used are asynchronous models, synchronous models and optimisation models.

As already explained, asynchronous models reproduce precisely operating processes but they do not perform a time-step simulation. Trains are included according to their hierarchical rank and, if disturbances occur and conflicts between train runs arise, these are solved following a chronological order.

Therefore, at every step of the simulation a new conflict timetable is reproduced. An example of asynchronous model for automatic conflict detection and resolution was presented in Jacobs (2004) and used in the ASDIS traffic-regulation tool. In particular, when conflicts between trains are detected, the tool provides solution options such as using alternative routes, extending dwell time at scheduled stops, adding extra stops or increasing running times.

Synchronous simulation models, by contrast, do not give the possibility to generate automatic non-conflicting schedules. The process considers all trains simultaneously and there is no roll-back in the simulation. For this reason, this method can be only used for the evaluation of the effects of possible recovery strategies using a 'what-if' approach, meaning that only a limited set of intervention scenarios can be evaluated.

Optimisation methods for conflict resolution are part of an active area of operational research. Many models presented in the literature showed the effectiveness of these approaches in terms of solution quality and computational time (Cacchiani et al., 2013). In particular, these procedures differ each other according to the delay severity (disturbances or disruptions), the scale of network representation (microscopic or macroscopic) and the analysis targets (minimising train delays or customer dissatisfaction).

The majority of the models concerns the rescheduling problem in the case of disturbances adopting the Alternative Graph model (Mascis and Pacciarelli, 2002; D'Ariano, 2008) for a microscopic description of the infrastructure. Basically, the alternative graph ($G = (V, F \cup A)$) is defined as follows. Each node corresponds to an operation and the passing of a train through a block section and then through the successive block section in the fixed route for the train is represented by a fixed arc belonging to set F. The length of the arc is the train running time on that block section. Two trains requiring the same block section at the same time cause a conflict. Therefore, a processing order and sufficient headway for the corresponding conflicting operations is modelled by pairs of alternative arcs belonging to the set A (indicated in Figure

3.21 as dashed lines). In this way, train separation and safety requirements are reproduced accurately. However, it is worth noting that this model can be considered as a '*fixed-speed*' microscopic model, since it assumes that train travel times on a block section are deterministic parameters, whose values correspond to undisturbed running times. Based on this method, D'Ariano et al. (2007a) proposed a branch-and-bound algorithm for scheduling trains in real-time. Since the model can include hundreds of block sections and trains, the authors decided to add a dynamic and static rules to the algorithm so as to produce outputs in a very short computational time although the approach is based on a microscopic scale.

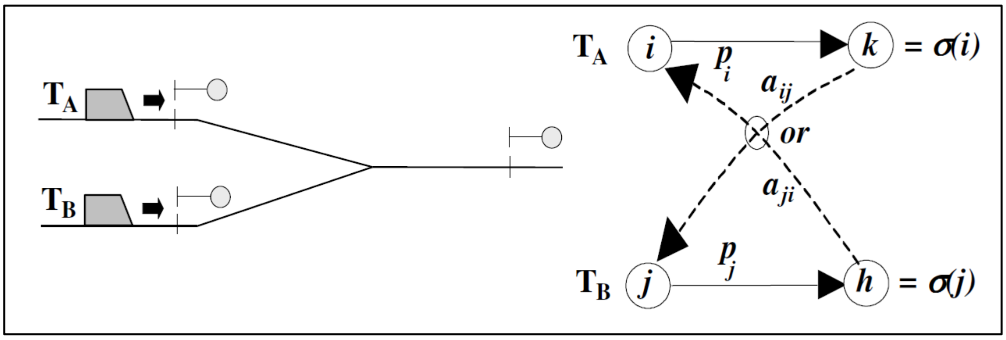


Figure 3.21 The 'Alternative-graph' representation.

The static rules are computed off-line on the basis of the network topology. An initial partial solution (possibly empty) is calculated and a list of partial solutions is maintained during the process. At each step, the algorithm chooses a partial solution from the list and an unselected pairs of alternative graphs $((i, j),(h,k))$. Taking into account the static implication rules, two partial solutions are then built up, one containing the arc (i, j) and the other one containing the arc (h,k). Dynamic implications are therefore computed. If the lower bound is smaller than the value of the best solution found, the two partial solutions are added to the list. Applications of this algorithm were performed on a real network area around Schiphol Amsterdam Airport, in the Netherlands. Results demonstrated the efficiency of the procedure to provide feasible dispatching solutions in short computational time.

In D'Ariano et al. (2008a), the alternative graph model was implemented in ROMA (Railway traffic Optimization by Means of Alternative Graphs). This tool aims to help dispatchers solving conflicts caused by delays and disturbances.

ROMA was then improved with numerous contributions (Corman et al., 2009; 2010a,b,c; 2011b; 2012; D'Ariano et al., 2007b; 2008b; D'Ariano and Pranzo, 2009). Basically, all these papers concerned the dispatching problem in the case of larger and busier railway networks than a single area or severe disturbances to the service.

In particular, in D'Ariano et al. (2008b) some timetabling constraints are relaxed in the ROMA tool in order to produce real time flexible timetables. This means that, contrary to current practice based on the formulation of a rigid timetable where many variants are discussed in depth and all possible conflicts among trains are solved by means of an off-line procedure, the author showed how flexible timetables are preferable than rigid ones. Indeed, flexibility, which consists of time windows with minimum, maximum arrival/departure times and a set of feasible platform tracks for each train and for each station, offers more freedom to solve conflicts and increases punctuality. Furthermore, the use of advanced optimisation algorithms (as the one implemented in ROMA) for conflict resolution improves the benefits of flexible timetables in terms of delay minimisation.

However, the tool ROMA does not take into account the dynamic evolution of randomly disturbed traffic conditions. In fact, the implementation of dispatching solutions is subjected to stochastic events and hence the system can react differently from what is expected. Quaglietta et al. (2013) combined ROMA with the microscopic simulation software EGTRAIN so as to evaluate the stability of recovery plans within different time horizons. More in detail, at regular time intervals, optimal plans are computed by ROMA on the basis of updated traffic information gathered from EGTRAIN. Experiments on the Dutch railway corridor Utrecht-Den Bosch demonstrated that shorter prediction horizons give a more stable but less effective control strategy, since optimal plans computed by ROMA mostly suggest retiming. Larger horizons consent by contrast, also reordering train strategies and give the possibility to manage traffic more effectively, but lead to more unstable solutions. Therefore, the paper showed that time horizon has to be considered in the definition of automated dispatching systems when random and dynamic traffic conditions are included in the analysis.

Other similar contributions in the definitions of recovery solutions based on a microscopic approach are Flamini and Pacciarelli (2008), Mannino and Mascis (2009), Caimi et al. (2012) and Pellegrini et al. (2012).

In particular, Flamini and Pacciarelli (2008) focused on the problem of real-time management of rail traffic in a metro terminal. The purpose is to develop an automated train traffic control system able to directly implement most traffic control actions without the authorisation of the local area manager. The procedure is divided in two steps. The first one provides a feasible schedule (i.e. define a routing in the terminus, and a departure time from the terminus in a given time horizon) with minimum train delays adopting an alternative graph model. In the second step, the solution evaluated is improved optimising the headways between trains.

Mannino and Mascis (2009) developed a branch-and-bound similar to the one adopted in D'Ariano (2007) but with a different objective function. Indeed, the proposed method consists of enumerating all the feasible routings for the trains, and then solving for each routing the one which minimises the deviations of the actual schedule from the original plan and the costs due to violating regularity. The algorithm was tested on an application provided by Azienda Trasporti Milanesi (ATM), namely the major municipal public transport company of Milan in Italy, showing the benefits of this new approach in terms of increase of punctuality and regularity.

Caimi et al. (2012) described a model predictive control framework for railway traffic management in complex central railway station areas. Basically, the procedure manages traffic by retiming and rerouting of trains as well as partial coordination of speed profiles. The approach is based on a closed-loop discrete-time control system. A forecast module computes a forecast of the evolution of the system. The rescheduling model computes a new disposition timetable, which can be combined with the forecasted time effects in the system for the dispatcher. Once the dispatcher takes a decision, this is forwarded to the trains and the intended commercial offer is changed accordingly, thus closing the control loop. The possibility to obtain realistic solutions in short computational time makes this model viable for practise.

Pellegrini et al. (2012) proposed a mixed integer linear programming formulation for solving the timetable rescheduling problem modelling each route through track circuits. In this way, the control area is represented realistically. Furthermore, this solution enables to select, among all possible routes, only the ones which can be practically exploited and it considers all possible train orderings.

As can be seen, microscopic procedures consent the analysis of restricted areas in order to provide feasible plans during real time operations but rarely deal with large networks. For this reason, other works focused on the possibility to adopt a macroscopic representation for reaching this task. Two different kinds of analysis were developed: one considering exclusively train delays, the other one more oriented towards passenger needs.

Concerning the first typology, Törnquist and Persson (2007), presented a mixed linear integer programming model for re-scheduling railway traffic in a geographically large and fine-grained railway network with highly interacting traffic. The model is composed of the following variables: start and end times of an event as well as the delay of an event are represented by continuous variables while binary variables are used to express whether an event uses a track, or to decide the order of trains. Fixed headway times between trains and fixed running times along segments between stations are also considered. The goal is to minimise a cost function based on train delays. Although it is a macroscopic approach, the complete model may require a high computational effort. Moreover, since the sequence of trains on the tracks as specified by the initial timetable will mostly remain the same, only a few modifications may be necessary to achieve a significant improvement. Therefore, the solution evaluates four different strategies which mainly consist of maintaining the train order with some changes.

Due to these difficulties, in Törnquist (2012) following the same structure of the previous work, the author proposed a heuristic procedure to increase the speed of the model. In particular, given a computational time, the algorithm first performs a depth search of a good and feasible solution and builds up a tree with possible other actions which can be adopted. Then, considering the time left, the algorithm try to improve the solution obtained exploring the tree. Applications on a real network in Sweden highlighted the benefits of this approach which, according to the author, can be further improved.

Acuna-Agost et al. (2011a) formulated a similar mixed integer linear programming extending the model presented by Törnquist and Persson (2007) with two main contributions. First of all, in this case travel times between stops is not consider fixed but takes into account also acceleration and braking phases; secondly, it is now possible to admit more than one train in the same section running in the same direction. To limit the time required to reach an optimal solution, the authors decided to limit the search space around the original non-disrupted timetable. The method is tested on instances based on railway networks in France and Chile. Results showed that solutions with an average optimality gap of less than 1% may be obtained in less than 5 minutes of computation time. In a following paper (Acuna-Agost et al., 2011b), the authors reduced the calculation time introducing a new approach called SAPI (Statistical Analysis of Propagation of Incidents). This method is based on the estimation of the probability that an event in the railway network is affected by a disturbances and reduces the search space of the solution accordingly. The tests performed on the same networks in France and Chile demonstrated that solutions with an average optimality gap of 0.5% can be found in about two minutes.

Kecman et al. (2013) proposed four different macroscopic models based on different level of details for solving the rescheduling problem in the case of large networks. Basically, the authors adopted a Time Event Graph (TEG) to describe rail operations. This is a representation of a discrete-event dynamic system consisting of events connected by processes that are described by the minimum process times. In a TEG an individual train run is hence modelled as a series of events and processes that connect them. Every node is an event, defined by the train number, the timetable point, type (departure, arrival or through) and the scheduled event time. Every arc is a process, defined by the train number, type (run or dwell), start and completion event, and the minimum process time. Interactions between trains are modelled with headways or connection processes. A TEG can be represented by a max-plus linear system as shown by Goverde (2007; 2010). The TEG is then converted to four simplified formulation of the alternative graph presented by D'Ariano (2008) so as to apply the algorithm previously described in D'Ariano et al. (2007a). The different macroscopic models were then tested on the whole Dutch railway service showing the possibility to deal with national rail networks in reasonable time even with the most complex macroscopic models.

The abovementioned procedures focus on train delays and punctuality neglecting passenger satisfaction during perturbed service conditions. Recently, this topic has been largely discussed in the literature especially by means of macro-optimisation models.

Schöbel (2007) for instance, studied the delay management problem which consists of deciding if connecting trains in a station should wait or not in case feeder trains are delayed. The objective is to decide what connections can be maintained minimising passenger delay. To this purpose, the author proposed a path-oriented mixed integer programming model and an activity based mixed integer programming model. The former has a binary variable for each path, used for deciding whether all connections on the path are maintained or not. Two sets of constraints guarantee that the delay at the start of an activity is transferred to its end, where it can be reduced by the slack time of the activity (i.e. the time that can be saved by performing the activity as rapidly as possible). Other constraints are used to satisfy passenger time requirements in case of a changing of the timetable. This formulation brings to a quadratic objective function which expresses the total passenger delay and can be linearized. The activity-based model is equivalent to the previous one since both provide the same solution. In particular, the activity based model contains nodes for all arrival and departure events, and a set of arcs, called activities, corresponding to waiting or driving of the trains, and to changing of the passengers from one train to another at a station. A timetable is hence obtained assigning each event a time so that the minimal duration for performing each activity could be satisfied. In this case, the model provides a cubic objective function which can be linearized, and presents a larger number of variables and constraints. However, it has the advantage that a branch-and-bound algorithm can be easily implemented with branching on arc variables. In Schöbel (2009) and in Schachtebeck and Schöbel (2010), the complexity of the model was increased including also constraints on the limited capacity of the tracks. A branch-and-bound algorithm and several heuristic approaches were therefore developed in order to solve this new delay management problem.

Dollevoet et al. (2012) on the basis of the activity model presented in Schöbel (2007), extended the delay management problem with rerouting of passengers. In other words, it is assumed that passengers are aware of the connections which will be maintained in the near future. They can therefore decide to take an alternative route minimising their total travel time. This new aspect is taken into account in the model by using binary variables which express whether a connection is used by passengers in the path between a given origin–destination pair.

Kumazawa et al. (2010) presented a rescheduling algorithm whose aim is to minimise passenger inconvenience. The procedure consists of two main parts: the first one creates the train plan by changing the arrival and departure times of the trains, while the other one evaluates the plan based on an estimation and simulation of passenger behaviour. In addition to a conventional passenger flow analysis, the passenger overflow, defined as the waiting time experienced by a passenger while waiting on a platform because of the capacity limit of trains, is considered. In this way, the algorithm provides rescheduling solutions considering a more realistic analysis of passenger flow.

Kanai et al. (2011) developed a model for optimal delay management trying to minimise dissatisfaction of all passengers in the whole railway network. The procedure is composed of a macroscopic simulation part and an optimisation part. The simulation part is a train traffic simulator which works in parallel with a passenger flow simulator. The first one forecasts future train diagrams considering the dynamic interaction between trains and passengers. The second one reproduces passenger behaviour on the platform at each station. Number of boarding and alighting customers is also calculated as well as the dwell time values necessary to complete this process which are then transferred to the traffic simulator. The optimisation part is based on a '*tabu*' search algorithm which is able to find good strategies for the management of train connections in case of delays. Test cases performed on Japanese railway network showed the effectiveness of the proposed approach.

In case of disruptions, the train timetable rescheduling problem is more complex to solve. Different papers in the literature focus on this particular task of the railway management. Also in this case, it is possible to distinguish between models which consider the infrastructure at a microscopic level and models which consider it at a macroscopic level.

Wiklund (2007) studied the effect of different recovery strategies in case of a disruption adopting the Railsys microscopic simulation tool. In particular, the author reproduced a specific real case study, namely a case in June 2000 where a fire caused excessive damage to the interlocking system at the Järna station, in the southwest of Stockholm. The effectiveness of five strategies were compared in terms of train traffic mileage and propagated delays. This application demonstrated the usefulness of micro-simulation for testing recovery strategies in such complicated scenario.

Hirai et al. (2009) considered the problem of train stop deployment in case of blockage of the line for a long time, when decisions have to be made about where the obstructed trains should stop so that the unobstructed trains can still reach their destinations in other parts of the network. The network is represented through a Petri net model which enables to consider rail infrastructure with a high level of details and to model the transitions between potential stop locations for the trains. These do not necessarily include only platforms in stations but also track segments outside stations. The problem is formulated as mixed integer programming model whose purpose is that of minimising the number of stops outside stations and the deviations from the original timetable.

Corman et al. (2011a) worked on the disruption problem in large and busy railway networks adopting a centralised and a distributed approach. In the centralised approach the entire rescheduling problem is solved while the distributed approach is presented to manage effectively larger networks, in which a coordinator sets constraints between different areas and delegates the scheduling decisions to local schedulers. Computational experiments on a large network in the Netherlands showed that, for a time horizon up to one hour, both methods compute good quality solutions in a very short time of computation, with the distributed approach resulting in a better feasibility performance than the centralised ones. Increasing the time horizon provided scenarios too difficult to solve and required high computational time.

Contrary to what just discussed with microscopic models, macroscopic models allow tackling disrupted service events in real time.

Shimizu (2008) for instance, presented a constraint programming approach for real-time reordering of trains in a case study involving the Shinkansen railway network in the north of Tokyo in Japan. The simulated scenario reproduced the effects of an earthquake which caused the closure of a section between two major stations, affecting the entire railway network. The model considers changing of train orders in order to minimise delays and provides feasible solutions in less than 15 seconds.

Nakamura et al. (2011) proposed an algorithm for train rescheduling during disruptions taking into account three pre-determined factors: input train groups, train cancelation sections, and return patterns. Input train groups are sets of trains which share the same assignment of rolling stock. Train cancelation sections are sections of the rail infrastructure within two stations in which all trains are cancelled if a disruption occurs inside the section. Finally, a return pattern defines the connections between trains in the same group at stations that bound a disrupted section. In case a section is blocked due to a disruption, the algorithm computes a new timetable by cancelling trains, combining return patterns, and changing the train departure order at stations in a series of steps. The effectiveness of the rescheduling plan is evaluated in terms of passenger dissatisfaction caused by propagated delays. The algorithm was tested on a railway line in a metropolitan area in Japan highlighting the possibility to use the algorithm for real time applications since, during the applications, results were obtained in few minutes.

Albrecht et al. (2013) analysed the problem of disruptions due to track maintenance. In particular, the paper discussed how a problem space search meta-heuristic can be used to create quality timetables considering both train movements and scheduled track maintenance simultaneously. Furthermore, the heuristic procedure was used as an operational tool for generating revised timetables according to the new state of a disrupted system.

Canca et al. (2014) developed an optimisation model for introducing short-turning shuttle operations after the occurrence of disruptions while maintaining the timetable of previously programmed services. The aim is to increase the frequency in particular critic sections where travel demand levels have become so high that cannot be served with the planned timetable. In particular, by means of mixed integer linear optimisation model, turn-back points location, departure and arrival times and short-turning offsets are taken into account. Experiments on a real line of the Madrid commuter railway system showed the benefits of short-turning policies in mitigating the increase in average waiting time resulting from an increased demand.

Louwerse and Huisman (2014) developed a model for designing alternative periodic timetables when disruptions prevent the planned service from being performed. The author focused on some regularity aspects which increase the feasibility of the recovery solutions. For example, in case of a partial blockade of track, the model aims at operating approximately the same numbers of trains in each direction. Furthermore, another constraint consists of keeping both intercity and regional trains connection so as to carry the greatest number of passengers. In this way, the model implicitly takes into account rolling stock and crew feasibility. Computational experiments based on instances of Netherlands railways indicated that this method leads to less cancelations of trains than the contingency plans that are currently used in practice by dispatchers.

The rolling stock and the crew allocation is a difficult problem to face after disruption events. Indeed, due to the cancelation of a number of trips, several rolling stock units may not be able to carry out certain tasks in their duties. Therefore, the rolling stock needs to be rescheduled, using the updated timetable and the original rolling stock allocation as input. To this purpose, it is worth mentioning the work carried out by Cadarso et al (2013). In this paper, the authors combined an integrated optimisation model (for the timetable and rolling stock) with a model for the passengers' behaviour. The first one is formulated as a mixed integer programming model where the objective function includes costs related to the operation of train services and empty movements, the number of unattended passengers, and the allocation of additional rolling stock. The second one by contrast, simulates the dynamic behaviour of the passengers due to the disruption by means of a multinomial logit model. In this way, the model adjusts the timetable and the rolling stock assignment considering explicitly passengers' reaction to the disruption. The proposed approach first computes the anticipated passenger demand and then solve jointly the timetabling and rolling stock scheduling problem. After the new timetable is computed, the two approaches are embed in an iterative framework so as to converge to an equilibrium between offered service and travel demand. The model was tested on instances from RENFE's '*Cercanías Madrid*' obtaining feasible solutions to the disrupted scenarios in few minutes.

In conclusion, in this chapter it has been discussed the state of art of rail simulation analysing different models which differ each other according to the network scale (i.e. macroscopic – microscopic – mesoscopic), the processing technique (i.e. synchronous – asynchronous) and the assumptions on input variables (i.e. deterministic – stochastic). In addition, an assessment of dispatching and rescheduling models is also provided and numerous procedures proposed in the literature for managing the rail systems after the occurrence of disturbances or disruptions are illustrated.

In the following chapter, the limits of these models are summarised so as to highlight the key elements which have inspired this thesis.

Chapter 4: The definition of a model framework for the planning and the management phases of the rail system in any kind of service condition.

In the previous chapter, different simulation approaches have been described showing possible applications in railway contexts. In particular, it has been analysed how simulation models can be used to solve problems concerning the design or the management of the rail service. Furthermore, an in-depth analysis of the procedures for tackling the rescheduling problem has been presented. In this context, two different approaches can be distinguished: the first one is based on procedures which provide feasible solutions in real time while the second one requires more computational time and therefore, can be just used as ‘off-line’ supporting tools. It is evident that the main differences between the two methods are related to the accuracy of the models as well as to the efficiency of the calculation methods. Although a more precise estimation of network performance can be achieved by means of microscopic models, these ones are not largely used except from limited applications inasmuch as the amount of data involved in the simulation makes this approach not suitable when promptness in finding feasible solutions is required. The alternative graph formulation is a particular case. As already explained, the procedures based on this representation are ‘fixed-speed’ microscopic models which consider only undisturbed train running times and this assumption supports the development of optimisation procedures. For this reason, less detailed models (e.g. macroscopic, mesoscopic and fixed speed microscopic models) are generally preferred for real time management of the network. However, these approaches require some approximations which are not admissible when congestion levels within the network increase. In addition, travel demand is often neglected since it is considered that increasing punctuality is always the best solution to achieve even from passengers’ point of view. This coincidence is not always proved. Indeed (see for instance Quaglietta et al., 2011a; D’Acierno et al., 2012), neglecting travel demand can bring to solutions which optimise operational aspects but reduce the service quality (see paragraph

2.1.9) perceived by customers. As already shown, in the literature there are some works which consider passengers' needs (Cadarso et al., 2013; Schöbel, 2007; Schachtebeck and Schöbel, 2010) but they are mainly based on macro-optimisation procedures and hence, travel demand is not simulated realistically. Likewise, just few examples of microscopic simulation approaches which take into account also passenger flow have been developed. One of the most complete is the model proposed by Kunimatsu et al. (2012) which evaluates different timetable configurations from users' viewpoint combining the micro-simulation of train operation and passenger flow.

The objective of this thesis is to propose a Decision Support System (DSS) for planning or managing the rail network in any kind of service conditions but focusing on disruption events. In particular, the model simulates the whole rail system microscopically assigning passengers to the network and analysing their influences on the service. This process causes dynamic interactions which have to be estimated so as to reproduce network conditions as closely as possible to the reality. In this way, the procedure provides the possibility to look for the intervention strategies which optimise passenger satisfaction.
Due to the complexity of the problem, the whole procedure is not feasible to obtain results in short time and it is therefore proposed as an 'off-line' methodology.

The chapter is organised as follows. First of all, the general architecture of the approach is described in detail. Then, each of the models involved in the procedure is further analysed providing more information about the analytic formulations and their modes of operation. Finally, in order to explain how to solve fixed point problems resulting from the interactions between travel demand and transportation performance, the theory and the resolution techniques of these particular complex problems are described.

4.1 Framework of the proposed approach.

The problem of identifying the optimal intervention strategy to adopt in the rail system in case of disruptions can be viewed as a bi-level multidimensional constrained optimisation problem, whose analytic formulation is the following:

$$\hat{\boldsymbol{y}} = \underset{\boldsymbol{y} \in \boldsymbol{S}_y}{arg\ min}\, Z(\boldsymbol{y}, \boldsymbol{fc}, \boldsymbol{tnp}, \boldsymbol{rnp}, \boldsymbol{td}) \tag{4.1}$$

Subject to:

$$[\boldsymbol{tnp}, \boldsymbol{rnp}, \boldsymbol{td}]^T == \Lambda(\boldsymbol{y}, \boldsymbol{fc}, \boldsymbol{tnp}, \boldsymbol{rnp}, \boldsymbol{td}, \boldsymbol{in}^0, \boldsymbol{rs}^0, \boldsymbol{ss}^0, \boldsymbol{pt}) \tag{4.2}$$

with:

$$\begin{aligned} &Z(\boldsymbol{y}, \boldsymbol{fc}, \boldsymbol{tnp}, \boldsymbol{rnp}, \boldsymbol{td}) = \\ &= \sum_i \beta^i_{VOT} \cdot \left(\beta^i_{waiting} \cdot \sum_s \sum_p \sum_r tw^{i,r}_{s,p}(\boldsymbol{y}, \boldsymbol{fc}, \boldsymbol{tnp}, \boldsymbol{rnp}) \cdot fw^{i,r}_{s,p}(\boldsymbol{td}) + \right. \\ &\qquad \left. + \beta^i_{on-board}(\boldsymbol{td}) \cdot \sum_l \sum_r tb^{i,r}_l(\boldsymbol{y}, \boldsymbol{fc}, \boldsymbol{tnp}, \boldsymbol{rnp}) \cdot fb^{i,r}_l(\boldsymbol{td}) \right) \end{aligned} \tag{4.3}$$

where $\boldsymbol{y}$ is the vector of parameters which identifies the intervention strategy; $\hat{\boldsymbol{y}}$ is the optimal value of vector $\boldsymbol{y}$; S_y is the feasibility set of vector $\boldsymbol{y}$ (i.e. the set identifying all feasible operational strategies); Z is the objective function to be minimised; $\boldsymbol{fc}$ is the vector of parameters identifying the failure context; $\boldsymbol{tnp}$ is the vector of parameters identifying the transportation network performance; $\boldsymbol{rnp}$ is the vector of parameters describing network performance of the rail system; $\boldsymbol{td}$ is the vector of parameters characterising travel demand; Λ is the simulation function; $\boldsymbol{in}^0$ is the vector defining rail infrastructure in non-perturbed conditions; $\boldsymbol{rs}^0$ is the vector describing rolling stock in non-perturbed conditions; $\boldsymbol{ss}^0$ is the vector representing the signalling system in non-perturbed conditions; $\boldsymbol{pt}$ is the vector reproducing the planned timetable; $\beta^i_{waiting}$ is a parameter which expresses the relevance (i.e. relative weight) given by users belonging to category i to waiting times; $tw^{i,r}_{s,p}$ is the average user waiting time of user category i at station s, on platform p between run (r–1) and run r; $fw^{i,r}_{s,p}$ is the number of passenger of user category i waiting at station s, on platform p between run (r–1) and run r; $\beta^i_{on-board}(\boldsymbol{td})$ is a parameter which expresses the relevance (i.e. relative weight) given by users belonging to category i to on-board time and depends on the crowding level within the coach; $tb_l^{i,r}$ is the time spent by user of category i on board the rail convoy associated to run r for travelling on link l; $fb_l^{i,r}$ is the number of passengers belonging to category i who travels on the rail convoy associated to run r while crossing link l; β_{VOT}^{i} expresses, for each user category i, the amount of money people are willing to spend for saving one hour of travel time.

In other words, the problem consists of finding the optimal solution within the set of all possible strategies, minimising the user generalised cost (objective function described by formula 4.3) and considering any single element of the rail network. The complexity of the problem is highlighted by formula (4.2) which is the kernel of the proposed procedure. It can be viewed as a consistency constraint between transportation performance and travel demand flow, whose formulation requires the adoption of four different models:

- Failure Model;
- Service Simulation Model;
- Supply Model;
- Travel Demand Model.

The Failure Model (*FM*) estimates the probability of failure for any element of the network and calculates the effects of these failures on the rail system. Depending on the breakdown typology, the effects can be different. Some examples are the unavailability of a rail convoy, the interruption of a track section or the reduction of performance of the whole line or part of it.

The Service Simulation Model (*SeSM*) is nothing but a microscopic synchronous simulation model which simulates rail network performance. According to the targets of the analysis, this model can perform both deterministic or stochastic simulation.

The Supply Model (*SM*) estimates the performance of all transportation systems within the study area, including the rail system.

The Travel Demand Model is the most innovative part of the procedure. The aim is to simulate microscopically also passenger flow and its interaction with the service. To this purpose, the Travel Demand Model is split into other two sub-models (see Figure 4.1): a Pre-Platform Model (*PPM*) and an On-Platform Model (*OPM*). The former is dedicated to the estimation of passenger flow entering the rail system and hence going to the platform. Obviously, this process is the result of a decision made by users considering the performance of all available means of transport (i.e. through the interaction with the Supply Model). The latter by contrast, considering the capacity constraints of each train approaching a station, evaluates the number of boarding and alighting passengers as well as the time required to complete this process (i.e. dwell time estimation). In addition, the model gives the possibility to consider crowding levels within each coach of the train and, in case the maximum capacity of the train is reached, to estimate the number of passengers who are forced to stay on the platform waiting for following trains.

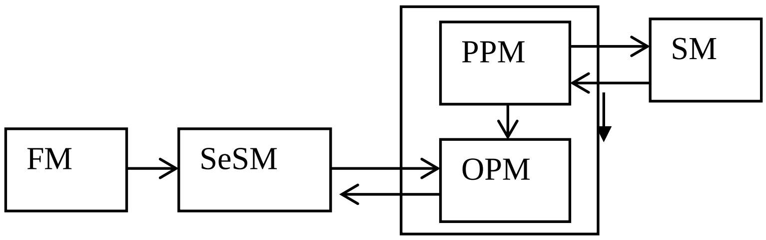


Figure 4.1 Framework of the proposed decision support system

Considering explicitly the four models and their interaction, relation (4.2) can be rewritten as:

$$\begin{cases} \mathbf{tnp} = \mathbf{SM}(\mathbf{rnp}, \mathbf{td}) \\ \mathbf{rnp} = \mathbf{SeSM}(\mathbf{y}, \mathbf{FM}(\mathbf{fc}, \mathbf{in}^0, \mathbf{rs}^0, \mathbf{ss}^0), \mathbf{td}, \mathbf{pt}) \\ \mathbf{td} = \mathbf{OPM}(\mathbf{rnp}, \mathbf{PPM}(\mathbf{tnp}, \mathbf{rnp}), \mathbf{FM}(\mathbf{fc}, \mathbf{in}^0, \mathbf{rs}^0, \mathbf{ss}^0)) \end{cases} \tag{4.4}$$

In order to obtain the objective function values, the following procedure has to be performed:

1. First of all, once estimated the probability of breakdown, the Failure Model provides the input data of the simulation so as to reproduce the failure scenario.
2. After setting up all information, the Service Simulation Model performs the simulation of the failure scenario.
3. According to the outputs of the simulation, an estimation of every possible intervention strategy (including the non-intervention) or a subset of them (obtained by means of a suitable algorithm) is carried out. Afterwards, each selected solution is simulated.
4. For each simulated scenario, the number of users arriving on the platform, which is the effect of user individual choices (Pre-Platform Model), has to be determined as a function of all transportation systems performances (Supply Model), including the rail system (Service Simulation Model).
5. Finally, the behaviour of passenger flow on the platform and its influence on the service are evaluated. These processes are strongly related to the performances of the rail system (e.g. frequency of the line) and to rolling stock features (e.g. vehicle capacities, number of doors per coach etc.)
6. After simulating the number of passengers boarding/alighting the trains together with the rail service, information about user trips (e.g. waiting times, travel times, crowding levels experienced within the train, platform congestion) is known. Therefore, for each intervention strategy, the objective function values (4.3) can be evaluated.

This process is based on the deterministic simulation of the network. Obviously, reality is far from this assumption which means that the strategies adopted might not be robust enough to ensure effectiveness of the intervention. In order to provide a sensitivity analysis of the recovery strategies, the previous methodology needs to be increased of a stochastic phase.

To this aim, parameters of relation (4.2) are considered random variables which can be indicated as follows:

$$\boldsymbol{X} = \overline{\boldsymbol{X}} + \boldsymbol{\varepsilon}_X \tag{4.5}$$

With:

$$\overline{\boldsymbol{X}} = E[\boldsymbol{X}] \tag{4.6}$$

$$\varepsilon_X \sim \boldsymbol{\Omega}_X(\boldsymbol{h}_X) \tag{4.7}$$

where $\boldsymbol{X}$ is the considered multivariate random variable (i.e. a vector of random variables); $\overline{\boldsymbol{X}}$ is the fixed vector whose elements are the mathematical expectations (i.e. first moments or means) of elements of $\boldsymbol{X}$; ε_X is the random residual of $\boldsymbol{X}$; $\boldsymbol{\Omega}_X(\cdot)$ is the statistical distribution of $\boldsymbol{\varepsilon}_X$; $\boldsymbol{h}_X$ is the vector of parameters of statistical distribution $\boldsymbol{\Omega}_X(\cdot)$. Thus, relation (4.2) becomes:

$$\begin{aligned}&\left[\boldsymbol{y}, \boldsymbol{fc}, \boldsymbol{tnp}, \boldsymbol{rnp}, \boldsymbol{td}, \boldsymbol{in}^0, \boldsymbol{rs}^0, \boldsymbol{ss}^0, \boldsymbol{pt}\right]^T = \\ &= \left[\overline{\boldsymbol{y}}, \overline{\boldsymbol{fc}}, \overline{\boldsymbol{tnp}}, \overline{\boldsymbol{rnp}}, \overline{\boldsymbol{td}}, \overline{\boldsymbol{in}}^0, \overline{\boldsymbol{rs}}^0, \overline{\boldsymbol{ss}}^0, \overline{\boldsymbol{pt}}\right]^T + \\ &+ \left[\boldsymbol{\varepsilon}_y, \boldsymbol{\varepsilon}_{fc}, \boldsymbol{\varepsilon}_{tnp}, \boldsymbol{\varepsilon}_{rnp}, \boldsymbol{\varepsilon}_{td}, \boldsymbol{\varepsilon}_{in^0}, \boldsymbol{\varepsilon}_{rs^0}, \boldsymbol{\varepsilon}_{ss^0}, \boldsymbol{\varepsilon}_{pt}\right]^T\end{aligned} \tag{4.8}$$

Therefore, including also the robustness assessment of recovery solutions into account, the complete procedure can be summarised as follows:

- for each failure context, the deterministic optimisation problem (i.e. problem 4.1 subject to 4.2 and 4.3) is implemented in order to obtain the optimal intervention strategy (i.e. $\hat{\boldsymbol{y}}$);
- a neighbourhood of $\hat{\boldsymbol{y}}$, indicated as $N(\hat{\boldsymbol{y}})$, which consists of all corrective actions providing objective function values close to the minimum cost (i.e. objective functions calculated in the case of strategy $\hat{\boldsymbol{y}}$) is analysed;
- *n* vectors describing random residual ε_X are extracted;
- for each single extracted vector $\boldsymbol{\varepsilon}_X$, the new objective function values for all the intervention strategies of set $N(\hat{\boldsymbol{y}})$ are calculated. Obviously,

calculation of the objective function requires the solution of problem (4.4);

- finally, the distribution of the objective function values for each element of set $N(\hat{\mathbf{y}})$ is analysed for performing the stability analysis of the recovery solutions.

Before providing more details, it is worth adding some comments about the procedure in order to highlight the benefits of the proposed framework.

As already said, the resolution of the problem is quite complex and requires long computational times. This can be viewed as a drawback, since other examples in the literature focused on the possibility to obtain results in few seconds. However, this methodology provides precise outputs and does not approximate the simulation of the rail system. Hence, it can be used to solve also complex disrupted scenarios involving large networks and find reliable solutions which increase the service quality.

Another remark on the procedure is related to the simulation of possible failure scenarios. Real time approaches, as the ones showed in the previous chapter, analyse disrupted events immediately after receiving information about the service from the network. Thus, the failure scenarios are based on what is happening in the network in the same moment. Off-line procedures generally work on failure scenarios which are similar to critic events happened in the past (see for instance Wiklund, 2007). The proposed procedure instead, adopts the Failure Model in order to analyse possible breakdowns which can affect the network. This is a great advantage. Indeed, thanks to a RAMS analysis (see the following paragraph for more details), it is possible to investigate weaknesses of the infrastructure, the rolling stock and the signalling system and to simulate the events with higher probability of occurrences. Once examined all possible breakdowns and run the procedure several times, the intervention strategies for any possible event can be selected. Based on these analyses, dispatchers will be able to react promptly to any occurrence increasing user satisfaction.

Further comments concern the analysis of possible recovery solutions. As previously stated, the methodology could be based on the adoption of an exhaustive approach which aims at assessing the effects of all feasible intervention strategies. Obviously, this procedure is convenient when the number of scenarios which have to be investigated is limited providing a complete and precise analyses of the network in any kind of service conditions. When the set of solutions is huge instead, it is necessary to implement optimisation algorithms in order to limit the number of simulations. This task is dealt with in the last part of this thesis where the combination of a macro-optimisation model with the abovementioned procedure is shown.
In addition, by means of several stochastic simulations, sensitivity analysis of all possible intervention strategies can be carried out giving important indications about the robustness and the reliability of the solutions provided. This method overcomes the limits of previous rescheduling models which are not able to take into account the dynamic evolution of dispatching strategies due to the randomness of the events (Quaglietta et al., 2013).
Finally, the microscopic simulation of the network together with that of travel demand flows needs to be highlighted. In fact, this approach enables the dynamic assignment of passenger flows to the service providing two important results. First of all, the procedure allows the evaluation of reliable strategies which fulfil also customers' needs. In other words, the microscopic simulation of the network guarantees that the solution found can be applied in a real context while the assignment makes sure that quality perceived by customers is considered. Service quality is estimated through the objective function ***Z*** which, as already said, expresses on average the total generalised cost perceived by customers of the rail service. In particular, this function specifies the users' disutility during the travel in terms of costs and level of service. Generally, the user generalised cost is the sum of several performance attributes or variables (Cascetta, 2009). In this case, only two attributes (i.e. waiting and travel times) are considered. In fact, these two variables are the only ones affected by variations during the simulation of the different

scenarios. Monetary costs and other attributes associated to the rail system remain constant and they are therefore neglected. However, it is worth noting that in the analytical relation (4.3), the $\beta_{on-board}$ parameter is not considered constant but it is function of the rail crowding. In this way, although implicitly, the comfort perceived on-board is also taken into account.

Another benefit of the procedure is the possibility of estimating the 'answer' of the system to different demand profiles during the day. In fact, adopting a within-day dynamic (or intra-period dynamic, see Cascetta, 2009 for more details), it is possible to simulate demand peaks, temporary capacity variations, temporary over-saturation of supply elements, and formation and dispersion of queues. As a consequence, the proposed models can be useful not only for the management of perturbed service conditions, but also for the planning of timetable and rolling stock according to customer requirements. In addition, the design of customer information about the service conditions (e.g. Intelligent Transportation Systems – ITS) as well as new supervision or management procedures of travel demand could be implemented. However, in order to perform the dynamic assignment, it is first necessary to estimate rail passenger flows as function of time *t* or, in other words, the arrival rate at stations. These data are often unknown by rail operators. Although electronic ticketing services are becoming more and more popular, only if the validation process is mandatory also in the exit from the system, this process is able to provide precise information about passenger flows. Therefore, in the majority of the cases, it is necessary to estimate travel demand by means of mathematical models for obtaining current demand profiles or foreseeing future demand patterns (due for instance to a variation in service quality or performance). To this purpose, the proposed procedure requires the split of Travel Demand Model into two levels. The Pre-Platform Model estimates only the amount of flow concerning the rail system as a result of a multi-modal assignment process involving all transportation systems within the study area. Although this method implies the increasing of the problem complexity, it enlarges the applications of the proposed framework enabling the evaluations of the

interactions among the different transport modes. Indeed, especially in urban contexts, the railway system is part of the public transportation network and cannot be considered individually.
After this phase, the users' on-platform choices and their influence on the service are simulated by the On-Platform Demand Model concluding the assignment process.

Starting from the following paragraph, each model will be described in detail providing the analytical formulation and the mode of operation.

4.2 The Failure Simulation Model.

The Failure Simulation Model estimates the probability of failure occurrences and their effects on the rail network. Analytically, it consists of a function, indicated as $\boldsymbol{FM}$, which provides parameters describing infrastructure ($\boldsymbol{in}$), rolling stock ($\boldsymbol{rs}$) and signalling system ($\boldsymbol{ss}$) depending on their non-perturbed values ($\boldsymbol{in^0}$, $\boldsymbol{rs^0}$, $\boldsymbol{ss^0}$) and failure context ($\boldsymbol{fc}$):

$$[\boldsymbol{in}, \boldsymbol{rs}, \boldsymbol{ss}]^T = \boldsymbol{FM}\left(\boldsymbol{in}^0, \boldsymbol{rs}^0, \boldsymbol{ss}^0, \boldsymbol{fc}\right) \tag{4.9}$$

Basically, the procedure requires the application of Reliability, Availability, Maintainability and Safety (RAMS) techniques as indicated by the norm EN 50126 (CENELEC, 1999). This norm defines procedures for railway companies, the rail industry and its suppliers within the European Union to implement a management system for reliability, availability, maintainability and safety. The adoption of a RAMS analysis (as will be stated in the following paragraph) provides the prediction, at any life cycle step and for each component of the system, of the expected failure rate and its effects on the system in case of occurrence. It is therefore useful for many aspects, such as:

- Evaluation of reliability and robustness of future system design.
- Identification of parts of the system which are likely to have the major impacts on system level failure, and also which failure modes to expect and which risks they pose to the users, clients, or society.
- Planning of cost-effective maintenance and replacement operations.

- Reducing the probability of hazards and accidents.
- Assessment of possible investments to improve the system.

Hence, the Failure Model not only assesses the breakdown contexts which are worth simulating but gives also indications about network conditions in the case of degraded service operations.

More details about RAMS are described in the following paragraph.

4.2.1 RAMS analysis.

In any engineering field, since the beginning of the industrial age, engineers have struggled to create reliable and durable equipment and systems. In that period, developments and improvements in the design process were due to the application of trial-and-error procedures. Nowadays, since the cost and risk of possible failures has considerably increased, it is more and more important to assess failure and risk and try to make predictions on these as early as in the design step. To this purpose, RAMS analysis studies the behaviour of a new system, equipment or design improvement in order to assess failure modes and their causes.

In order to understand the procedures and techniques involved in a RAMS analysis, it is worth focusing on the single terms composing the abbreviation. As already said, RAMS states for Reliability, Availability, Maintainability and Safety. Frequently, the abbreviation adopted is RAMS(S), including also the Security as parameter to consider. The difference between safety and security has to be defined. Safety means the functional safety within the system and protection against hazardous consequences caused by technical failure and unintended human mistakes. Security, by contrast, is the protection against hazardous consequences due to wilful and unreasonable human actions. The majority of the components in railway systems are safety related. However, failures can be caused also by security reasons (e.g. copper thieves) and therefore, measures to protect each component of the network have to be considered.

Reliability, availability and maintainability are strongly related. According to EN 50126 (CENELEC, 1999), the term availability is defined as:

'The ability of a product to be in a state to perform a required function under given conditions at a given instant of time, or over a given time interval, assuming that the required external sources of help are provided.'

In other words, the system (called 'product') will fulfil the required tasks (called 'functions') under the defined framework conditions. In railway contexts, the main function is the safe transport of persons and goods. The required external sources of help are the technical components of the system (e.g. signalling system, track clear detection etc.) and the railway staff in undertaking their tasks.

Here comes the importance of reliability in achieving availability. Indeed, it is defined as (IEC 2001):

'The probability that an item can perform a required function under given conditions for a given time interval (t1, t2)'.

This results in the requirement of failure-free working of the components during a specified time period. Obviously, to achieve this task, maintainability is another factor to take into account. EN 50126 defines it as:

'The probability that a given active maintenance action, for an item under given conditions of use, can be carried out within a stated time interval when the maintenance is performed under stated conditions and using stated procedures and resources'.

Reliability and maintainability are both probability values related to a defined time period. The former leads to failure rates while the latter leads to maintenance rates. Both components influence availability which is an important requirement of the railway system. This is strongly related to safety inasmuch as the more available a technical system is, the lower is the probability to operate in degraded mode. Clearly, this increases safety. However, as risk can never be zero, safety can never be perfect. The railway system has to run in the zone of safety but a certain level of remaining risk cannot be avoided. In particular, the risk can be defined as the product of hazard rate multiplied by damage (Theeg and Vlasenko, 2009). Since damage in railway accident counts in most cases as ‘high’, risk can only be reduced by lowering the hazard rate. Once again, this will be achieved by high availability which is responsible for the safety of the railway system.

This brief introduction has explained the importance and the meaning of RAMS but how can it be implemented?

Basically, RAMS analysis is performed according to a cycle composed of three steps (Medeiros, 2008): data compilation, simulation and impact on system life cycle. Failure data compilation is the basis of any RAMS simulation or process. This phase gives the possibility to obtain data on failure rate and other reliability parameters which are the input for the simulation. In this second step, the objective is to model the system in terms of reliability aspects. In other words, the simulation, given the values of failure rate and reliability of each single component, estimates causes and effects of failures of the same components but while interacting with each other in a system and a stated environment. The main simulation methods can be divided in qualitative and quantitative approach as follows (Medeiros, 2008):

- Quantitative approach:
    - Preliminary Hazard Analysis (PHA),
    - Failure Modes and Effects Analysis (FMEA).
- Quantitative approach:
    - Success Diagram Method (SDM),
    - Cause Tree Method (CTM),
    - Truth Table Method (TTM),
    - Gathered Fault Combination Method (GFCM),
    - Consequence Tree Method (CQTM),
    - Cause-Consequence Diagram Method (CCDM).

Examples of RAMS software analysis packages are ‘RAM Commander’, which is used for project (system) level analysis and ‘Weibull++’, which is used for the statistical studies of components with non-constant failure rates.

Finally, the last step is the life cycle assessment whose goal is to assess benefits and costs the whole system will bring. It consists of analysing the cost of breakdowns and of corrective maintenance operations as well as the cost of accidents.

As previously described, specific reliability parameters have to be determined for performing step 1 and 2. Generally, reliability and safety are represented most adequately by quantitative parameters (Theeg and Vlasenko, 2008). In particular, the failure rate is defined as follows:

- Failure Rate $\lambda(t)$: $\lambda(t) = \dfrac{n(\Delta t)}{N_m \Delta t}$ (4.10)

- Dangerous failure rate $\lambda_D(t)$: $\lambda_D(t) = \dfrac{n_D(\Delta t)}{N_{sm} \Delta t}$ (4.11)

where:

$n(\Delta t)$ and $n_D(\Delta t)$ are quantity of samples of the system having a failure (hazardous failure) in a given time (Δt);

$N_m = \dfrac{N_i + N_{i+1}}{2}$ is the mean quantity of failure free systems in a given time (Δt);

N_i, N_{i+1} are quantity of failure free systems in the time $t - \dfrac{\Delta t}{2}$ and $t + \dfrac{\Delta t}{2}$;

$N_{sm} = \dfrac{N_i + N_{i+1}}{2}$ is the mean quantity of failure-free systems not having any hazardous failures in the time interval (Δt) (on condition that samples of the system having protected failure are immediately replaced by the new samples).

Assuming a constant failure rate, which means that failures occur at approximately similar time intervals, the following equations can be considered:

$\lambda(t) = \lambda = const$ and $\lambda_D(t) = \lambda_D = const$ (4.12)

Then, on the basis of exponential law of reliability, the other parameters that can be estimated are:

- Failure-free operation probability $P(t)$: $P(t)=e^{-\lambda t}$ (4.13)
- Probability of failure $Q(t)$: $Q(t)=1-e^{-\lambda t}$ (4.14)
- Mean operating time to failure T: $T=\frac{1}{\lambda}$ (4.15)
- Probability of safety $P_S(t)$: $P_S(t)=e^{-\lambda_D t}$ (4.16)
- Probability of dangerous failure $Q_D(t)$: $Q_D(t)=1-e^{-\lambda_D t}$ (4.17)
- Mean operating time to hazardous failure T_D: $T_D=\frac{1}{\lambda_D}$ (4.18)

Once known all these parameters, the RAMS analysis can be completed and the most likely failure events can be evaluated providing their effects on the rail service. This phase completes the process of the Failure Model.

4.3 The Service Simulation Model (SeSM).

The Service Simulation Model (***SeSM***) determines rail system performance depending on rail infrastructures, rolling stock, signalling system, timetable and user flows on the network. In other words, this model is intended to calculate running times and headways of each simulated train. Under the assumption of a micro-simulation approach, the problem consists of a system of differential equations whose solution requires the adoption of numerical procedures (see paragraph 2.3.4).

Analytically, the ***SeSM*** model is described as follows:

$$\boldsymbol{rnp}=\boldsymbol{SeSM}(\boldsymbol{y},\boldsymbol{td},\boldsymbol{in},\boldsymbol{rs},\boldsymbol{ss},\boldsymbol{pt}) \tag{4.19}$$

Where, once again, ***(rnp)*** is rail network performance, ***(y)*** is the implemented strategy, ***(td)*** indicates the travel demand, ***(in)*** is the infrastructure, ***(rs)*** is the rolling stock, ***(ss)*** is the signalling system and ***(pt)*** is the planned timetable.

This is the case all events like departures, arrivals and running times are constant values established in the timetable, that is the model is deterministic. When the simulation is influenced by the realisation of random (i.e. stochastic) processes, the parameters are not constant but follow a statistical random distribution.

In the previous chapter, several examples of microscopic models have been described. Any of those programs can be useful for the implementation of the ***SeSM***. However, due to copyrights, commercial software often cannot be modified and adapted to research targets. This is a drawback since it is highly recommendable that the ***SeSM*** is an opened code language. In this way, it is possible to improve the model and simplify the interaction with other applications for performing the passenger flow assignment.

4.4 The Supply Model (SM).

The Supply Model (***SM***) is a system of models which simulates the performances and the flows resulting from users' demand and the technical aspects of the physical transportation supply (Cascetta, 2009). Specifically, the ***SM*** combines traffic flow theory and network flow theory models. The former ones analyse and simulate the performances of the main supply elements while the latter represent the topological and functional structure of the system.

In analytical terms, the ***SM*** can be viewed as:

$$\boldsymbol{tnp}(t) = \boldsymbol{SM}(\boldsymbol{td}(t)) \tag{4.20}$$

where, based on congested network flow models, transportation network performances (***tnp***) are function of travel demand flows (***td***). In this case, ***tnp*** and ***td*** refer to all transportation systems including the rail system (see Figure 4.2). Furthermore, each transport mode is simulated assuming that all relevant characteristics such as traffic flows and supply performances are not stationary, but dependent on the time instant *t* internal to the reference period (within-day dynamic formulation).

Basically, the idea is to simulate the different alternative choices within the study area modelling the dynamic nature of both flows and network performances. In fact, disruption events during the service (no matter what transportation system is affected), due to their dynamic evolution, are phenomena that cannot be analysed statically. Their effects can originate for example, temporary service interruptions and thus, demand profiles and performances can change rapidly according to the new conditions.

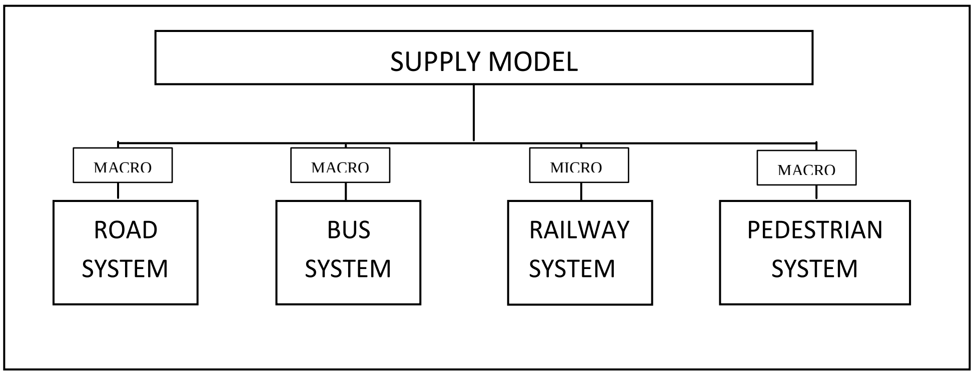


Figure 4.2 Supply Model definition.

For example, passengers waiting for a train on the platform can decide to modify their trip (adaptive choice) if no trains arrive after a long period. Likewise, if something negative happens within the road or bus systems, it is likely that the number of customers arriving at stations increases. For this reason, it is worth considering all different transportation systems since it enables a more realistic estimation of the arrival rate at stations. Obviously, this approach is useful for the study of urban contexts where public transport systems are strongly integrated. Conventional railway networks, by contrast, can be analysed as a closed system without interactions with other transport modes.

The general structure of the Supply Model concerning railway systems has already been explained in the previous chapter. Basically, it is depicted by means of a microscopic model. As far as the other public transport networks are concerned, the Supply Model is represented macroscopically and can be divided in different sub-models. First of all, the graph defines the topology of the transportation system under study. The flow propagation model describes the relations among path and link flows. The link performance model specifies the physical and functional characteristics of the transportation system considering their relationships with user flows. Finally, the path performance model, for any origin-destination pair, expresses the connections between the performances of single links and those of the whole trip.

Analytically, the abovementioned sub-models can be divided into two different groups. The first one refers to continuous service system (e.g. road system), whose relations are the following (Cascetta, 2009):

$$f = \phi[t(\tau), h(\tau)] \tag{4.21}$$

$$t(\tau) = t\left(f\left(\tau^{'}\right), \tau^{'} \le \tau\right) \tag{4.22}$$

$$TT^{f}(\tau) = \Gamma\left(t\left(\tau^{'}\right), \tau^{'} > \tau\right) \tag{4.23}$$

where $t(\tau)$ is the vector of link travel times at time τ; $TT(\tau)$ is the vector of forward path travel at times τ (i.e. the time needed to traverse a general path starting at time τ); $f(\tau)$ denotes the vector of relevant flow or occupancy input variables for travel time functions at time τ; $h(\tau)$ is the path flow vector at time τ; Γ expresses symbolically the relationship between link and path travel time, namely it describes the fact that the time to traverse a general path starting at time τ is dependent on the time to traverse all the links belonging to the path in the following instant of time; Φ is the function simulating how time-varying continuous path flows propagate through the network inducing time-varying in-flows, out-flows and link occupancies.

The second group refers to public transport services with high frequency (e.g. bus system). In this case, the supply model can be modelled adopting a diachronic graph where the departure and arrival times are random variables (indicated as β_d and β_a) whose average values correspond to the planned departure and arrival time. Obviously, these values depend on running and waiting times which are considered random variables (i.e. $\boldsymbol{y_r}$ and $\boldsymbol{y_s}$) as well.

Therefore, on the running link l between stop s-1 and s the arrival time is formulated as:

$$b_{a,rs} = b_{p,r(s-1)} + y_{r,l} \; l \equiv ((s-1),s) \tag{4.24}$$

While the departure time is:

$$b_{p,rs} = b_{a,rs} + y_{r,s} \tag{4.25}$$

Assuming that $\boldsymbol{b^t}$ is the vector obtained as a random enumeration of $\boldsymbol{b}$, G^t the relative diachronic graph and $\boldsymbol{\Delta^t}$ the incidence link-path matrix, the dynamic supply model can be formulated similar to the static formulation as (Cascetta, 2009):

$$\boldsymbol{g}^t = \boldsymbol{\Delta}^T \boldsymbol{c}^t + \boldsymbol{g}^{NA^t} \tag{4.26}$$

$$\boldsymbol{f}^t = \boldsymbol{\Delta}^t \boldsymbol{h}^t \tag{4.27}$$

where $\boldsymbol{g}^t$ is the cost path vector at time instant t; $\boldsymbol{g}^{NA^t}$ is the non-additive cost path vector at time instant t; $\boldsymbol{c}$ is the cost link vector at time instant t, $\boldsymbol{f}$ is the link flow vector at time instant t and $\boldsymbol{h}$ is the path flow vector at time instant t.

4.5 The Travel Demand Model.

A transportation Demand Model is a mathematical relationship associating the average values of demand flows with their relevant characteristics to given activity and transportation supply systems (Cascetta, 2009).

As already said, in this proposal, it is necessary to simulate two levels of choice which have to be dealt with separately.

The Pre-Platform Model (***PPM***) estimates the rail passenger flow depending on performances of all transportation systems. In analytical terms, it can be formulated as follows:

$$\boldsymbol{upf}(t) = \boldsymbol{PPM}(\boldsymbol{tnp}(t), \boldsymbol{rnp}(t)) \tag{4.28}$$

Where ***upf(t)*** states user flows arriving to the platform, ***tnp(t)*** is the abovementioned transportation network performance vector and ***rnp(t)*** specifies rail network performance. Due to the dynamic of the phenomenon, all elements of relation (4.28) are not constant but depend on the instant of time considered *t*. However, the service typology strongly influences the passenger behaviour on the platform. Indeed, in the case of metro-rail systems, due to the high frequency of the service, passengers do not know the timetable and go to the platforms waiting for the first arriving train. Hence, the number of passengers willing to board the train at each station is estimated according to the performance of the service. As a consequence, the higher the headway is, the more crowded is the platform. This assumption is not valid for long distance trains where frequencies are low. In fact, in this case, passengers generally arrive at stations some minutes before the planned departure of the trains. The arrival rate does not affect the number of boarding passengers who will be the same even if the train is delayed (until a certain value of time).

Function (4.28) is influenced by (4.20) through a fixed point problem (Cantarella, 1997; Cascetta, 2009). In fact, in the ***SM*** model disutilities and costs perceived by passengers during their travel depend on the flow levels and these, in turn, depend on the disutilities and costs experienced. However, due to the assumption of within-day dynamic, there is another dependence between flows and costs producing a double feedback loop. The first one is external and involves costs and path flows like the static case; the second one instead, is typical of the dynamic user equilibrium problem and concerns flows and link costs at the instant of time considered (see Bellei et al., 2005; Cascetta, 2009; Frederix et al., 2013; Trozzi et al., 2013).

The On-Platform Model (***OPM***) describes in detail what happens on the platform. In particular, it analyses, for each train approaching a station, whether the residual capacity is greater than the number of boarding passengers. If this condition is not satisfied, only a portion of travel demand (i.e. waiting passengers) equal to the residual capacity, is able to board the train while the surplus has to wait for the following trains. This principle is true in the case of high frequency services which are the main target of this thesis. Obviously, other assumptions have to be considered for the analysis of different kinds of rail systems, such as conventional rail lines.

As regards the analytical formulation, the On-Platform Model can be expressed through a function, indicated as ***OPM***, which provides user flows within the network (***td***) depending on user platform flows (***upf***), rolling stock (***rs***) and rail network performance (***rnp***), that is:

$$\boldsymbol{td}(t) = \boldsymbol{OPM}(\boldsymbol{rnp}(t), \boldsymbol{upf}(t), \boldsymbol{rs}(t)) \tag{4.29}$$

As stated by (4.19) (i.e. the ***SeSM***) rail network performance depends on travel demand which, as confirmed by (4.29), depends again on the service performance. Therefore, the interactions between the ***SeSM*** and the ***OPM*** generates a new fixed point problem. In other words, the number of passengers on the platform influences the dwell times of trains at stations. These, in turn, cause an increasing of delays which produces an increasing in headways. As already explained, in the case of metro-rail contexts higher headways could generate more passenger flows on the platform producing a further extension of the dwell times. This phenomenon describes the dynamic interaction between passengers and rail service and it is called '*snowball effect*' (Kanai et al., 2011) since delays increase at each station as a snowball.

Another important benefit provided by the ***OPM*** is the determination of crowding level within each coach. Indeed, the dwell time estimation problem, provides the load diagrams for single coach. This information can be extremely useful for train operating companies. Indeed, it is thus possible to plan and organise the fleet composition according to customers' attitude or to design Intelligent Transportation System (ITS) to inform passengers how to place themselves on the platform while waiting for the approaching train.

The following paragraphs are addressed to the specification of both Pre-Platform and On-Platform models.

4.5.1 The Pre-Platform Demand model.

The Pre-Platform Model can be viewed as a path-mode choice model which provides user platform flows. This is the result of a choice process which involves all transportation systems within the study area. In other words, users generally choose how to move from a generic origin *o* to a generic destination *d* evaluating among all possible alternatives, the one which maximises their utility (i.e. lower generalised cost). Obviously, it is still necessary to consider that the dynamic evolution of the flow propagation makes this utility be not constant but dependent on the service performances at time *t*.

Therefore, let:

- $h_{od,m}(t)$ be the path flow vector related to the pair *od* and the mode *m* at time t;
- $V_{od,m}(t)$ be the vector of systematic utilities for paths related to the pair *od* and the mode *m* at time τ;
- $p_{od,m}$ be the vector of path choice probabilities for the *od* and the mode *m*, whose elements are the probabilities of choosing path *k*, given *od* and *m*;
- $d_{od,m}$ be the demand flow of the users between the pair *od* with mode *m*.

The Pre-Platform Model can be written as follows:

$$h_{od,m}(t) = d_{od,m} p_{od,m}(V_{od,m}(t)) \; \forall od,m \qquad (4.30)$$

with:

$$V_{od,m} = -g_{od,m}(t) + V^{\circ}{}_{od,m} \; \forall od,m \qquad (4.31)$$

where $g_{od,m}(t)$ is the path cost vector while $V^{\circ}{}_{od,m}$ is a vector whose elements consist of the systematic utility components depending on any other attributes differing from path costs (e.g. socio-economic attributes).

The demand flow $d_{od,m}$ for the pair *od* on mode *m* is generally defined by the product of three sub-models:

- An emission model which simulates the choice of whether or not travelling in a time period;
- A distribution model, which estimates the probability of going to a generic place;
- A mode choice model, which provides the probability of travelling by private car, public transport (e.g. bus trolley-bus, train) or simply walking.

However, the choice decision related to the transport mode is function of the path choice EMPU (Expected Maximum Perceived Utility) at time *t* (Cascetta, 2009):

$$d_{od,m} = d_{od,m}(s(t)) \ \forall od,m \tag{4.32}$$

Relation (4.32) describes analytically the choice process of users about the transport mode. Indeed, according to the random utility theory (Cascetta, 2009), the Expected Maximum Perceived Utility of the generic decision maker *I* associated to a given choice context is defined as the expected value of the maximum perceived utility over the alternatives available in the choice set.

4.5.2 The On-Platform Demand model.

The On-Platform Model (OPM) enables the dynamic assignment of rail passenger flows to the rail service. This process can be divided into two phases. The first phase concerns the introduction of rail capacity constraints of trains. As mentioned before, the model checks whether the number of passengers willing to board the train at the generic station exceeds the residual capacity of the train coming from the previous station. If the residual capacity is not enough, some passengers have to wait for the following rail vehicles on the platform. In this case, the model adopts a FIFO (First In – First Out) rule which means that these passengers have priority to board the following trains.

In fact, although passengers generally tend to mingle on the platforms, in the case of high levels of crowding, the freedom of movement is limited and hence the priority in being served (i.e. in boarding) is strongly correlated with the sequence of arrivals on platform, especially if part of the users are unable to board the first approaching train. Hence, the adoption of a FIFO rule could generate more realistic simulations. In addition, remembering that the target is the evaluation of the objective function (4.3), this assumption provides the possibility to estimate for each passenger precise waiting and running times experienced during the trip. A RIFO (Random In – First Out) approach, by contrast, although could take into account effects of mixing of passengers on platforms, would change waiting times and the related determination of objective function values into random variables with the effect that the optimal strategy would not be determined with an absolute certainty but it would be associated to a confidence interval (i.e. probability which expresses the reliability of the value). Obviously, the FIFO logic takes into account that on the same platform passengers may have different destinations and hence different alighting stations. Therefore, not only does the model estimate the surplus of passengers on the platform but computes also their destinations according to the attractiveness of the following stations.

In order to describe the basic principles of the ***OPM***, the following variables have to be introduced:

- The flow matrix $\boldsymbol{P}^t$ defined for each train *t*, whose generic element p^t_{od} is the number of passengers willing to board the train *t* in order to reach station *d* starting from station *o*;
- The matrix of passengers surplus $\boldsymbol{S}^t$ defined for each train *t*, whose generic element s^t_{od} indicates the number of passengers willing to board the train *t* in order to reach station *d* starting from station *o*, who is forced to remain on the platform waiting for a following run;

- The vector of residual capacity $\boldsymbol{RC}^t$, defined for each train *t*, whose generic element rc_s^t is the residual capacity of train *t* at station *s* and it is calculated as follows:

$$rc_s{}^t = mc^t - obp_{(s-1;s)} + ap_s \qquad (4.33)$$

  where mc^t is the maximum capacity of train *t*, $obp_{(s-1;sj)}$ is the number of passengers on-board between stations *s-1* and *s*, and ap_s is the number of alighting passengers;

- the matrix of the actual boarding passengers $\boldsymbol{BP}^s$, defined for each station *s*, which is an upper triangular matrix whose generic elements $bp^s_{t,t+k}$ is the number of passengers willing to board train *t* and getting on train *t+k* with $k \in [0; n-t]$, where *n* is the total number of runs during the whole daily service. This matrix is important for estimating the actual waiting time of each user on the platform.

Basically, at each station *s* the residual capacity of the generic train t $rc_s{}^t$ is compared to the number of passengers p^t_{od} willing to board the train. If $rc_s{}^t$ is higher than the passenger flow on the platform, all customers get on the train, s_{ts} is equal to zero and $bp_{tt}{}^s$ is equal to p_{od}; otherwise s_{ts} is not null and these users will have priority in boarding the following runs on the new passenger flows $p_{od}{}^{t+1}$ .of the train *t+1*. The basic principle (which can be easily modified according to the targets of the analysis) is that passengers do not leave the system until they manage to take a train. This assumption, although is not valid in reality especially in places where public transport systems are highly developed, enables the definition of a model which can be adapted to any contexts. Following this principle, the matrices $\boldsymbol{BP}^s$ can be determined providing important information about users' trips and the possibility to estimate the passenger waiting times on the platform.

Another important hypothesis has to be considered for the correct estimation of the generic flow matrix $\boldsymbol{P}^t$. In fact, when $rc_s^{\ t} < p_{od}^{\ t}$, it is necessary to know the destination choices of the '*not served*' users, in order to correct the matrix $\boldsymbol{P}^t$ of train *t* and to load the following flow matrices $\boldsymbol{P}^{t+i}$ of train *t+1*, *t+2* and so on. Obviously, it is impossible to know exactly the individual choice of passengers. However, it is assumed that $s_{od}^{\ t}$ can be divided proportionally to the attractiveness (***att***) of each possible destination, calculated as follows:

$$\forall \text{ train } t \text{ and station } o\text{: } att_d^{\ t} = \frac{p_{.d}^{\ t}}{\sum_d p_{.d}^{\ t}} \tag{4.34}$$

Therefore, multiplying $s_{od}^{\ t}$ by the attractiveness of all the following destinations which can be reached starting from station *o*, the right amount of flow can be subtracted from the matrices $\boldsymbol{P}^t$.

The second phase of the assignment process is related to the determination of passenger behaviour while boarding and alighting from the train and the dwell time necessary to complete this process. As already explained, this consists of solving a fixed point problem because of the reciprocal dependence between headways and dwell times. To this purpose, let:

- $\boldsymbol{dwt} = \vartheta(\boldsymbol{td})$ be the function expressing the dependence of dwell times on the number of boarding/alighting passengers. This function is a continuous function which has to be calibrated on the system analysed since it takes into account particular features related to rolling stock, stations as well as platform configuration;
- $\boldsymbol{hd} = \psi(\boldsymbol{dwt})$ be the function describing the simulation of the network for estimating the headways' variation according to the dwell time values.

As already said, the frequency of a metro rail service strongly influences the congestion level on the platform. Hence, assuming within a short time interval a constant arrival rate (***upf***) at station, it is possible to calculate the travel demand at each station s as:

$$\boldsymbol{td}_s = upf_s \cdot \boldsymbol{hd}_s \tag{4.35}$$

This relation expresses the direct dependence of dwell times on headways. Therefore, combining the abovementioned functions, it is obtained:

$$\begin{cases} \boldsymbol{dwt}^* = \vartheta(\boldsymbol{hd}^*) \\ \boldsymbol{hd}^* = \psi(\boldsymbol{dwt}^*) \end{cases} \tag{4.36}$$

In other words, the problem formulated by system of equations (4.36) highlights the necessity of finding the dwell time vector which produces the headway vector which generates again the same dwell time vector. According to the theory of fixed point problem (see Cantarella, 1997; Cascetta, 2009) this particular case is called '*compound fixed point problem*'. It involves two vectors, $\boldsymbol{x} \in S_x \subseteq E^n$ and $\boldsymbol{y} \in S_y \subseteq E^m$ with $n \neq m$, influencing reciprocally each other:

$$\begin{cases} \boldsymbol{y}^* = \eta(\boldsymbol{x}^*)\ \boldsymbol{x}^* \in S_x\ \boldsymbol{y}^* \in S_y \\ \boldsymbol{x}^* = \rho(\boldsymbol{y}^*)\ \boldsymbol{y}^* \in S_y\ \boldsymbol{x}^* \in S_x \end{cases} \tag{4.37}$$

This system can be written also as follows:

$$\boldsymbol{x}^* = \rho(\eta(\boldsymbol{x}^*))\ \boldsymbol{x}^* \in S_x \tag{4.38}$$

Likewise, problem (4.36) becomes:

$$\boldsymbol{dwt}^* = \vartheta(\psi(\boldsymbol{dwt}^*)) \tag{4.39}$$

The mathematical conditions for the solution of (4.39) are expressed by the Brouwer's theorem according to which the compound fixed point problem has at least one solution if both functions $\vartheta(.)$ and $\psi(.)$ are continuous, the definition set is a nonempty, compact and convex set, and:

$$S_{dwt} \subseteq \vartheta(\psi(S_{dwt})),\ \text{i.e.}\ \vartheta(\psi(\boldsymbol{x})) \in S_{dwt},\ \forall \boldsymbol{x} = \boldsymbol{dwt} \in S_{dwt} \tag{4.40}$$

The dwell time estimation problem (i.e. problem 4.39) fulfils partially the assumptions of Brouwer's theorem. Regarding the continuity property, both $\boldsymbol{dwt} = \vartheta(\boldsymbol{td})$ and $\boldsymbol{hd} = \psi(\boldsymbol{dwt})$ satisfy this condition. Indeed, function ψ is nothing but the motion differential equations described in paragraph 2.3 which are composed of <u>continuous functions of time *t*</u>. Likewise, function ϑ, although may vary from system to system due to the characteristics of the line and the rolling stock, evaluates the dwell times as function of passengers by means of continuous calibrated formula (see paragraph 5.2 for some examples).
Furthermore, the fulfilment of the other properties can be easily demonstrated. First of all, the definition set is nonempty since:

$$S_{dwt} = \{dwt_i \geq 0 \quad \forall i\} \tag{4.41}$$

which states that, for evident reasons, dwell time values are always defined and positive.
Furthermore, this set is limited and closed since:

$$dwt_i \in \left[0\,;max_j\left(dwt_j\right)\right] \quad \forall i \tag{4.42}$$

In fact, dwell times are limited due to the capacity constraints of rolling stock and the border values are included in the definition set.
The set S_{dwt} is also convex. Indeed, for each couple of points belonging to S_{dwt}, the joining segment is completely included in S_{dwt}. That is:

$$(1-\mu)\cdot \boldsymbol{dwt}' + \mu \cdot \boldsymbol{dwt}'' \in S_{dwt} \quad \forall \boldsymbol{dwt}', \boldsymbol{dwt}'' \in S_{dwt} \quad \forall \mu \in [0\,;1] \tag{4.43}$$

Finally, also relation (4.40) is fulfilled since set S_{dwt} includes all the values which can be obtained by means of function ψ. In effect, the latter consists of a microscopic simulation of the service with new dwell times as input variables, providing new headways which, in turn, generate new permissible dwell times.
Furthermore, sufficient condition which guarantees the uniqueness of the solution is that the two functions $\vartheta(.)$ and $\psi(.)$ must be monotone in the opposite direction. In other words, $\boldsymbol{dwt} = \vartheta(\boldsymbol{hd})$ must be a strictly increasing

function and $\boldsymbol{hd} = \psi(\boldsymbol{dwt})$ must be a non-increasing function. This condition is not satisfied since both functions are strictly increasing and neither of the two is monotone in the opposite direction:

$$\left(\psi\left(dwt^{'}\right)-\psi\left(dwt^{''}\right)\right)^{T}\left(dwt^{'}-dwt^{''}\right)>0\ \forall dwt^{'},dwt^{''}\in S_{dwt} \qquad (4.44)$$

$$\left(\vartheta\left(hd^{'}\right)-\vartheta\left(hd^{''}\right)\right)^{T}\left(hd^{'}-hd^{''}\right)>0\ \forall hd^{'}\neq hd^{''}\in\vartheta\left(S_{dwt}\right) \qquad (4.45)$$

This does not mean that the solution is not unique though, but there are no mathematical bases to be sure of this assumption according to traditional proof proposed in the literature. Hence, in order to give effectiveness to this approach, it is necessary to look for numerical evidences that the solution of this fixed point problem could be unique. To this purpose, some applications are presented in chapter 5. As far as the convergence solution is concerned, the analysis of possible algorithms for solving fixed point problems must be provided. In particular, two different methods have been considered, namely the '*iterative algorithm*' and the '*MSA* (Method of Successive Average) *framework algorithm*' (Sheffi and Powell, 1981; Cantarella, 1997).

The iterative algorithm is a procedure largely used in several engineering fields. Basically, the problem consists of starting from an initial value x_0 and then generating the sequence $\boldsymbol{dwt}^{0}$, $\boldsymbol{dwt}^{1}$, $\boldsymbol{dwt}^{2}, \ldots \boldsymbol{dwt}^{n}$ which is hoped to converge to the fixed point solution problem:

$$\boldsymbol{dwt}^{n+1,*} = \vartheta\left(\psi\left(\boldsymbol{dwt}^{n,*}\right)\right), \text{ with } n = 0,1,2,..... \qquad (4.46)$$

Unfortunately, the convergence of this procedure cannot be demonstrated analytically and therefore, it is not possible to exclude that the algorithm diverges. For this reason, it is necessary to implement a termination test so as to prevent the algorithm from performing an infinite number of iterations. Finally, the steps of the iterative algorithm can be summarised as follows:

1. Initialise the algorithm with a starting value of dwell time vector $\mathbf{dwt}^{i}$; this step is considered as iteration 0 (i.e. $i = 0$);

2. Run the ***SeSM*** for analysing the performance of the rail service (i.e. evaluates the new headways) according to the dwell time vector $\boldsymbol{dwt}^{i}$ provided at iteration *i*;
3. The ***OPM*** estimates a new dwell time vector, that is $\boldsymbol{dwt}^{i+1}$ based on the headways obtained by ***SeSM***;
4. The termination test checks whether the convergence is reached or the algorithm has diverged in the following way:
   a. If $max_j \left(\frac{dwt_j^{i+1} - dwt_j^{i}}{dwt_j^{i}} \right) < 0.01$ or $i > M$, then stop the algorithm;
   b. Else set *i* = *i* + 1 and start again from point 2.

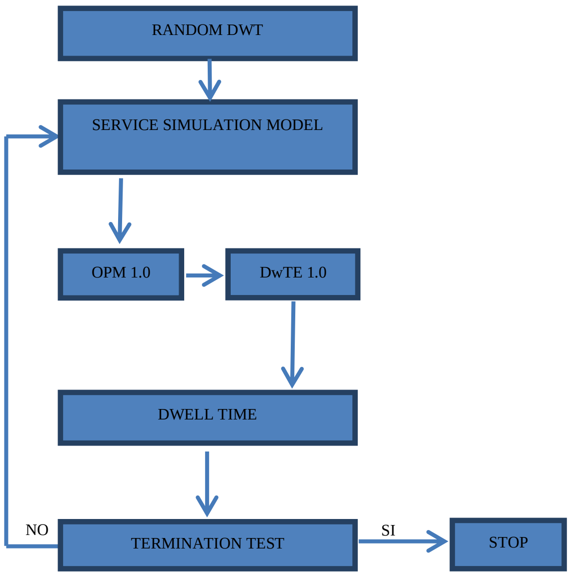


Figure 4.3 Graphic representation of the iterative algorithm.

The MSA algorithm is generally adopted for the resolution of stochastic traffic assignment over congested networks (Sheffi and Powell, 1981; Cantarella, 1997). In particular, it is based on the recursive equation:

$$x^{k+1} = (1 - \mu_k) x^k + \mu_k \, \psi(x^k) \tag{4.47}$$

where $\{\mu_k = 1/k\}_{k>0}$ is a sequence satisfying the conditions:

$$\sum_{k>0}\mu_k=\infty,\quad \sum_{k>0}\mu^2{}_k<\infty \tag{4.48}$$

The convergence of this algorithm for the resolution of the compound fixed point problem is guaranteed by the Blum's theorem (Blum, 1954) according to which if the assumptions of the Brouwer's theorem are satisfied, the sequence 4.47 defines a sequence convergent to the fixed point x^* (Sheffi and Powell, 1982; Cantarella, 1997). Therefore, even in this case, the convergence of the algorithm cannot be demonstrated since, as already explained, the uniqueness of the solution cannot be assured mathematically.

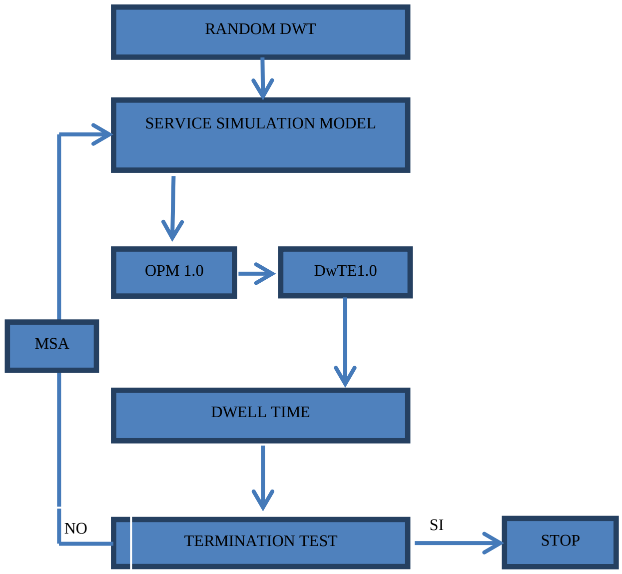


Figure 4.4 Graphic representation of the MSA algorithm.

The steps of the MSA procedure are the following:

1. Initialise the algorithm setting up a starting value of dwell time vector $\overline{\boldsymbol{dwt}}^i$ considering iteration i equal to 1;
2. The algorithm is initialised and the ***SeSM*** analyses the performance of the rail service (i.e. evaluates the new headways) according to the dwell time vector $\overline{\boldsymbol{dwt}}^i$;
3. The ***OPM*** estimates a new dwell time vector, that is $\boldsymbol{dwt}^i$ based on the headways obtained by ***SeSM***;

4. The new dwell time vector $\overline{\boldsymbol{dwt}}^{i+1}$ which is used in the following iteration is:

   $$\overline{\mathbf{dwt}}^{i+1} = \frac{i-1}{i}\overline{\mathbf{dwt}}^{i} + \frac{1}{i}\mathbf{dwt}^{i}$$

   where $\overline{\boldsymbol{dwt}}^{i}$ is equal to the average value of all previous $\boldsymbol{dwt}^{i}$;
5. The termination test checks whether the convergence is reached or the algorithm has diverged in the following way:
   a. If $\dfrac{\boldsymbol{dwt}^{i} - \overline{\boldsymbol{dwt}}^{i}}{\overline{\boldsymbol{dwt}}^{i}} < 0.01$ or $k > M$, then stop the algorithm;
   b. Else set $i = i + 1$ and start again from point 2.

Contrary to the iterative algorithm, the MSA generates at each iteration decimal values. Obviously, these values need to be rounded up/down to the integer part before being set up within the simulation model. This process can cause theoretical problem for the achievement of the convergence solution. Indeed, at each iteration, the MSA algorithm moves from a permissible point to another one within the convex definition set S_{dwt}. Rounding these values does not assure that the new point is still included within this set. However, it is worth highlighting that the new values obtained are included within a very little neighbourhood of the solution and hence, it is likely that this theoretical problem does not exist. However, this process can also cause a slowdown in reaching the convergence solution and makes the iterative algorithm be easier and faster than the MSA procedure.

In conclusion, in this chapter the framework of the Decision Support System (DSS) for planning or managing the rail network in any kind of service conditions is discussed in detail, providing the analytic formulation of each model (i.e. ***FM, SeSM, SM, PPM*** and ***OPM***).

Next chapter is instead dedicated to the formulation of a specific application which performs the dynamic assignment of passenger flows to the rail service.

CHAPTER 5: DEVELOPMENT OF AN APPLICATION FOR THE ASSIGNMENT OF TRAVEL DEMAND TO THE RAIL SYSTEM.

In order to perform the dynamic assignment of passenger flow to the service, the On-Platform Model has to be implemented. To this purpose, this chapter presents the development of a specific application which works in combination with a general microscopic simulation software. It can be viewed as an API (Application Programming Interface) which focuses on the travel demand assignment. The name of this '*own-built*' software is OPM 1.0 (On-Platform Model) which is developed in C++ language. In addition, this tool can be combined with another module, whose name is DwTE 1.0 (Dwell Time Estimation), for the estimation of the dwell time at stations as previously described. Concerning the simulation software for the analysis of the rail performances, it is worth highlighting that this thesis is not focused on the development of a specific microscopic model. Therefore, the analysis of the network performances is achieved by means of OPENTRACK® (see paragraph 3.3.2). The great advantage provided by the adoption of this software is the possibility to manage and modify some input and output values (e.g. simulation results, dwell times) also from outside the program by exporting txt files. This is extremely important to let different tools interact each other. However, due to the fact that commercial programs are protected by copyrights, it is not possible to modify their code and make this interaction be easier and completely automatic. Therefore, OPM 1.0 and DwTE 1.0 have a particular structure which guarantees their correct mode of operation with any kinds of simulation software.

5.1 OPM 1.0

OPM 1.0 reproduces exactly what has been explained in the previous chapter dealing with the On-Platform Model. Hence, it is the application enabling the flow assignment to the network according to the rail service performances.

The OPM 1.0 architecture is organised as described in Figure 5.1. Both input and output data are organised in specific folders by means of txt files. The input data are divided in four modules:

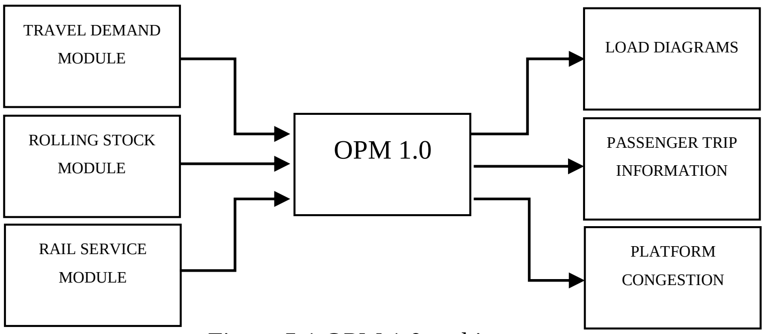


Figure 5.1 OPM 1.0 architecture.

- Travel demand module: the input data of this module concerns the passenger flow rate per minute at each station which is expressed by means of the Origin and Destination matrices $\mathbf{P}^t$. As explained in the previous chapter (see paragraph 4.5.2), the generic element $p_{od}^{\,t}$ represents the number of passengers willing to board the train t in order to reach station d starting from station o.

  This configuration manages to take into account the variation of the arrival rate on the platform during the service. Figure 5.2 shows an example of this file for a line with 17 stations;

matrix_0 - Blocco note

File Modifica Formato Visualizza ?

0.00	0.51	0.77	0.51	2.57	0.51	2.06	2.57	3.86	1.80	0.51	0.51	1.98	5.15	3.86	3.09	7.72
0.31	0.00	0.80	0.54	2.94	0.54	2.14	3.21	4.02	1.87	0.80	0.80	1.85	5.09	4.02	3.21	8.03
0.05	0.03	0.00	0.07	0.28	0.09	0.09	0.19	0.37	0.09	0.07	0.04	0.07	0.37	0.37	0.30	0.74
0.08	0.05	0.03	0.00	0.05	0.09	0.09	0.27	0.45	0.14	0.05	0.09	0.15	0.20	0.45	0.36	0.91
0.31	0.19	0.03	0.09	0.00	0.10	0.39	0.97	0.97	0.48	0.24	0.15	0.29	0.82	0.97	0.77	1.94
0.33	0.22	0.07	0.07	0.04	0.00	0.44	0.87	1.09	0.22	0.13	0.09	0.27	0.87	1.09	0.87	2.18
0.13	0.13	0.04	0.04	0.07	0.04	0.00	0.08	0.47	0.16	0.13	0.13	0.13	0.30	0.47	0.37	0.94
1.17	0.67	0.23	0.27	0.33	0.33	0.33	0.00	1.31	0.44	0.44	0.35	0.37	0.96	1.31	1.05	2.61
0.93	0.74	0.18	0.18	0.74	0.18	0.18	0.56	0.00	0.20	0.40	0.40	0.74	1.19	0.83	0.58	7.33
1.29	1.29	0.15	0.26	0.51	0.21	0.41	0.26	0.77	0.00	0.93	0.93	0.59	1.39	0.77	0.62	1.54
0.18	0.18	0.05	0.05	0.14	0.04	0.03	0.07	0.09	0.09	0.00	0.30	0.30	0.45	0.09	0.07	0.18
0.05	0.07	0.01	0.01	0.01	0.01	0.01	0.10	0.12	0.02	0.05	0.00	0.60	0.98	0.12	0.10	0.25
0.47	1.41	0.19	0.19	1.88	0.09	0.09	1.41	1.41	0.94	0.94	0.38	0.00	16.21	1.41	1.13	2.82
0.42	0.83	0.17	0.17	0.83	0.17	0.17	0.83	0.83	0.33	0.33	0.33	1.22	0.00	0.83	0.67	1.67
0.93	0.74	0.18	0.18	0.74	0.18	0.18	0.56	1.19	0.20	0.40	0.40	0.74	1.19	0.00	0.96	8.05
0.74	0.59	0.15	0.15	0.59	0.15	0.15	0.44	0.96	0.16	0.32	0.32	0.59	0.96	0.76	0.00	6.44
1.85	1.48	0.37	0.37	1.48	0.37	0.37	1.11	2.39	0.40	0.80	0.80	1.48	2.39	2.39	1.91	0.00

Figure 5.2 Input txt file concerning passenger arrival rate at stations for the generic train t.

- Rolling stock module: very detailed train characteristics are requested by this module (Figure 5.3), such as:
  - the maximum number of passengers per coach;

- the maximum number of sitting and standing passengers per coach;
- number of doors per coach;
- maximum number of passengers per train;
- available space within each coach.

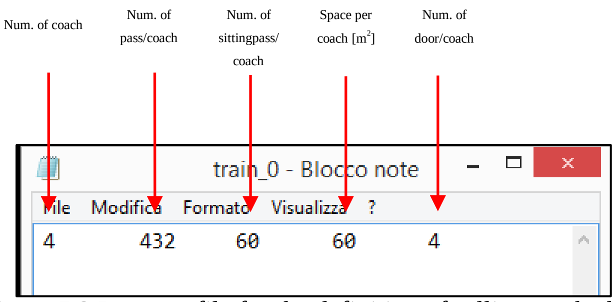


Figure 5.3 Input txt file for the definition of rolling stock characteristics.

- Rail service module: this module includes information about the simulated rail service. In particular, the data required are:
  - the fleet composition (i.e. the sequence of train implemented during the simulation divided by typology and number of coaches);
  - the headways (in minutes) and the running times (in seconds) of the simulated trains organised as shown in Figures 5.4 and 5.5;

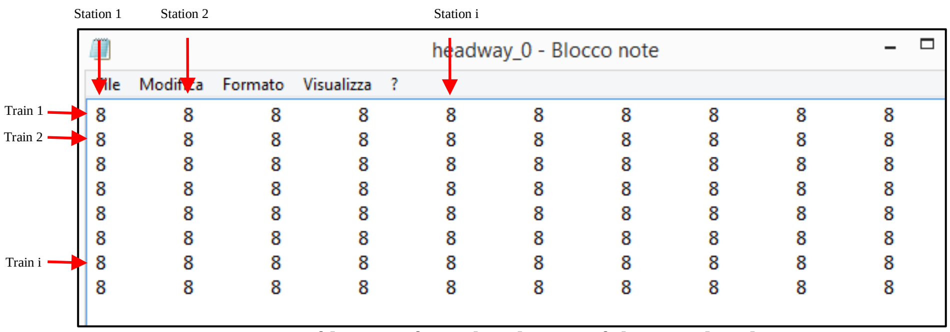
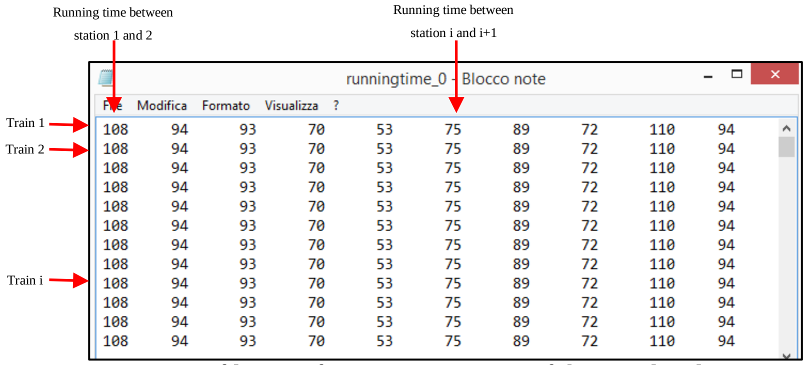


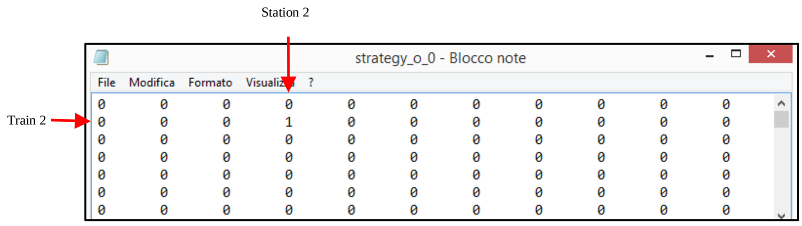
Figure 5.4 Txt file specifying headways of the simulated trains.

  - the operational service, that is the data concerning the path of each train. Empty movements are not taken into account since there is no influence with travel demand. However, in case a train is forced to stop its run at a station (because of a breakdown for example) passengers have to alight the train. This and other similar phenomena are reproduced through specific txt files whose name is ‘strategy.txt’, where the stations in which trains stop their runs are identified by code 1. The code 0 by contrast, characterises ordinary service conditions. Figure 5.6 shows the situation where train 2 stops its run at station 4;



108	94	93	70	53	75	89	72	110	94
108	94	93	70	53	75	89	72	110	94
108	94	93	70	53	75	89	72	110	94
108	94	93	70	53	75	89	72	110	94
108	94	93	70	53	75	89	72	110	94
108	94	93	70	53	75	89	72	110	94
108	94	93	70	53	75	89	72	110	94
108	94	93	70	53	75	89	72	110	94
108	94	93	70	53	75	89	72	110	94
108	94	93	70	53	75	89	72	110	94
108	94	93	70	53	75	89	72	110	94
108	94	93	70	53	75	89	72	110	94

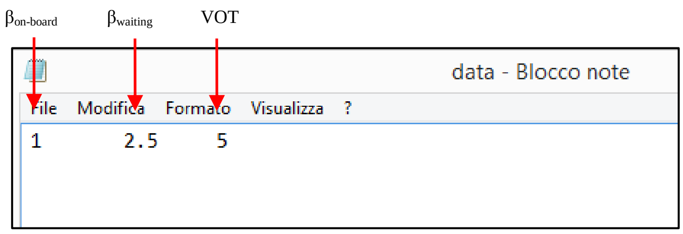
Figure 5.5 Txt file specifying running times of the simulated trains.



0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

Figure 5.6 Txt file ‘strategy’ for the definition of the operational service.

- User disutility module: in this module, the values of the β parameters related to the disutility perceived by passengers during their trip are gathered, as well as the VOT (Value Of Time) value (Figure 5.7).



1	2.5	5

Figure 5.7 Txt file defining the user disutility module.

In case the discomfort experienced on board the train depends on the rail crowding level, the $\beta_{on-board}(rc)$ is expressed by means of another txt file (Figure 5.8) which considers the passenger density within each coach (i.e. $\frac{pax}{m^2}$).

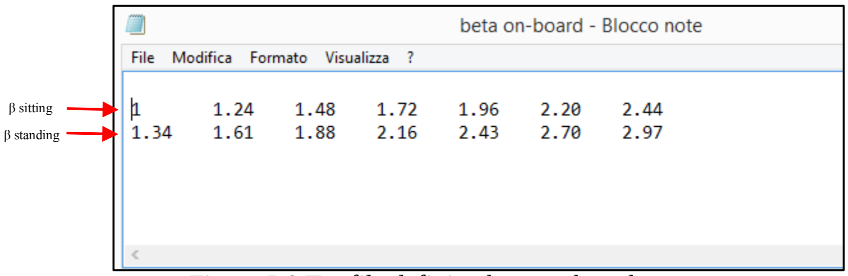


Figure 5.8 Txt file defining beta on-board parameters.

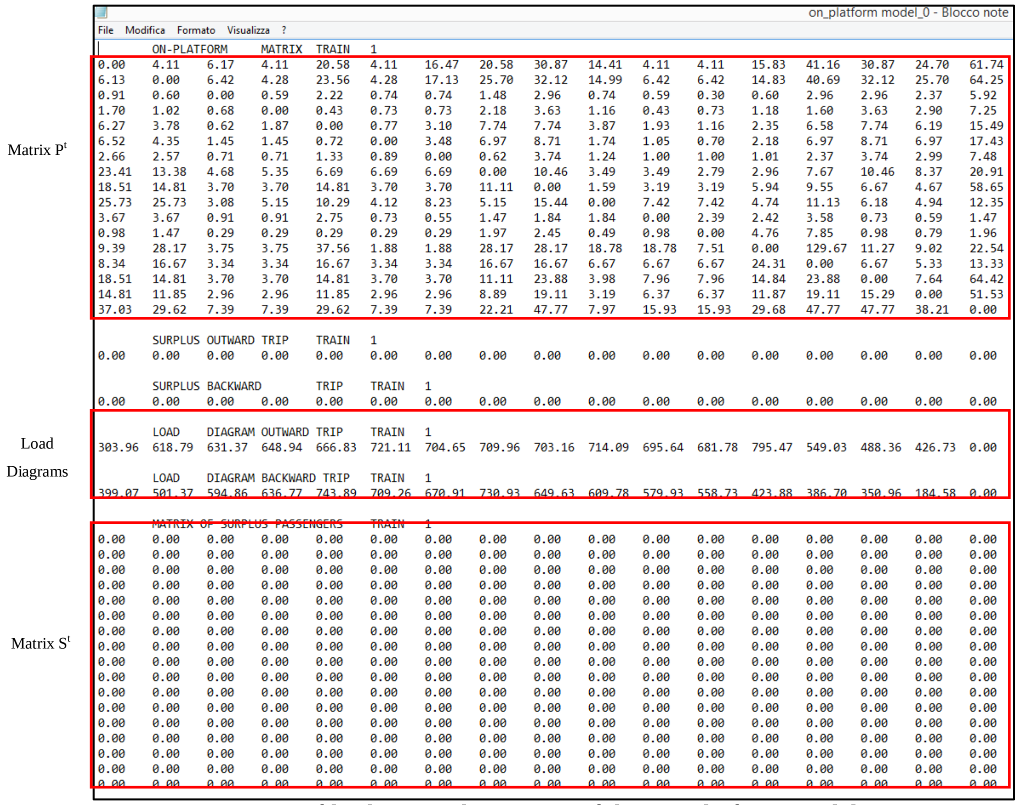
In this way, it is possible to use the proposed weights independently of the different interior layout of the trains (Wardman and Whelan, 2011). The values showed in Figure 5.8 are drawn from previous surveys carried out by the MVA Consultancy in Britain (see MVA Consultancy, 2008) whose results are summarised in Table 5.1.

	Non-business		Business		LSE		Regional		Interurban	
Pass./m^2	Sitting	Standing	Sitting	Standing	Sitting	Standing	Sitting	Standing	Sitting	Standing
0	1	1.48	1	1.91	1	1.43	**1**	**1.34**	1	1.77
1	1.1	1.58	1.13	1.95	1.09	1.56	**1.24**	**1.61**	1.11	1.81
2	1.21	1.68	1.27	1.99	1.18	1.69	**1.48**	**1.88**	1.23	1.85
3	1.31	1.77	1.4	2.03	1.27	1.82	**1.72**	**2.16**	1.34	1.89
4	1.41	1.87	1.54	2.08	1.36	1.95	**1.96**	**2.43**	1.46	1.92
5	1.52	1.97	1.67	2.12	1.45	2.08	**2.2**	**2.7**	1.57	1.96
6	1.62	2.06	1.81	2.16	1.54	2.21	**2.44**	**2.97**	1.69	2

Table 5.1 Different values of the $\beta_{on-board}$

As outputs, OPM 1.0 provides the actual load diagram for each train, information about passengers trip (which train they managed to take, how many trains they were forced to wait for, waiting times at stations), platform congestion and the user generalised cost for each train and for the whole simulated scenario. In particular, the output files are organised as follows:

- 'opm.txt': this file summarises the results of the On-Platform Model including the actual load diagrams (even and odd runs) with the correct form of matrix $\boldsymbol{P}^t$ (i.e. taking into account the maximum capacity of each train) and matrix $\boldsymbol{S}^t$ (Figure 5.9);




```
on_platform model_0 - Blocco note
File  Modifica  Formato  Visualizza  ?
        ON-PLATFORM     MATRIX  TRAIN   1
0.00    4.11    6.17    4.11    20.58   4.11    16.47   20.58   30.87   14.41   4.11    4.11    15.83   41.16   30.87   24.70   61.74
6.13    0.00    6.42    4.28    23.56   4.28    17.13   25.70   32.12   14.99   6.42    6.42    14.83   40.69   32.12   25.70   64.25
0.91    0.60    0.00    0.59    2.22    0.74    0.74    1.48    2.96    0.74    0.59    0.30    0.60    2.96    2.96    2.37    5.92
1.70    1.02    0.68    0.00    0.43    0.73    0.73    2.18    3.63    1.16    0.43    0.73    1.18    1.60    3.63    2.90    7.25
6.27    3.78    0.62    1.87    0.00    0.77    3.10    7.74    7.74    3.87    1.93    1.16    2.35    6.58    7.74    6.19    15.49
6.52    4.35    1.45    1.45    0.72    0.00    3.48    6.97    8.71    1.74    1.05    0.70    2.18    6.97    8.71    6.97    17.43
2.66    2.57    0.71    0.71    1.33    0.89    0.00    0.62    3.74    1.24    1.00    1.00    1.01    2.37    3.74    2.99    7.48
23.41   13.38   4.68    5.35    6.69    6.69    6.69    0.00    10.46   3.49    3.49    2.79    2.96    7.67    10.46   8.37    20.91
18.51   14.81   3.70    3.70    14.81   3.70    3.70    11.11   0.00    1.59    3.19    3.19    5.94    9.55    6.67    4.67    58.65
25.73   25.73   3.08    5.15    10.29   4.12    8.23    5.15    15.44   0.00    7.42    7.42    4.74    11.13   6.18    4.94    12.35
3.67    3.67    0.91    0.91    2.75    0.73    0.55    1.47    1.84    1.84    0.00    2.39    2.42    3.58    0.73    0.59    1.47
0.98    1.47    0.29    0.29    0.29    0.29    0.29    1.97    2.45    0.49    0.98    0.00    4.76    7.85    0.98    0.79    1.96
9.39    28.17   3.75    3.75    37.56   1.88    1.88    28.17   28.17   18.78   18.78   7.51    0.00    129.67  11.27   9.02    22.54
8.34    16.67   3.34    3.34    16.67   3.34    3.34    16.67   16.67   6.67    6.67    6.67    24.31   0.00    6.67    5.33    13.33
18.51   14.81   3.70    3.70    14.81   3.70    3.70    11.11   23.88   3.98    7.96    7.96    14.84   23.88   0.00    7.64    64.42
14.81   11.85   2.96    2.96    11.85   2.96    2.96    8.89    19.11   3.19    6.37    6.37    11.87   19.11   15.29   0.00    51.53
37.03   29.62   7.39    7.39    29.62   7.39    7.39    22.21   47.77   7.97    15.93   15.93   29.68   47.77   47.77   38.21   0.00

        SURPLUS OUTWARD TRIP    TRAIN   1
0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00

        SURPLUS BACKWARD        TRIP    TRAIN   1
0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00

        LOAD    DIAGRAM OUTWARD TRIP    TRAIN   1
303.96  618.79  631.37  648.94  666.83  721.11  704.65  709.96  703.16  714.09  695.64  681.78  795.47  549.03  488.36  426.73  0.00

        LOAD    DIAGRAM BACKWARD TRIP   TRAIN   1
399.07  501.37  594.86  636.77  743.89  709.26  670.91  730.93  649.63  609.78  579.93  558.73  423.88  386.70  350.96  184.58  0.00

        MATRIX OF SURPLUS PASSENGERS    TRAIN   1
0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00
0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00    0.00
```

Figure 5.9 Txt file showing the outputs of the On-Platform Model.

- 'cost.txt': this file gathers the user generalised cost of each simulated train and the total cost of the of the strategy (Figure 5.10);
- 'boarding.txt': this file provides for each station the boarding passenger matrix $\boldsymbol{BP}^s$ which is extremely important for the estimation of passenger waiting times. In fact, assuming that users arrive at the station according to a Poisson process with a constant arrival rate, the average waiting time is evaluated as (Cascetta, 2009):

$$t_w = \frac{\vartheta}{\varphi} \tag{5.1}$$

Where φ is the frequency of the line and ϑ is equal to 0.5 in case of a perfectly regular service (i.e. the headways between successive vehicle arrivals are constant) and is equal to 1 if the line is ‘completely irregular’ (i.e. the headways between successive arrivals are distributed according to a negative exponential random variable). Formula (5.1) calculates the waiting times of passengers who get on the first arriving train. In case this event does not occur, knowing the matrix $\boldsymbol{BP}^s$, it is possible to evaluate the exact increase of waiting time.

```
cost_0 - Blocco note
File  Modifica  Formato  Visualizza  ?
TOTAL COST TRAIN 1: 4970.12 euro
TOTAL COST TRAIN 2: 9385.95 euro
TOTAL COST TRAIN 3: 7921.13 euro
TOTAL COST TRAIN 4: 6418.25 euro
TOTAL COST TRAIN 5: 5611.97 euro
TOTAL COST TRAIN 6: 5698.42 euro
TOTAL COST TRAIN 7: 5760.03 euro
TOTAL COST TRAIN 8: 5957.87 euro
TOTAL COST TRAIN 9: 6449.66 euro
TOTAL COST TRAIN 10: 1402.04 euro
TOTAL COST TRAIN 11: 10204.2 euro
TOTAL COST TRAIN 12: 13305.3 euro
TOTAL COST TRAIN 13: 12627 euro
TOTAL COST TRAIN 14: 9775 euro
TOTAL COST TRAIN 15: 8203.76 euro
TOTAL COST TRAIN 16: 8287.7 euro
TOTAL COST TRAIN 17: 7898.5 euro
TOTAL COST TRAIN 18: 8362.58 euro
TOTAL COST TRAIN 19: 12211.3 euro
TOTAL COST TRAIN 20: 13150.5 euro
TOTAL COST TRAIN 21: 13666.4 euro
TOTAL COST TRAIN 22: 14670.2 euro
TOTAL COST TRAIN 23: 15632.2 euro
TOTAL COST TRAIN 24: 13014.4 euro
TOTAL COST TRAIN 25: 10423.9 euro
TOTAL COST TRAIN 26: 13200.5 euro
TOTAL COST TRAIN 27: 12942.5 euro
TOTAL COST TRAIN 28: 12956.8 euro
TOTAL COST TRAIN 29: 13144.1 euro
TOTAL COST TRAIN 30: 12816.7 euro
TOTAL COST TRAIN 31: 12663.7 euro
TOTAL COST TRAIN 32: 12243.1 euro
TOTAL COST TRAIN 33: 12152.4 euro
TOTAL COST TRAIN 34: 11666.5 euro
TOTAL COST TRAIN 35: 11681.5 euro
TOTAL COST TRAIN 36: 11734.9 euro
TOTAL COST TRAIN 37: 12804.7 euro
TOTAL COST TRAIN 38: 12053.5 euro
TOTAL COST TRAIN 39: 12289 euro
TOTAL COST TRAIN 40: 13282.3 euro
TOTAL COST TRAIN 41: 20355.1 euro
TOTAL COST TRAIN 42: 29004.2 euro
TOTAL COST TRAIN 43: 29378.4 euro
TOTAL COST TRAIN 44: 26972.8 euro
TOTAL COST TRAIN 45: 25722.8 euro
TOTAL COST TRAIN 46: 25043.9 euro
TOTAL COST TRAIN 47: 23793.4 euro
TOTAL COST TRAIN 48: 23433.4 euro
TOTAL COST TRAIN 49: 21583.7 euro
TOTAL COST TRAIN 50: 18825.1 euro
TOTAL COST TRAIN 51: 19286.1 euro
TOTAL COST TRAIN 52: 21751.9 euro
TOTAL COST TRAIN 53: 24974 euro
TOTAL COST TRAIN 54: 22159.3 euro
TOTAL COST TRAIN 55: 18899.8 euro
TOTAL COST TRAIN 56: 17808.1 euro
TOTAL COST TRAIN 57: 22561.3 euro
TOTAL COST TRAIN 58: 20327.1 euro
TOTAL COST TRAIN 59: 24660.7 euro
TOTAL COST OF STRATEGY: 830520.98 euro
```

Figure 5.10 Txt file with the strategy costs.

This particular structure, once having set up all input data, enables the analysis of more scenarios simultaneously. In fact, before starting the computations, the program asks the user about the number of scenarios reproduced by the microscopic simulation model and the number of stations involved (Figure 5.11). Thus, OPM 1.0 can be launched one time and the output files which refer to different strategies are gathered in the same folders with the same name but with a different subscript (Figure 5.12).This solution is extremely useful in the case of stochastic analyses, where simulations are numerous and OPM 1.0 can be launched just one time providing all the results.

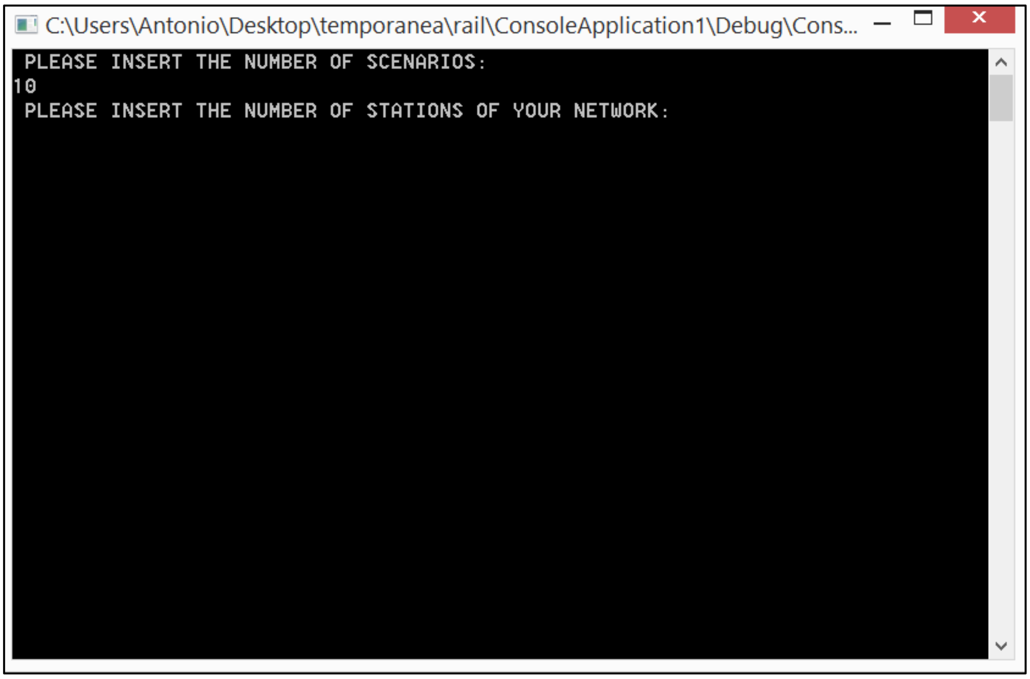


Figure 5.11 The OPM 1.0 console.

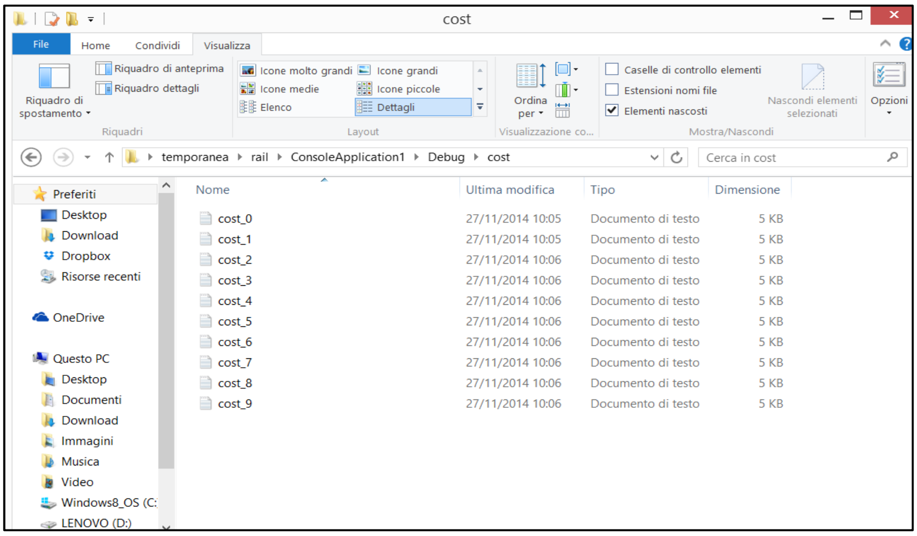


Figure 5.12 Cost folder with all outputs of the simulated scenarios.

5.2 DwTE 1.0

The DwTE module is a specific tool for the estimation of dwell times at stations. It can be viewed as a module of OPM 1.0, but the idea of considering the two applications separated is due to the fact that the analysis of dwell times differs according to the rail system considered and sometimes it is neglected. As already explained, metro-rail lines are strongly influenced by passenger flows insomuch as the boarding/alighting process requires the major part of the dwell times. Generally, although erroneously, the dwell time at station of regional or long distance trains is not analysed in detail. It is usually considered as a constant time window (generally equal to 60 seconds) useful for recovering from small delays. Hence, in most cases, a more correct estimation is not required. However, due to the great increase of travel demand on conventional rail lines, few examples in the literature showed the importance of considering the influence of crowding also on conventional line services (Nash, et al., 2006; Buchmueller et al., 2010) in order to avoid conflicts between trains. Furthermore, it is unquestionable that a better estimation of the dwell time is useful to increase the operational efficiency of rail transportation system in terms of speed and energy consumption (Hansen and Pachl, 2008). In fact, dwell times have a key role since they are part of a time window which allows drivers to adapt their driving behaviour in order to reduce the energy consumption (Albrecht et al., 2013). Hence, knowing the time necessary to complete the boarding/alighting process can be a great advantage for several tasks.

DwTE 1.0 computes the dwell time as a result of passengers' behaviour while boarding and alighting from the train. To this aim, the model is organised as shown in Figure 5.13. The input data are divided in three modules:

- Passenger flow module: basically, this module includes data regarding the passenger flow on the platform and all information about their trip. This data are provided by the OPM 1.0 tool through matrices $\boldsymbol{P}^t$;

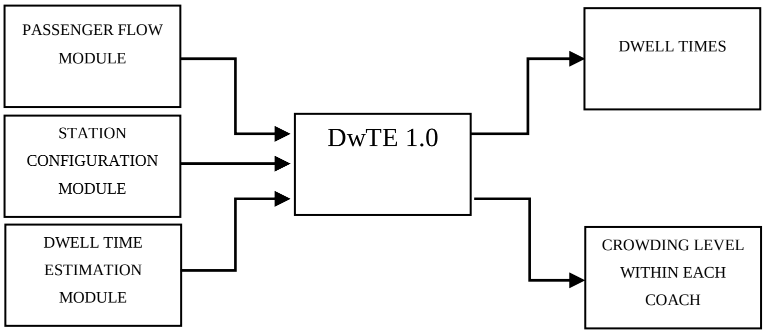


Figure 5.13 DwTE 1.0 architecture.

- Station configuration module: generally, passengers prefer boarding the train considering their final destination (Kunimatsu et al., 2012). In fact, especially commuters, who know the system very well, seek to minimise their walking distance at the station they want to reach and therefore, they choose the door of the train which is as close as possible to the stairs or the elevators (Figure 5.14). For each station, these preferences are described through file ‘chosen_door.txt’ (Figure 5.15).

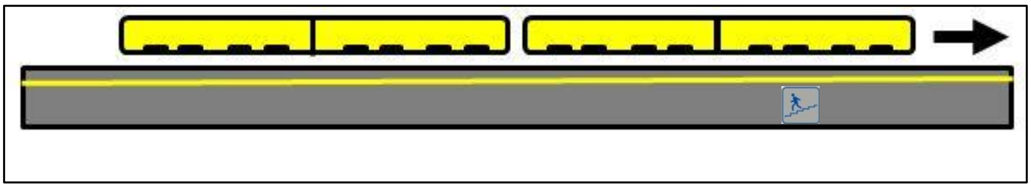


Figure 5.14 Example of the stair location in a metropolitan station.

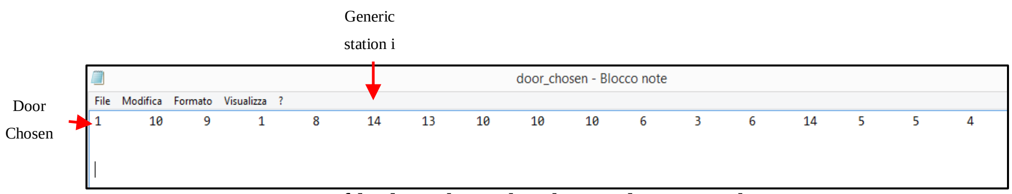


Figure 5.15 Txt file describing the chosen door at each station.

- Dwell time estimation module: as already explained, the dwell time is mainly influenced by the number of boarding and alighting passengers. Indeed, the function expressing this dependence must be calibrated on the analysed system. Figure 5.16 and 5.17 represent two examples of dwell time estimation formula. The first one refers to Japan regional railway lines, whose formulation is the following (Toriumi et al., 2005):

$$max\left(\left(1+\rho\right)\left(21.9\,log\left(x\right)-37.1\right), minimum\,dwell\,time\right) \qquad (5.2)$$

Where x is the average number of passengers per door who get on/off at the station and ρ is a concentration ratio, namely the ratio of the number of passengers at the most congested door to the average number of passengers who get on/off from one door. This formula has been calibrated in the case of trains composed of 20 meters long coaches with four doors. Obviously, for different kinds of trains all the parameters must be changed.

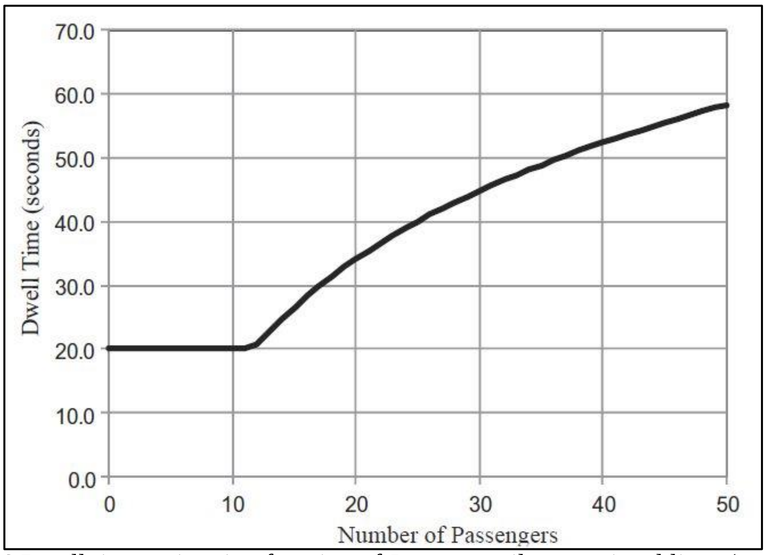


Figure 5.16 Dwell time estimation function of Japanese railway regional lines (source: Toutiumi et al, 2005).

The second example describes the dwell time function estimated for Line 1 of Naples metro system, that is:

$$\begin{cases} y=5 & if \;\; x \leq 3.5279 \\ y=0.8602x+1.9653 & if \;\; x > 3.5279 \end{cases} \qquad (5.3)$$

Where x is the number of passengers boarding/alighting the train on the most loaded door and y is the correspondent dwell time. As for the previous case, even formula (5.3) is specific for the station configuration and the rolling stock of this line. This means that other systems need a new calibration process so as to determine their suitable function.

More in detail, DwTE 1.0 works according to the following estimation procedure. First of all, on the basis of the station configuration module, passengers move toward the door they prefer. When, in front of a door, there is a group of passengers higher than a prefixed value (which, obviously, can be modified according to the network characteristics), they start moving to the adjacent doors which belong to the same coach. In fact, in the case of crowded situation, the customer target is to get on the train as rapidly as possible trying to remain close to the first favourite door. If a coach becomes full, passengers move toward the others which attract flow according to their available capacity.

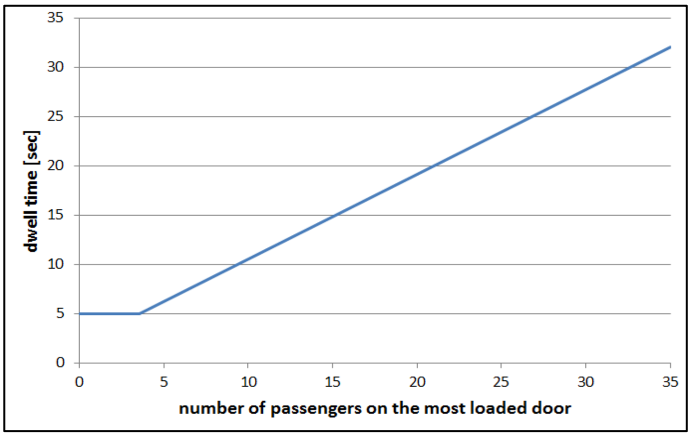


Figure 5.17 Dwell time estimation function of Line 1 of Naples metro system.

This means that the more a coach is empty, the more passengers will get on it. It is worth noting that this process is simultaneously with that analysed by the OPM 1.0 which checks the number of passengers that can board the whole train. DwTE 1.0 by contrast, focuses on how these passengers get on/off from each coach and the door chosen to complete this action. An important assumption is necessary though: following the explained procedure, the door chosen to board the train will be the same to alight it. However, especially in crowded situations, usually passengers find many difficulties to move within the coach and therefore, it is likely that this assumption is satisfied. Finally, once the number of boarding and alighting passengers is known, the model adopts the function provided as input data to estimate the dwell time of each simulated train and at each stations .These values are organised as vectors distributed in a txt file (Figure 5.18). In addition, by estimating the door chosen to get on the train, the model provides information about crowding level within each coach. This result, as already stated, is extremely useful for the calculation of the disutility perceived on-board by passengers. Furthermore, it enables to plan a fleet composition closer to customers' necessity and to design Intelligent Transportation Systems (ITS) which could help to reduce the amount of dwell time.

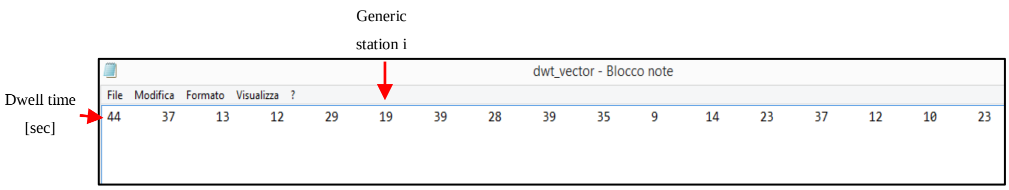


Figure 5.18 Txt file showing the estimated dwell time at each station.

In conclusion, this chapter has dealt with the definition of an application which performs the dynamic assignment of passenger to the rail service working in combination with a microscopic simulation software.

In particular, OPM 1.0 reproduces the On-Platform Model and gives the possibility of loading the trains according to the performances of the network and considering their limited capacity.

DwTE 1.0 instead is a specific module for the evaluation of the dwell times at stations. Basically, it reproduces passenger behaviour on the platforms and estimates the number of boarding and alighting users per door. By means of a specific calibrated formula, it is thus possible to obtain the amount of time the train has to stop within the station.

Both applications represent an important enhancement of microscopic simulation programs which are mainly based on the calculation of train motion and neglect the interaction with flows which, as will be demonstrated in the following chapter, is often essential to carry out complete analyses of the rail service considering at the same time rail operator and passenger perspectives.

CHAPTER 6: APPLICATIONS OF THE PROPOSED PROCEDURE TO SUPPORT THE MANAGEMENT OF THE RAILWAY SYSTEM.

The aim of this chapter is to demonstrate the benefits of the proposed model for supporting the management of the rail system in any kind of service conditions. In particular, all applications performed and published during the course of the PhD are described in order to show the effectiveness of the procedure and its applicability over different kinds of railway contexts. Specifically, the majority of the applications refers to Line 1 of Naples metro system. Therefore, a short introduction of this network is presented so as to clarify the weaknesses of this system and to justify the different test cases analysed.

Then, the chapter is organised as follows. First of all, the planning of recovery solutions considering the service quality is described (D'Acierno et al., 2012). In particular, impacts on travel demand due to system breakdowns are analysed under varying levels of performance (D'Acierno et al., 2013a), travel demand flow (D'Acierno et al., 2013b) and fleet composition (Placido et al., 2014c).

The second group of applications focuses on the Pre-Platform Model, showing some procedures already known in the literature to estimate different arrival rates on the platform during the daily service (Ercolani et al., 2014).

The following test cases concern the robustness of intervention strategies taking into account the stochastic distribution of train performance and delays (Placido et al., 2015b).

An application on recovery strategies analysis considering the stochastic variability of dwell times is described (Placido et al., 2015a). In this case, the convergence of the dwell time estimation problem is demonstrated trough the adoption of both iterative and MSA algorithms.

Then, the development of a new objective function for increasing the efficiency and the effectiveness of the rail system is presented.

Another important improvement is the combination of the proposed methodology with a macro-optimisation model for managing rail systems in case of disruptions. This enhancement increases the computational efficiency of the microscopic procedure and reduces the number of intervention strategies which should be assessed (Placido et al., 2014a).

Finally, a test case on a conventional regional rail line in the south of Italy is also discussed in order to highlight the possibility to adopt this decision support system in other railway contexts.

6.1 The Line 1 of the Naples metro system.

Line 1 of the Naples metro system is famous all over Europe thanks to the architectural and artistic beauty of its stations. It is more than a simple transportation system since it is completely integrated within the urban fabric. Indeed, all the stations have been designed and constructed like museums hosting permanent art exhibitions in perfect harmony with the historical identity of the city and of the areas in which they are realised. Actually, the impact of station beauty on travellers' behaviour has been remarkable insomuch as it has considerably increased their propensity to use rail services (Cascetta et al., 2013; Cascetta and Cartenì, 2014; Cascetta et al., 2014).

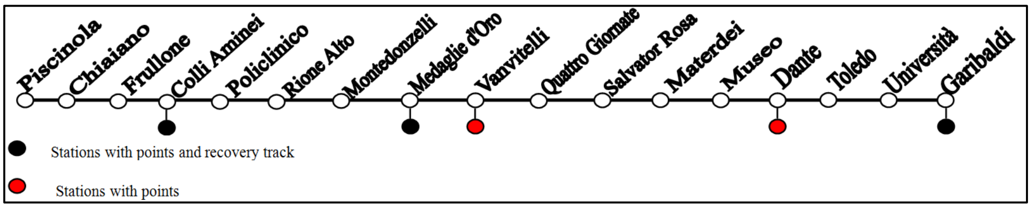


Figure 6.1 Line 1 stations.

The Line 1 is operated by METRONAPOLI (recently absorbed by ANM transport company) and it is nowadays (since December 31st, 2013) composed of 17 stations (Figure 6.1). The infrastructure is extremely complex because of the hilly terrain in the city. Indeed, steep slopes and low radius curves have led to the construction of two completely separate tunnels, one per direction. Only certain stations are equipped with points and/or recovery tracks which reduce the elasticity of the system in the event of failure. More in detail:

- Piscinola station has the points which allow trains to change track and connect the line to the depot;

- Colli Aminei station has the points for the track change and presents also a recovery track where faulty train can be driven to;
- likewise, Medaglie d'Oro station has both the points and the recovery track;
- Dante station has just the points for making the track change possible;
- after Garibaldi station there is a specific area with points for changing track and reversing the train motion.

In order to clarify the applications which are then described, it is necessary to divide the line into three parts:

- the Piscinola–Dante section, 13.47 km long, consisting of 14 stations;
- the Dante–Università section, 1.87 km long, consisting of 3 stations;
- the Università–Garibaldi section, 1.68 km long, consisting of 2 stations.

It is worth noting that until December 31st 2013, the third section (i.e. Università–Garibaldi) was under construction and since the second section (i.e. Dante–Università) had two separate tunnels without any point which enabled a metro convoy to change tracks, the metro services were performed as follows: a regular metro service between Piscinola and Dante (first section) ran on a double-track section; a shuttle service between Dante and Università (second section) ran on just one of the two tracks of the section. In 2014, the opening of the third section provided the possibility to have a metropolitan service over the whole line.

From a technological point of view, the whole infrastructure is extremely advanced. The signalling system is composed of (see 2.4 for more details concerning the signalling system):

- Electronic interlocking machines for routing trains within the stations;
- A.T.I.S. system (Audio-frequency Transmission and Interlocking System), whose main functions are: train detection on the tracks, check of track integrity and computation and transmission to the trains of the

information regarding the state of the signalling devices so as to allow the on board instrumentation to regulate the train running;

- On board equipment, namely continuous ATP system, discontinuous ATP system and ATO system.

The rolling stock is constituted of trains which can be composed of one, two or three traction units. Each traction unit is, in turn, composed of two carriages with a maximum capacity of 432 passengers (120 sitting and 312 standing passengers). Therefore, a complete train can carry at most 1296 passengers.

The Line 1 plays a key role in the Naples public transportation system inasmuch as it connects the high density suburbs with the city centre. For this reason, especially during peak hours, the line is extremely crowded since customers cannot rely on the performance of alternative means of transport (e.g. buses or trams) which is generally lower due to the high congestion level of the main roads. However, due to a lack of rolling stock and to the complexity of the network, especially in case of breakdowns, re-establishing ordinary conditions could involve inconveniently long travel times. Furthermore, there is just one depot located near Piscinola station and spare trains are not always available. Indeed, when there is a faulty train in the network, dispatchers prefer to close the whole line and remove the broken-down convoy or, should it be possible, just leave the service without any kind of intervention. Obviously, this results in great discomfort for passengers who are not considered at all. Although this strategy is the easiest to implement and is also optimal from an operational point of view, it does not fulfil customers' needs. Therefore, the Line 1 is the perfect test case for evaluating the effectiveness of the proposed approach.

Figure 6.2 and 6.3 show the infrastructure, the signalling system and the rolling stock reproduced in OpenTrack software (i.e. the ***SeSM***) faithfully to the reality.

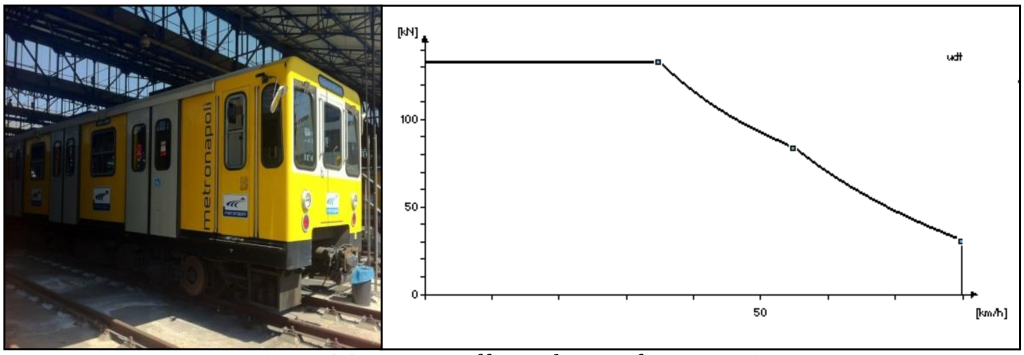


Figure 6.2 Tractive effort/velocity of Line 1 train.

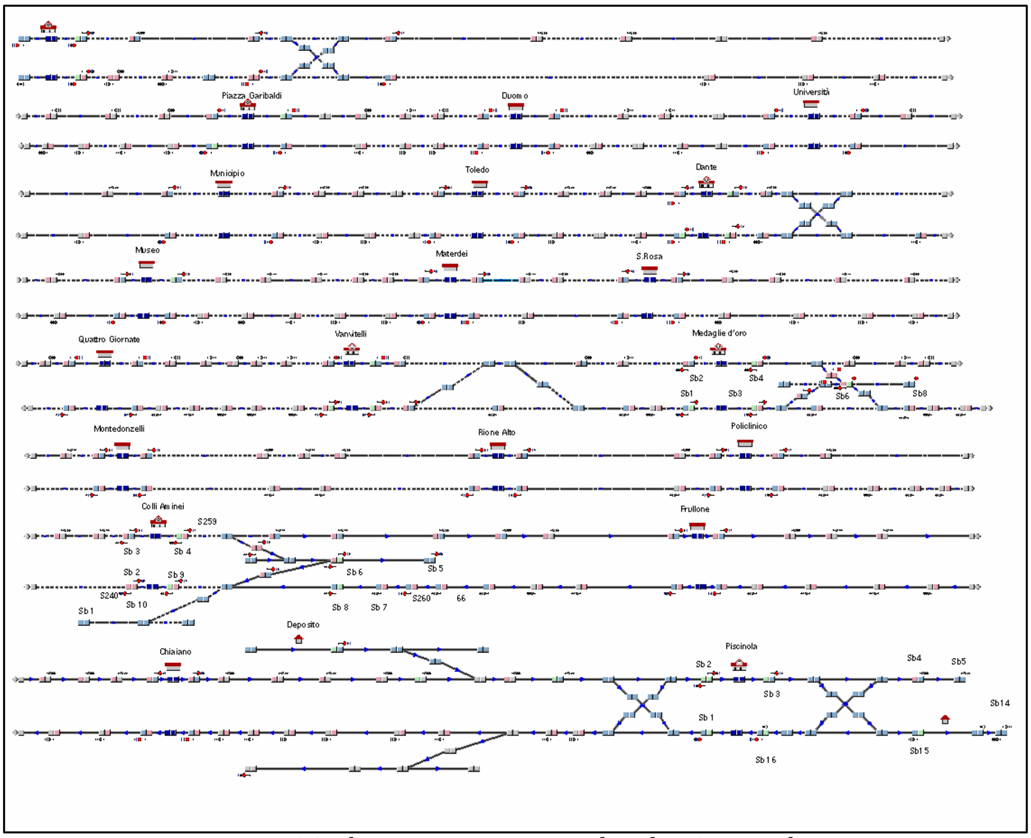


Figure 6.3 The Line 1 system reproduced in Opentrack.

6.2The management of the rail system considering the service quality.

The first application on the Line 1 metro system is intended to demonstrate the importance of considering the service quality perceived by passengers as the main target to achieve during the management of the network.

In particular, the analysed scenarios concern the old service involving the section from Piscinola station to Dante station and neglecting the shuttle service (Figure 6.4).

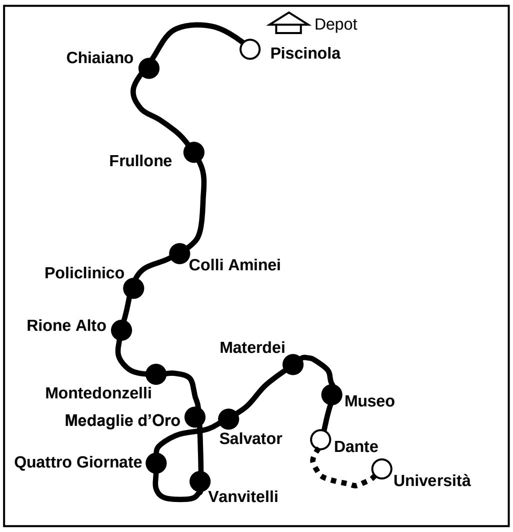


Figure 6.4 Particular of the Line 1 before the opening of the third section.

As shown in paragraph 4.2, the failure contexts which is worth analysing are provided by the Failure Model. However, for the sake of simplicity, RAMS analyses are not considered. Usually these surveys are carried out by rail operators and/or manufactures and hence, these data are already available. Nevertheless, although invented, the breakdowns simulated are always plausible. In this case, as first application, a general breakdown to the traction units of a train performing its service during the morning rush hour (i.e. 7.00 a.m. – 9.00 a.m.) is evaluated. As a consequence, two different failure scenarios are considered according to which the speed of the faulty train is limited respectively to the 80% and to the 20% of the maximum speed. Obviously, the simulation concerns a wider time period (from 6.00 a.m. to 12.00 p.m.) for analysing network loading (people and trains on the network at 7.00 generally started before) and discomfort duration (failure effects could last also beyond the peak-hour). In particular, the timetable in terms of headways simulated by OpenTrack is:

- 12 minutes (i.e. 5 trains/hour) between 6.00 am and 7.00 am;
- 7 minutes (i.e. 8.6 train/hour) between 7.00 am and 9.00 am;
- 10 minutes (i.e. 6 trains/hour) between 9.00 am and 12.00 pm.

As first application, the passenger arrival rate at station (which is the result of the interaction between the Pre-Platform and the Supply Model) is not investigated in detail.

Stations	Pis	Chi	Fru	Col	Pol	Rio	Mon	Med	Van	Qua	Sal	Mat	Mus	Dan
Pis	0.00	1.21	1.81	1.21	6.03	1.21	4.82	6.03	9.04	4.22	1.21	1.21	10.25	12.05
Chi	0.81	0.00	2.11	1.41	7.74	1.41	5.63	8.44	10.55	4.92	2.11	2.11	10.55	13.37
Fru	0.08	0.05	0.00	0.12	0.46	0.15	0.15	0.31	0.62	0.15	0.12	0.06	0.31	0.62
Col	0.23	0.14	0.09	0.00	0.15	0.25	0.25	0.75	1.24	0.40	0.15	0.25	1.00	0.55
Pol	0.25	0.15	0.03	0.08	0.00	0.08	0.31	0.78	0.78	0.39	0.20	0.12	0.59	0.67
Rio	1.20	0.80	0.27	0.27	0.13	0.00	1.61	3.21	4.01	0.80	0.48	0.32	2.41	3.21
Mon	0.22	0.21	0.06	0.06	0.11	0.07	0.00	0.13	0.76	0.25	0.20	0.20	0.50	0.48
Med	4.02	2.30	0.80	0.92	1.15	1.15	1.15	0.00	4.49	1.50	1.50	1.20	2.99	3.29
Van	1.89	1.51	0.38	0.38	1.51	0.38	0.38	1.14	0.00	0.41	0.81	0.81	3.66	2.44
Qua	3.67	3.67	0.44	0.73	1.47	0.59	1.17	0.73	2.20	0.00	2.64	2.64	3.97	3.97
Sal	0.49	0.49	0.12	0.12	0.37	0.10	0.07	0.20	0.24	0.24	0.00	0.80	1.99	1.20
Mat	0.10	0.15	0.03	0.03	0.03	0.03	0.03	0.19	0.24	0.05	0.10	0.00	2.90	1.93
Mus	1.23	3.68	0.49	0.49	4.90	0.25	0.25	3.68	3.68	2.45	2.45	0.98	0.00	42.29
Dan	0.98	1.95	0.39	0.39	1.95	0.39	0.39	1.95	1.95	0.78	0.78	0.78	6.84	0.00

Table 6.1 Passenger arrival rates during rush hours.

Two different arrival rates are therefore assumed for analysing rush and weak hours, whose numerical values are synthesised in Table 6.1 and 6.2.

The input data of the OPM 1.0 model, which are considered in this application, can be summarised as follows:

- maximum number of passengers per coach: 432 passengers;
- fleet composition: all trains are composed of 3 traction units;
- maximum number of passengers per train: 1296 passengers;
- $\beta_{waiting}$ equal to 2.5, $\beta_{on\text{-}board}$ constant and equal to 1 and VOT equal to 5,00 €/h (These values are drawn from previous surveys proposed in the literature (Wardman, 2004) and highlights the fact that generally waiting time is almost three times more burdensome than running time.)

For each failure scenario simulated, three different recovery solutions are analysed:

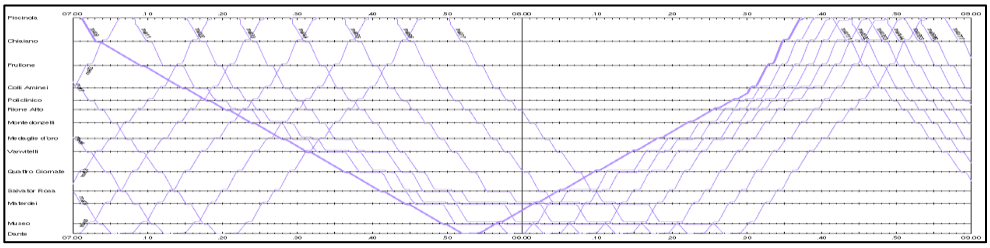
- the faulty train continues the service until Medaglie d'Oro and then, after unloading passengers on the platform, it is driven onto the maintenance track;
- the faulty train continues the service until the following terminus, i.e. Dante, and is then driven onto the maintenance track (indeed, one platform is available since Dante station corresponds to the terminus of the first section);
- the faulty train continues the whole service until the depot, i.e. Piscinola.

Stations	Pis	Chi	Fru	Col	Pol	Rio	Mon	Med	Van	Qua	Sal	Mat	Mus	Dan
Pis	0.00	0.60	0.90	0.60	3.01	0.60	2.41	3.01	4.52	2.11	0.60	0.60	5.12	6.03
Chi	0.40	0.00	1.06	0.70	3.87	0.70	2.81	4.22	5.28	2.46	1.06	1.06	5.28	6.68
Fru	0.04	0.03	0.00	0.06	0.23	0.08	0.08	0.15	0.31	0.08	0.06	0.03	0.15	0.31
Col	0.12	0.07	0.05	0.00	0.07	0.12	0.12	0.37	0.62	0.20	0.07	0.12	0.50	0.27
Pol	0.13	0.08	0.01	0.04	0.00	0.04	0.16	0.39	0.39	0.20	0.10	0.06	0.29	0.33
Rio	0.60	0.40	0.13	0.13	0.07	0.00	0.80	1.61	2.01	0.40	0.24	0.16	1.20	1.61
Mon	0.11	0.10	0.03	0.03	0.05	0.04	0.00	0.06	0.38	0.13	0.10	0.10	0.25	0.24
Med	2.01	1.15	0.40	0.46	0.57	0.57	0.57	0.00	2.24	0.75	0.75	0.60	1.50	1.65
Van	0.95	0.76	0.19	0.19	0.76	0.19	0.19	0.57	0.00	0.20	0.41	0.41	1.83	1.22
Qua	1.83	1.83	0.22	0.37	0.73	0.29	0.59	0.37	1.10	0.00	1.32	1.32	1.98	1.98
Sal	0.24	0.24	0.06	0.06	0.18	0.05	0.04	0.10	0.12	0.12	0.00	0.40	1.00	0.60
Mat	0.05	0.07	0.01	0.01	0.01	0.01	0.01	0.10	0.12	0.02	0.05	0.00	1.45	0.97
Mus	0.61	1.84	0.25	0.25	2.45	0.12	0.12	1.84	1.84	1.23	1.23	0.49	0.00	21.15
Dan	0.49	0.98	0.20	0.20	0.98	0.20	0.20	0.98	0.98	0.39	0.39	0.39	3.42	0.00

Table 6.2 Passenger arrival rates during weak hours.

In all cases, no spare trains are included into the service and hence, the timetable has to be revised. In particular, when a train breaks down, it increases the headway with the preceding convoy because it travels with a lower speed than the previous one. Likewise, the faulty train decreases the headway with the following convoy because the following train initially travels with a higher speed than the faulty one. Signalling systems tend to ensure that the following train then travels at the same speed as the faulty train. The slowing-down wave spreads progressively over all following trains.

In terms of travel demand, an increase in headways provides an increase in boarding passengers and a possible exceeding of train capacity. Likewise, a decrease in headways provides a decrease in passengers boarding trains. Therefore, the faulty train tends to be saturated by passengers, while the



following convoys tend to be empty except for people who were unable to board the previous train.

Figure 6.5 The slowing-down wave spreading progressively over all following trains

When the faulty train is eliminated from the service earlier than planned (i.e. it reaches the maintenance track), if there are passengers on board, they have to alight onto the appropriate platform and wait for another run. Therefore, train elimination will increase network performance, since following convoys are not constrained anymore by the signalling system and are able to reach their maximum speed, but will produce a combined deterioration in service quality since user waiting times and vehicle crowding will increase.

Results of the six strategies are compared in Table 6.3 and 6.4 with the ordinary service.

		Regular service scenario	Strategy 1 (Medaglie d'Oro station)	Strategy 2 (Dante station)	Strategy 3 (Piscinola station)
User generalised costs [k€]		80.942	88.432	**85.823**	86.214
peak-hour time period	minimum headway [minutes]	7.00	5.55	5.02	3.17
	maximum headway [minutes]	7.00	14.00	14.00	14.00
weak-hour time period	minimum headway [minutes]	10.00	10.00	10.00	10.00
	maximum headway [minutes]	10.00	10.00	10.00	10.00

Table 6.3 Strategy analysis with 20% reduction in train performance (source: D'Acierno et al., 2012)

In the first scenario (Table 6.3), performance reduction does not lead to substantial increases in user travel times. Hence the best strategy in this case is to complete the service until Dante station then drive the faulty train onto the maintenance track. Indeed, the backward trip (i.e. from Dante to Piscinola) during the morning rush hour is not affected by a large number of passengers and removing a run does not produce a high increase of user generalised cost. In addition, it is worth noting that this strategy is not optimal in terms of service availability (i.e. punctuality). In fact, it produces higher delays than strategy 1 since following trains are constrained to a degraded speed for a longer section (see the difference of minimum headways).

		Regular service scenario	Strategy 1 (Medaglie d'Oro station)	Strategy 2 (Dante station)	Strategy 3 (Piscinola station)
User generalised costs [k€]		80.942	**103.931**	136.106	192.524
peak-hour time period	minimum headway [minutes]	7.00	1.28	20.00	20.00
	maximum headway [minutes]	7.00	20.20	33.87	51.47
weak-hour time period	minimum headway [minutes]	10.00	10.00	4.80	1.92
	maximum headway [minutes]	10.00	10.00	6.02	5.47

Table 6.4 Strategy analysis with 80% reduction in train performance (source: D'Acierno et al., 2012)

In the second scenario (Table 6.4), performance reduction is considerable and large increases are generated in user travel times. Therefore, the best strategy consists in excluding the faulty train from the system as soon as possible so as to let the service reach the ordinary conditions in less time than the other cases.
The application demonstrated that it cannot be stated a priori (i.e. without any model implementation) which is the optimal operational strategy. Looking at the results, it comes to light that, especially when the performance reduction is not considerable, minimising users' discomfort can bring to adopt solutions far from the ones which are optimal in terms of service availability.

However, these results have to be investigated deeper so as to produce reliable conclusions. In fact, each solution can be influenced by travel demand levels, breakdown severity and fleet composition. Other applications are therefore necessary to establish with more accuracy the optimal recovery strategies.
To this purpose, the same application was performed more times limiting the maximum train speed in each scenario to a value between 10 % and 80 %. In addition, all the three maintenance tracks (at Colli Aminei, Medaglie d'Oro and

Dante) within the network are considered. As a consequence, six strategies are now feasible, namely:

- the train continues the service until Colli Aminei or Medaglie d’Oro and then, after unloading passengers on the platform, it is driven onto the nearest maintenance track;
- the train continues the service until it reaches the following terminus, i.e. Dante, and is then driven onto the maintenance track;
- the train completes the outward trip and starts the return trip until Medaglie d’Oro or Colli Aminei and then, after unloading passengers on the platform, it is driven onto the nearest maintenance track;
- the train completes the whole service until it reaches the depot, i.e. Piscinola.

Even in this case, no spare trains are included into the service. Thus, the runs which had to be performed by the faulty train are cancelled, extending passengers’ discomfort to off peak-hours. Furthermore, in order to highlight the importance of considering capacity constraints of rail convoys, numerical applications were performed twice, adopting also previous models proposed in the literature (Mazzeo et al., 2011; Quaglietta et al., 2011) based on a more simple formulation.

Tables 6.5 and 6.6 provide the objective function values (i.e. total travel time of passengers), expressed in terms of equivalent monetary costs, for each intervention strategy and for each speed reduction respectively without and with convoy capacity constraints. Bold values represent the minimum of the objective function for each breakdown occurring.

	Speed reductions							
Stations	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**
ColliAminei (outward)	€ 130,567	€ 130,533	€ 130,487	€ 130,425	€ 130,332	**€ 130,190**	**€ 130,352**	**€ 135,900**
Medaglied'Oro (outward)	€ 130,198	€ 130,185	€ 130,084	€ 130,047	**€ 129,960**	€ 130,343	€ 136,445	€ 161,676
Dante	€ 128,218	€ 128,247	€ 128,329	€ 128,571	€ 130,567	€ 137,883	€ 252,828	€ 470,171
Medaglied'Oro (backward)	€ 128,204	€ 128,257	€ 128,468	€ 129,141	€ 133,132	€ 147,756	€ 259,682	€ 496,950
ColliAminei (backward)	€ 127,662	€ 127,699	€ 128,098	€ 128,394	€ 133,001	€ 154,780	€ 272,416	€ 584,778
Piscinola (depot)	**€ 127,338**	**€ 127,374**	**€ 127,507**	**€ 128,171**	€ 135,757	€ 167,713	€ 252,762	€ 467,591

Table 6.5 Total travel passenger costs by neglecting capacity constraints of rail convoys (source: D'Acierno et al., 2013a).

Hence these values allow the optimal intervention strategy to be identified for each failure scenario. Figure 6.6 by contrast, provides a comparison of objective function values by adopting for each model (i.e. neglecting and considering capacity constraints) two different scales in order to highlight trends of the function in the neighbourhood of optimal values.

	Speed reductions							
Stations	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**
ColliAminei (outward)	€ 134,173	€ 134,140	€ 134,094	€ 133,469	€ 133,940	€ 133,799	**€ 133,420**	**€ 139,214**
Medaglied'Oro (outward)	€ 132,668	€ 132,654	€ 132,604	€ 132,559	€ 132,472	**€ 132,785**	€ 139,434	€ 171,950
Dante	€ 130,097	€ 130,126	€ 130,209	€ 130,451	€ 132,465	€ 141,936	€ 276,895	€ 649,768
Medaglied'Oro (backward)	€ 129,836	€ 129,816	€ 129,903	€ 131,112	**€ 132,454**	€ 142,794	€ 274,313	€ 650,892
ColliAminei (backward)	€ 129,618	€ 129,270	€ 129,698	**€ 130,029**	€ 132,661	€ 147,741	€ 276,047	€ 649,826
Piscinola (depot)	**€ 129,217**	**€ 129,253**	**€ 129,386**	€ 130,068	€ 137,833	€ 175,379	€ 276,549	€ 646,473

Table 6.6 Total travel passenger costs by considering capacity constraints of rail convoys (source: D'Acierno et al., 2013a).

Obviously, since there is a great difference between the minimum and maximum values of objective function and our aim is to identify the minimum of the objective function, it is represented only the part of the objective functions below the threshold of € 170,000 (in the upper part of the figure) and € 140,000 (in the lower part of the figure), not indicating higher values in the figure.

In terms of data analysis, it is worth noting that in some cases, the objective function has more than one local minimum (i.e. the objective function is not convex). Moreover, as expected, by considering capacity constraints, the objective function has values greater than neglecting them since some passengers would not be able to board the first arriving convoy and therefore have to wait for the followings, increasing their waiting times. These differences between the two approaches could provide different optimal strategies. Obviously, the adoption of capacity constraints yields an estimation of user disutility and hence identifies the optimal strategy closer to the real phenomenon. The simplified model by contrast, tends to calm down the negative effects of the failure in less time than it would be necessary. Indeed, only in extreme conditions (i.e. speed reductions lower than 30% or higher than 70%) do both approaches provide the same optimal strategy.

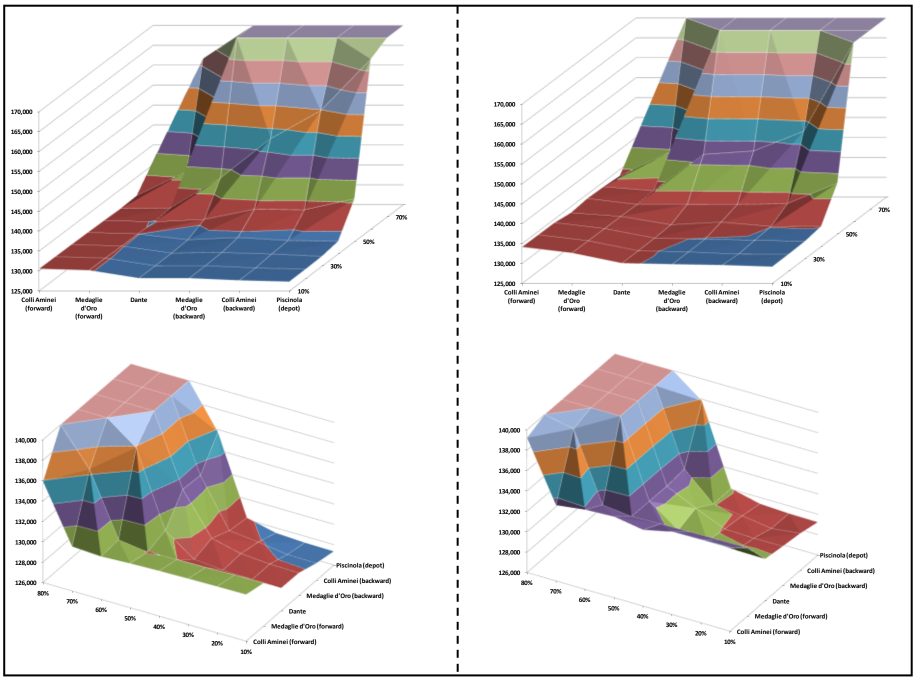


Figure 6.6 Objective function values by neglecting (left) and considering (right) capacity constraints of rail convoys (source: D'Acierno et al., 2013a)

Focusing on the numerical results, in the case of speed reductions between 10% and 40% (between 10% and 50 % in the case of the unconstrained approach), the application of the Colli Aminei (outward) as well as Medaglie d'Oro (outward) strategy provides a slight reduction in objective function values by increasing failure severity. This is due to the fact that an increase in breakdown severity yields a decrease in faulty train speed which generates a decrease in headway between this train and the following. Hence, there is an increase in travel times for passengers on the faulty train combined with a decrease in waiting time for the following rail convoys, once passengers are unloaded onto the platform. Since, as already stated, the discomfort perceived waiting on the platform is greater than the one experienced on board the train, a slight increase in travel times is more than compensated by the reduction in waiting times. However, this effect does not take place when breakdown severities, and therefore increases in travel times, are significant. Indeed, especially in the case of the constrained approach, waiting times could be higher, since some passengers might not be able to board the first arriving train.

A common result between the two approaches is that, in terms of optimal strategy, when the faulty train is fast (low reduction in maximum speed), it is best to conclude the trip at the depot so as to avoid passenger discomfort caused by alighting from the faulty train and boarding the following train. Likewise, when the faulty train is excessively slow (great reduction in maximum speed), it is best to position the faulty train on a maintenance track as soon as possible. However, the constrained approach is to prefer inasmuch as it provides reliable results and feasible strategies with any kind of performance reduction.

Finally, the application demonstrated that recovery solutions are influenced by the breakdown severity level, since points of convenience are different according to the simulated scenario.

In order to analyse also the effects of different travel demand rates on operational strategies, the method was then applied twelve times multiplying the estimated travel demand by a value varying between 10% and 120%, keeping the same input data and the same recovery solutions. In addition, the transposed travel demand (i.e. the return home demand) in the case of current travel demand (i.e. multiplier equal to 100%) was also considered so as to assess possible changes in the convenience of the optimal solution.

All the results are synthesised in Table 6.7–6.10. In particular, Table 6.7 provides values of the objective function (4.3) in the case of regular service;

Table 6.8 shows objective function values in the case of transposed travel demand; Table 6.9 and 6.10 presents generalised costs of passengers for each breakdown, for each demand level and each intervention strategy; likewise, bold values represent the minimum value of the objective function (i.e. identify the optimal strategy) for each breakdown and demand level occurring.

Analysing the numerical results, it is worth noting that, as already shown in the previous application, the objective function is not convex. In fact, in some cases, the function has more than one local minimum. As far as the influence of travel demand levels is concerned, it is possible to identify optimal intervention strategies which do not depend formally on travel demand. Indeed, for breakdown severities up to 30%, the faulty train is fast enough to make the final depot arrival (i.e. Piscinola) always the optimal strategy so as to avoid passenger discomfort caused by alighting from the faulty train and boarding the following train.

Travel demand	***10%***	***20%***	***30%***	***40%***	***50%***	***60%***	***70%***	***80%***	***90%***	***100%***	***110%***	***120%***
User costs	€ 11,819	€ 23,638	€ 35,457	€ 47,276	€ 59,092	€ 70,915	€ 82,735	€ 94,555	€ 106,371	€ 118,191	€ 130,015	€ 141,823

Table 6.7 Total travel passenger costs in the case of regular service (source: (source: D'Acierno et al., 2013a).

Interventionstrategy	***10%***	***20%***	***30%***	***40%***	***50%***	***60%***	***70%***	***80%***
Colli Aminei (outward)	€ 134,190	€ 134,181	€ 134,170	€ 134,154	€ 134,131	**€ 134,095**	**€ 134,372**	**€ 142,482**
Medaglie d'Oro (outward)	€ 134,055	€ 134,051	€ 134,012	€ 134,001	**€ 133,973**	€ 134,524	€ 143,632	€ 181,848
Dante	€ 133,350	€ 133,358	€ 133,385	€ 133,469	€ 136,202	€ 149,599	€ 302,656	€ 810,122
Medaglie d'Oro (backward)	€ 131,986	€ 131,988	€ 131,993	€ 132,265	€ 136,614	€ 159,114	€ 319,206	€ 742,429
Colli Aminei (backward)	€ 131,643	€ 129,936	€ 131,678	€ 132,434	€ 139,091	€ 170,229	€ 350,910	€ 710,605
Piscinola (depot)	**€ 129,206**	**€ 129,244**	**€ 129,427**	**€ 130,423**	€ 141,323	€ 188,076	€ 297,346	€ 626,543

Table 6.8 Total travel passenger costs in the case of return home travel demand (source: D'Acierno et al., 2013a).

Likewise, for breakdown severity no lower than 70%, the faulty train is excessively slow, and it is convenient to place it on a maintenance track as soon as possible (i.e. Colli Aminei outward). This phenomenon is evident even for the return home travel demand. In all other cases, i.e. when breakdown severity is greater than 30% and lower than 70%, different speed reduction values yield different points of convenience depending on travel demand levels. Therefore, intervention strategies are not invariant with respect to the arrival rate at stations. This highlights one more time the importance of simulating passengers travel choices for providing reliable recovery solutions.

Travel demand multiplier	***Intervention strategy***	***10%***	***20%***	***30%***	***40%***	***50%***	***60%***	***70%***	***80%***
10%	Colli Aminei (forward)	€ 13,054	€ 13,051	€ 13,047	€ 13,040	€ 13,031	**€ 13,017**	**€ 13,033**	**€ 13,586**
	Medaglie d'Oro (forward)	€ 13,013	€ 13,012	€ 13,002	€ 12,998	**€ 12,989**	€ 13,026	€ 13,633	€ 16,149
	Dante	€ 12,822	€ 12,825	€ 12,833	€ 12,857	€ 13,057	€ 13,788	€ 25,284	€ 47,021
	Medaglie d'Oro (backward)	€ 12,780	€ 12,783	€ 12,791	€ 12,821	€ 13,129	€ 13,068	€ 25,010	€ 46,841
	Colli Aminei (backward)	€ 12,767	€ 12,770	€ 12,779	**€ 12,812**	€ 13,182	€ 14,994	€ 25,182	€ 46,816
	Piscinola (depot)	**€ 12,734**	**€ 12,737**	**€ 12,751**	€ 12,817	€ 13,576	€ 16,772	€ 25,277	€ 46,763
20%	Colli Aminei (forward)	€ 26,109	€ 26,102	€ 26,093	€ 26,081	€ 26,062	**€ 26,034**	**€ 26,066**	**€ 27,172**
	Medaglie d'Oro (forward)	€ 26,027	€ 26,024	€ 26,004	€ 25,996	**€ 25,978**	€ 26,054	€ 27,268	€ 32,299
	Dante	€ 25,644	€ 25,650	€ 25,667	€ 25,715	€ 26,114	€ 27,577	€ 50,668	€ 94,953
	Medaglie d'Oro (backward)	€ 25,560	€ 25,566	€ 25,583	€ 25,642	€ 26,258	€ 28,753	€ 50,120	€ 94,594
	Colli Aminei (backward)	€ 25,534	€ 25,540	€ 25,560	**€ 25,625**	€ 26,364	€ 29,989	€ 50,467	€ 94,549
	Piscinola (depot)	**€ 25,468**	**€ 25,476**	**€ 25,502**	€ 25,635	€ 27,152	€ 33,544	€ 50,660	€ 94,437
30%	Colli Aminei (forward)	€ 39,164	€ 39,154	€ 39,140	€ 39,121	€ 39,093	**€ 39,050**	**€ 39,099**	**€ 40,757**
	Medaglie d'Oro (forward)	€ 39,040	€ 39,036	€ 39,006	€ 38,994	**€ 38,967**	€ 39,080	€ 40,901	€ 48,449
	Dante	€ 38,466	€ 38,475	€ 38,499	€ 38,572	€ 39,171	€ 41,366	€ 76,430	€ 144,068
	Medaglie d'Oro (backward)	€ 38,340	€ 38,348	€ 38,374	€ 38,462	€ 39,387	€ 43,129	€ 75,608	€ 143,529
	Colli Aminei (backward)	€ 38,301	€ 38,310	€ 38,339	**€ 38,437**	€ 39,546	€ 44,984	€ 76,127	€ 143,459
	Piscinola (depot)	**€ 38,202**	**€ 38,213**	**€ 38,253**	€ 38,452	€ 40,728	€ 50,356	€ 76,442	€ 143,294
40%	Colli Aminei (forward)	€ 52,217	€ 52,204	€ 52,185	€ 52,160	€ 52,123	**€ 52,066**	**€ 52,130**	**€ 54,342**
	Medaglie d'Oro (forward)	€ 52,053	€ 52,047	€ 52,007	€ 51,991	**€ 51,955**	€ 52,106	€ 54,534	€ 64,891
	Dante	€ 51,287	€ 51,299	€ 51,332	€ 51,429	€ 52,227	€ 55,153	€ 102,729	€ 194,268
	Medaglie d'Oro (backward)	€ 51,119	€ 51,131	€ 51,164	€ 51,282	€ 52,515	€ 57,546	€ 101,634	€ 193,550
	Colli Aminei (backward)	€ 51,067	€ 51,079	€ 51,118	**€ 51,248**	€ 52,727	€ 60,068	€ 102,327	€ 193,459
	Piscinola (depot)	**€ 50,935**	**€ 50,950**	**€ 51,003**	€ 51,269	€ 54,303	€ 67,423	€ 102,695	€ 193,236
50%	Colli Aminei (forward)	€ 65,268	€ 65,251	€ 65,228	€ 65,198	€ 65,151	**€ 65,080**	**€ 65,160**	**€ 67,924**
	Medaglie d'Oro (forward)	€ 65,063	€ 65,056	€ 65,005	€ 64,986	**€ 64,941**	€ 65,129	€ 68,165	€ 81,481
	Dante	€ 64,106	€ 64,120	€ 64,162	€ 64,283	€ 65,280	€ 69,115	€ 129,631	€ 246,953
	Medaglie d'Oro (backward)	€ 63,896	€ 63,910	€ 63,952	€ 64,099	€ 65,641	€ 72,158	€ 128,263	€ 246,048
	Colli Aminei (backward)	€ 63,831	€ 63,845	€ 63,894	**€ 64,057**	€ 65,905	€ 75,229	€ 129,128	€ 246,071
	Piscinola (depot)	**€ 63,666**	**€ 63,684**	**€ 63,750**	€ 64,083	€ 67,963	€ 84,487	€ 129,535	€ 245,638
60%	Colli Aminei (forward)	€ 78,326	€ 78,306	€ 78,278	€ 78,241	€ 78,185	**€ 78,099**	**€ 78,196**	**€ 81,756**
	Medaglie d'Oro (forward)	€ 78,080	€ 78,072	€ 78,011	€ 77,988	**€ 77,934**	€ 78,159	€ 82,093	€ 98,233
	Dante	€ 76,932	€ 76,949	€ 76,998	€ 77,144	€ 78,352	€ 83,239	€ 157,239	€ 301,264
	Medaglie d'Oro (backward)	€ 76,680	€ 76,696	€ 76,747	€ 76,924	€ 78,914	€ 86,778	€ 155,596	€ 300,605
	Colli Aminei (backward)	€ 76,601	€ 76,619	€ 76,677	**€ 76,873**	€ 79,265	€ 90,397	€ 156,636	€ 300,529
	Piscinola (depot)	**€ 76,404**	**€ 76,425**	**€ 76,505**	€ 76,904	€ 81,680	€ 101,929	€ 157,083	€ 299,683

Table 6.9 Total travel passenger costs in the case of disturbed service with travel demand level from 10% to 60% (source: (source: D’Acierno et al., 2013a).

Travel demand multiplier	Intervention strategy	10%	20%	30%	40%	50%	**60%**	**70%**	**80%**
70%	Colli Aminei (forward)	€ 91,508	€ 91,484	€ 91,452	€ 91,408	€ 91,343	**€ 91,243**	**€ 91,357**	**€ 95,680**
	Medaglie d'Oro (forward)	€ 91,182	€ 91,172	€ 91,101	€ 91,074	**€ 91,011**	€ 91,287	€ 96,024	€ 115,652
	Dante	€ 89,842	€ 89,862	€ 89,920	€ 90,090	€ 91,676	€ 97,361	€ 185,851	€ 371,053
	Medaglie d'Oro (backward)	€ 89,548	€ 89,568	€ 89,627	€ 89,833	€ 92,314	€ 101,476	€ 183,934	€ 370,677
	Colli Aminei (backward)	€ 89,457	€ 89,477	€ 89,545	**€ 89,774**	€ 92,691	€ 105,876	€ 185,147	€ 370,480
	Piscinola (depot)	**€ 89,226**	**€ 89,252**	**€ 89,345**	€ 89,878	€ 95,396	€ 119,714	€ 185,657	€ 369,134
80%	Colli Aminei (forward)	€ 105,179	€ 105,152	€ 105,115	€ 105,065	€ 104,990	€ 104,877	**€ 104,992**	**€ 109,640**
	Medaglie d'Oro (forward)	€ 104,622	€ 104,612	€ 104,524	€ 104,493	**€ 104,422**	**€ 103,012**	€ 110,034	€ 133,659
	Dante	€ 102,986	€ 103,009	€ 103,075	€ 103,269	€ 105,004	€ 111,777	€ 215,284	€ 449,293
	Medaglie d'Oro (backward)	€ 102,650	€ 102,672	€ 102,740	€ 102,975	€ 105,715	€ 116,684	€ 213,085	€ 449,590
	Colli Aminei (backward)	€ 102,546	€ 102,569	€ 102,647	**€ 102,908**	€ 106,118	€ 121,683	€ 214,487	€ 448,901
	Piscinola (depot)	**€ 102,282**	**€ 102,311**	**€ 102,417**	€ 103,005	€ 109,114	€ 137,852	€ 215,066	€ 447,091
90%	Colli Aminei (forward)	€ 118,898	€ 118,867	€ 118,826	€ 118,770	€ 118,686	€ 118,558	**€ 118,656**	**€ 123,909**
	Medaglie d'Oro (forward)	€ 118,227	€ 118,215	€ 118,117	€ 118,082	**€ 118,002**	**€ 116,374**	€ 124,388	€ 152,192
	Dante	€ 116,152	€ 116,178	€ 116,252	€ 116,470	€ 118,355	€ 126,431	€ 245,773	€ 549,282
	Medaglie d'Oro (backward)	€ 115,774	€ 116,497	€ 115,875	€ 116,140	€ 119,298	€ 132,186	€ 243,288	€ 549,973
	Colli Aminei (backward)	€ 115,657	€ 115,683	€ 115,770	**€ 116,064**	€ 119,727	€ 137,800	€ 244,918	€ 549,066
	Piscinola (depot)	**€ 115,360**	**€ 115,393**	**€ 115,512**	€ 116,154	€ 123,273	€ 156,339	€ 245,477	€ 546,562
100%	Colli Aminei (forward)	€ 134,173	€ 134,140	€ 134,094	€ 133,469	€ 133,940	€ 133,799	**€ 133,420**	**€ 139,214**
	Medaglie d'Oro (forward)	€ 132,668	€ 132,654	€ 132,604	€ 132,559	€ 132,472	**€ 132,785**	€ 139,434	€ 171,950
	Dante	€ 130,097	€ 130,126	€ 130,209	€ 130,451	€ 132,465	€ 141,936	€ 276,895	€ 649,768
	Medaglie d'Oro (backward)	€ 129,836	€ 129,816	€ 129,903	€ 131,112	**€ 132,454**	€ 142,794	€ 274,313	€ 650,892
	Colli Aminei (backward)	€ 129,618	€ 129,270	€ 129,698	**€ 130,029**	€ 132,661	€ 147,741	€ 276,047	€ 649,826
	Piscinola (depot)	**€ 129,217**	**€ 129,253**	**€ 129,386**	€ 130,068	€ 137,833	€ 175,379	€ 276,549	€ 646,473
110%	Colli Aminei (forward)	€ 150,473	€ 150,436	€ 150,385	€ 150,305	€ 150,201	€ 150,045	**€ 150,110**	**€ 156,216**
	Medaglie d'Oro (forward)	€ 148,501	€ 148,486	€ 148,358	€ 148,222	€ 148,123	**€ 148,417**	€ 155,598	€ 262,967
	Dante	€ 145,209	€ 145,241	€ 145,332	€ 145,586	**€ 147,783**	€ 157,926	€ 320,086	€ 794,932
	Medaglie d'Oro (backward)	€ 144,747	€ 144,778	€ 144,871	€ 146,212	€ 150,916	€ 169,852	€ 327,922	€ 764,755
	Colli Aminei (backward)	€ 144,604	€ 144,636	€ 144,743	€ 145,762	€ 150,920	€ 175,487	€ 327,922	€ 822,529
	Piscinola (depot)	**€ 144,241**	**€ 144,281**	**€ 144,427**	**€ 145,165**	€ 153,431	€ 194,914	€ 319,645	€ 727,877
120%	Colli Aminei (forward)	€ 171,653	€ 171,613	€ 171,557	€ 171,483	€ 171,369	€ 171,199	**€ 171,274**	**€ 327,379**
	Medaglie d'Oro (forward)	€ 166,163	€ 166,146	€ 165,998	€ 165,952	€ 165,844	**€ 167,287**	€ 175,407	€ 641,914
	Dante	€ 161,761	€ 161,796	€ 161,895	€ 162,185	**€ 164,653**	€ 176,097	€ 377,005	€ 986,174
	Medaglie d'Oro (backward)	€ 161,744	€ 161,809	€ 162,073	€ 162,869	€ 167,919	€ 188,779	€ 385,813	€ 829,275
	Colli Aminei (backward)	€ 161,333	€ 161,390	€ 161,626	€ 162,377	€ 167,901	€ 194,758	€ 385,812	€ 1,016,550
	Piscinola (depot)	**€ 160,705**	**€ 160,749**	**€ 160,908**	**€ 161,736**	€ 171,011	€ 216,143	€ 376,495	€ 787,093

Table 6.10 Total travel passenger costs in the case of disturbed service with travel demand level from 70% to 120% (source: (source: D’Acierno et al., 2013a).

Finally, the influence of fleet composition on users’ disutilities during disruption events has to be studied. In previous applications, it was assumed that all trains performing the service were composed of three traction units. Actually, METRONAPOLI does not have enough rolling stock to satisfy this requirement. Indeed, due to maintenance and regulation enforcements the vehicles available amount to 36 traction units per day. Hence, the service enterprise is forced to adopt, in the case of operations with 9 convoys, 6 tripleheader trains (i.e. with 3 traction units) and 3 doubleheader convoys (i.e. with 2 traction units). It is trivial to conclude that the best operational strategy consists in adopting a 3-3-2 convoy sequence and that the removal from service of a triple-header convoy generates a greater impact than a double-header convoy. Due to the randomness of the breakdown phenomenon, previous failure scenarios are simulated adopting three kinds of convoy sequences, i.e. 3-3-2, 3-2-3 and 2-3-3. Indeed, it is not possible to establish a priori which train will be the first to start in the daily service and/or which train will undergo the breakdown. Results of the application are shown in Table 6.11, 6.12, 6.13.

	Speed reductions							
Stations	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**
ColliAminei (outward)	€ 138,912	€ 138,879	€ 138,833	€ 138,771	€ 138,677	€ 138,535	**€ 138,591**	**€ 144,038**
Medaglied'Oro (outward)	€ 136,175	€ 136,158	€ 136,036	€ 135,992	€ 135,899	**€ 136,164**	€ 142,729	€ 178,184
Dante (terminus)	€ 132,385	€ 132,401	€ 132,464	€ 132,691	**€ 134,720**	€ 144,280	€ 294,519	€ 669,219
Medaglied'Oro (backward)	€ 131,962	€ 131,978	€ 132,043	€ 132,323	€ 135,756	€ 150,967	€ 291,621	€ 670,140
ColliAminei (backward)	€ 131,832	€ 131,849	€ 131,927	**€ 132,238**	€ 136,196	€ 156,771	€ 293,342	€ 668,924
Piscinola (depot)	**€ 131,504**	**€ 131,528**	**€ 131,642**	€ 132,322	€ 140,022	€ 177,785	€ 293,727	€ 665,037

Table 6.11 Total travel passenger costs in the case of convoy sequence 3-3-2 (source: Placido et al., 2014c).

	Speed reductions							
Stations	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**
ColliAminei (outward)	€ 138,022	€ 138,001	€ 137,972	€ 137,932	€ 137,871	€ 137,775	**€ 137,904**	**€ 143,406**
Medaglied'Oro (outward)	€ 136,679	€ 136,691	€ 136,614	€ 136,640	€ 136,658	**€ 137,062**	€ 143,566	€ 176,684
Dante (terminus)	€ 133,760	€ 133,795	€ 133,884	€ 134,125	**€ 136,186**	€ 145,328	€ 294,904	€ 660,061
Medaglied'Oro (backward)	€ 133,357	€ 133,392	€ 133,482	€ 133,775	€ 137,106	€ 152,077	€ 293,052	€ 662,653
ColliAminei (backward)	€ 133,227	€ 133,262	€ 133,365	**€ 133,691**	€ 137,600	€ 157,939	€ 294,695	€ 661,476
Piscinola (depot)	**€ 132,880**	**€ 132,923**	**€ 133,061**	€ 133,757	€ 141,251	€ 179,310	€ 294,916	€ 657,345

Table 6.12 Total travel passenger costs in the case of convoy sequence 3-2-3 (source: Placido et al., 2014c).

Stations	Speed reductions 10%	20%	30%	40%	50%	60%	70%	80%
ColliAminei (outward)	€ 141,094	€ 141,059	€ 141,011	€ 140,945	€ 140,846	€ 140,697	**€ 140,769**	**€ 146,111**
Medaglied'Oro (outward)	€ 138,724	€ 138,710	€ 138,579	€ 138,540	€ 138,445	**€ 138,743**	€ 145,089	€ 346,291
Dante (terminus)	€ 135,359	€ 135,392	€ 135,484	€ 135,741	**€ 137,742**	€ 146,546	€ 301,816	€ 770,003
Medaglied'Oro (backward)	€ 135,331	€ 135,389	€ 135,618	€ 136,298	€ 140,527	€ 157,633	€ 308,832	€ 699,452
ColliAminei (backward)	€ 134,989	€ 135,040	€ 135,246	€ 135,888	€ 140,502	€ 162,847	€ 309,037	€ 794,751
Piscinola (depot)	**€ 134,479**	**€ 134,520**	**€ 134,661**	**€ 135,389**	€ 142,757	€ 180,315	€ 301,504	€ 665,935

Table 6.13 Total travel passenger costs in the case of convoy sequence 2-3-3 (source: Placido et al., 2014c).

The main result is that optimal intervention strategies do not depend formally on convoy sequence since, except from one case, solutions producing lower generalised costs are the same. In particular, like the previous applications, for breakdown severities up to 30%, the optimal strategy is always to complete the whole service as far as the depot. Likewise, for breakdown severity no lower than 70%, the optimal strategy is always to drive the faulty train onto the maintenance track at Colli Aminei (outward strategy). In all other cases, i.e. when breakdown severity is greater than 30% and lower than 70%, apparently the different convoy sequences do not affect the points of convenience. The user generalised costs suggest recovering the faulty train on the maintenance track in Medaglie d'Oro during the backward trip (40% of speed reduction), in Dante (50% of speed reduction) and in Medaglie d'Oro again but during the outward trip (60% of speed reduction). Actually, in the case of fleet sequence 2-3-3 and 40% of speed reduction, the optimal recovery strategy is slightly dissimilar (i.e. completion of the service up to depot) from the one obtained with the other two fleet compositions (i.e. stop the train at Colli Aminei during the backward trip).

In conclusion, this paragraph proved that the DSS proposed in this thesis, by considering user generalised cost, can be useful for managing the rail network maximising the service quality perceived by passengers. Furthermore, although effects on travel demand are often neglected in the literature, the applications showed that this assumption is correct only in extreme cases (i.e. when the faulty convoy is fast enough or extremely slow). In other circumstances, travel demand levels as well as breakdown contexts can influence the recovery strategies.

6.3 The estimation of the passenger arrival rate at stations

The ***PPM*** through the interaction with the ***SM*** has the target of evaluating the amount of flow arriving at stations. This input data is extremely important since, as already explained, it influences the optimal intervention strategies. Hence, a generic disruption can yield different intervention strategies according to the period of the day affected. The previous applications are based on simplified assumptions related to passenger flow. Indeed, just two travel demand matrices have been considered during the whole day, one concerning the peak hours and another one representing the weak hours. Obviously, a better estimation of the different arrival rates during the day is necessary. This can be achieved through the travel demand O-D estimation using traffic counts (Cascetta, 2009). In particular, these techniques have received considerable attention in recent years because of the great cost and complexity of sampling surveys as well as the lack of precision related to model estimators. On the other hand, users' flows within the network in particular sections can be obtained very easily and often automatically (e.g. counts at turnstiles). More in detail, the problem consists of estimating an O-D travel demand matrix which is close to the O-D seed matrix and generates path flows similar to the ones observed. This procedure under the usual assumption of a within-day static system has been largely studied. The within-day dynamic framework by contrast, increases the complexity of the problem and it is still a very recent topic. However, in the literature, there are valid examples of dynamic estimation of O-D flows from traffic counts (Cipriani et al., 2011; Cascetta et al., 2013).

In this application, different passenger arrival rates for the Line 1 system have been determined which resulted in a better estimation of the flow variation during the day. In particular, according to the sequential estimator procedure (see Cascetta, 2009 for more details), the dynamic of the daily service is taken into account dividing the whole reference period into different time intervals and estimating the O-D demand vector statically for each of them. Thanks to several survey campaigns, METRONAPOLI collected 9 months of turnstile data which provides for each station and for five time periods of the day (i.e. from 6.00 to 9.00; from 9.00 to 12.00; from 12.00 to 14.00; from 14.00 to 19.00; from 19.00 to 23.30) the number of crossing passengers. Obviously the station data do not allow the travel direction to be identified, except in the case of the terminus. Using the previous O-D flows (i.e. Table 6.1 and 6.2) as seed matrices, the new surveyed data enable the estimation of 5 different arrival rates for representing the travel demand pattern of the day. However, in order to determine several flow levels for each time period analysed, a new approach is proposed. Basically, instead of considering a single flow value (e.g. the average value or the maximum value of the sample), it is possible to determine the statistical distribution which fits as best as possible the surveyed data. In this way, passenger flows are random variables with a known distribution and thus, different travel demand percentiles can be adopted for the estimation of the O-D matrices. To this purpose, it is first necessary to clean the sample from irrelevant data related to singular days such as holidays, pre-holiday and strikes. Then, the probability density function which best describes the selected data is obtained through the following approach:

- three kinds of distribution functions are considered, namely the Gamma, the Gumbel and the Normal distribution;
- for each station and for each of the five time periods, all three distribution functions are calibrated which means that the distribution parameters are calculated so as to minimise differences between surveyed data and model data;
- for each calibrated distribution function (i.e. a function for each station and for each time period), the term p-value is calculated. This value specifies how close the statistical distribution is to the physical phenomenon;

- finally, the distribution with the higher p-values is selected. Obviously, the selection is related only to the functional form while function parameters differ for each station and each time period.

Table 6.14-6.16 summarise the p-values obtained for the three considered statistical distributions.

	ST 01	ST 02	ST 03	ST 04	ST 05	ST 06	ST 07	ST 08	ST 09	ST 10	ST 11	ST 12	ST 13	ST 14
TP1	0.16	0.00	0.70	0.98	0.79	0.63	0.37	0.07	0.11	0.20	0.16	0.05	0.62	0.12
TP2	0.68	0.50	0.37	0.35	0.92	0.37	0.81	0.00	0.95	0.90	0.25	0.01	0.85	0.62
TP3	0.22	0.41	0.08	0.04	0.66	0.74	0.82	0.52	0.71	0.60	0.99	0.13	0.27	0.97
TP4	0.84	0.66	0.47	0.85	0.57	0.74	0.30	0.70	0.52	0.34	0.29	0.08	0.63	0.07
TP5	0.55	0.85	0.12	0.17	0.28	0.64	0.99	0.20	0.26	0.90	0.66	0.16	0.75	0.03

Table 6.14. P-value terms in the case of Gamma distribution function (source: Ercolani et al., 2014).

	ST 01	ST 02	ST 03	ST 04	ST 05	ST 06	ST 07	ST 08	ST 09	ST 10	ST 11	ST 12	ST 13	ST 14
TP1	0.03	0.00	0.45	0.99	0.84	0.55	0.46	0.02	0.07	0.05	0.06	0.08	0.26	0.13
TP2	0.18	0.60	0.15	0.10	0.56	0.06	0.82	0.00	0.60	1.00	0.07	0.05	0.19	0.07
TP3	0.06	0.59	0.15	0.01	0.38	0.44	0.99	0.14	0.26	0.25	0.98	0.34	0.69	0.79
TP4	0.71	0.70	0.27	0.45	0.32	0.56	0.23	0.28	0.19	0.07	0.09	0.10	0.11	0.01
TP5	0.43	0.64	0.12	0.21	0.23	0.67	0.99	0.25	0.05	0.67	0.78	0.08	0.18	0.07

Table 6.15. P-value terms in the case of Gumbel distribution function (source: Ercolani et al., 2014).

	ST 01	ST 02	ST 03	ST 04	ST 05	ST 06	ST 07	ST 08	ST 09	ST 10	ST 11	ST 12	ST 13	ST 14
TP1	0.28	0.00	0.63	0.24	0.14	0.17	0.08	0.28	0.11	0.77	0.74	0.00	0.69	0.09
TP2	0.92	0.36	0.53	0.69	0.90	0.70	0.72	0.02	0.66	0.35	0.80	0.00	0.87	0.61
TP3	0.29	0.14	0.02	0.27	0.42	0.55	0.33	0.90	0.99	0.62	0.33	0.01	0.10	0.63
TP4	0.62	0.43	0.71	0.83	0.55	0.44	0.60	0.38	0.75	0.69	0.88	0.01	0.83	0.22
TP5	0.02	0.10	0.00	0.45	0.01	0.07	0.21	0.03	0.73	0.99	0.20	0.19	0.97	0.01

Table 6.16. P-value terms in the case of Normal distribution function (source: Ercolani et al., 2014).

As can be seen, the Gamma function has p-values higher than the Gumbel function in 49 of 70 cases (i.e. in 70% of cases) and than the Normal function in 40 of 70 cases (i.e. 57% of cases). Therefore, the Gamma function is selected to represent passenger flows. Having fixed the statistical distributions, three different travel demand levels are considered: the 50th percentile, corresponding to the traditional approach based on the average condition; the 85th percentile, corresponding to a moderately high value of travel demand; the 95th percentile, corresponding to an exceptionally high value of travel demand. Figure 6.7 shows an example of the approach in the case of Piscinola station during the last time period of the day (i.e. from 19.00 to 23.30).

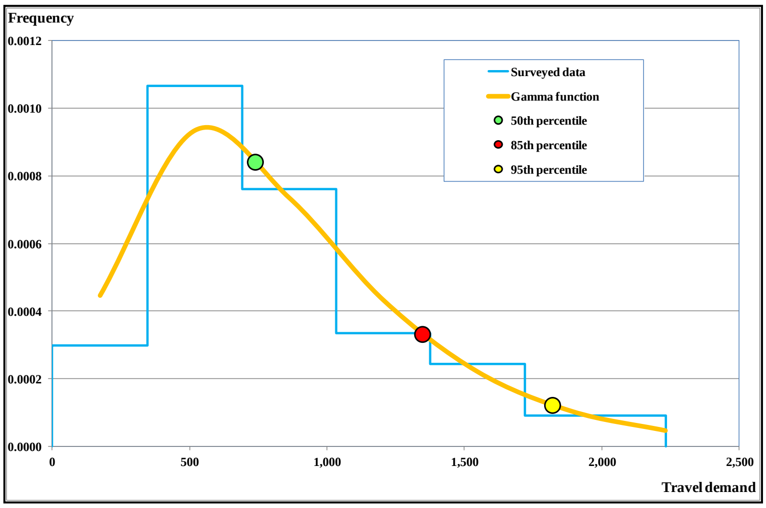


Figure 6.7 Gamma function distribution and demand levels in the case of station no. 1 (Piscinola) in time period no. 5 (19.00–23.30). (source: Ercolani et al., 2014).

Following surveyed turnstile data concerning the opening of the new service (after the 31st December 2013 the line opened up to Garibaldi station) enable the estimation of demand profiles of the new stations.

6.4 Application of the DSS in the case of failure contexts using pattern demand profiles

The DSS was then applied considering the new demand profiles. In particular, adopting the previous infrastructure framework (i.e. from Piscinola to Dante), the ordinary daily service and three specific disrupted scenarios were assessed, namely:

- at Dante station (the terminus far away from the depot) a breakdown occurs in the ATP (Automatic Train Protection) system of the train performing run 801 (i.e. the run which starts from Dante at 7.16 a.m.). Hence, the convoy is forced to travel at a maximum speed of 45 km/h;

- in the track section between Piscinola and Colli Aminei, the ATP system of the infrastructure breaks down for the whole day and hence all trains have to respect the speed limit of 45 km/h;
- at Colli Aminei station, the train performing run 602 (i.e. the run which starts from Piscinola at 7.37 a.m.) experiences a breakdown in the door closing system. Hence, the train is not allowed to travel with passengers on board.

Obviously, all results have been obtained by considering three different travel demand levels, i.e. 50th, 85th and 95th percentiles, as previous described. Moreover, the effect of fleet composition on the intervention strategies is evaluated again. Therefore, the simulations were repeated three times adopting the feasible convoy sequences, i.e. 3-3-2, 3-2-3 and 2-3-3.

Table 6.17 provides user costs in the case of different fleet compositions and different travel demand levels during the daily regular service.

Travel demand level	**Fleet 3-3-2**	**Fleet 3-2-3**	**Fleet 2-3-3**
50th percentile	172,102	172,102	172,102
85th percentile	222,919	222,919	223,263
95th percentile	258,406	258,293	259,017

Table 6.17 Daily user generalised costs [in Euros] in the case of regular service (source: Ercolani et al., 2014).

These outputs are important to assess the increase of costs in the case of failure scenarios. For the first malfunction, as shown by Table 6.18, three intervention strategies may be implemented:

1. the train completes the whole service until the last terminus (i.e. Piscinola) and then it is sent to the depot. A new convoy (replacement) will continue the service;

2. the train completes the whole service until the last terminus (i.e. Piscinola) and then it is driven to the depot. No convoy will replace the faulty train;
3. the faulty train will continue the service throughout the day.

Intervention strategy	Travel demand level	Fleet 3-3-2	Fleet 3-2-3	Fleet 2-3-3
Maintenance of the faulty train (depot) with a new replacement convoy	50th percentile	172,146	172,146	172,146
	85th percentile	222,979	222,979	223,324
	95th percentile	258,477	258,365	259,088
Maintenance of the faulty train (depot) without any replacement convoy	50th percentile	197,148	197,973	197,946
	85th percentile	258,062	259,027	258,516
	95th percentile	303,250	304,142	302,450
The faulty train continues the services throughout the whole day	50th percentile	174,370	174,370	174,370
	85th percentile	225,836	225,836	226,121
	95th percentile	261,404	261,308	261,898

Table 6.18 Daily user generalised cost [in Euros] in the case of train failure (source: Ercolani et al., 2014).

On analysing the simulation results of intervention strategies it emerges that the replacement of a faulty train with an efficient new convoy is always the best operational strategy. Obviously, this is possible only if there are additional convoys. Hence, this methodology allows quantification of the cost (purchase costs of new traction units) and benefits of having additional trains.

Moreover, the simulations show that if there are no additional convoys for replacement operations, the best strategy consists in using the faulty train because the speed limit (45 km/h) allows a fair service to be attained in any event (user discomfort is reduced).

In the second failure context, i.e. a signalling system failure, it is not possible to implement suitable intervention strategies for reducing user discomfort. However, in this case estimation of user disutilities can be useful for quantifying the costs and benefits of carrying out urgent maintenance operations to restore the regular service or, in the planning/design phases, the increase in redundancy of technological components to reduce failure probabilities. However, Table 6.19 shows that the increase in user generalised costs is always lower than 2%. In fact, the breakdown affects train running times for a short section which does not influence severely the service.

Travel demand level	Fleet 3-3-2	Fleet 3-2-3	Fleet 2-3-3
50th percentile	175,419	175,419	175,419
85th percentile	227,254	227,254	227,540
95th percentile	263,058	262,977	263,607

Table 6.19 Daily user generalised costs [in Euros] in the case of signalling (system failure source: Ercolani et al., 2014).

Finally, in the third case, i.e. a breakdown which reduces the functionality but not the performance of a convoy, there are generally two feasible strategies: leaving the train on the line (i.e. at the station platform) and trying to repair the damage; or making passengers alight on the platform, driving the faulty train onto the maintenance track and putting a replacement convoy (from the depot) in operation. However, repair times are generally unpredictable and it is therefore necessary to compare effects in the case of an a-priori unknown repair time of 10 minutes, 20 minutes and 30 minutes. In this case results were estimated only in the case of Fleet 3-3-2 because only a doubleheader convoy can be replaced on the maintenance track next to Colli Aminei station (in previous examples this particular restriction has not been considered). Table 6.20 shows the results only for the fleet sequence 3-3-2.

Travel demand level	Drive the faulty train onto the maintenance track	Repair the damage (10 min)	Repair the damage (20 min)	Repair the damage (30 min)
50th percentile	175,885	173,002	177,938	186,624
85th percentile	227,942	224,128	230,602	242,444
95th percentile	264,298	259,832	267,183	281,202

Table 6.20 Daily user generalised costs [in Euros] in the case of different repair Strategies (system failure source: Ercolani et al., 2014).

In fact, since the maintenance track can host at most a doubleheader train and the broken train (i.e. run 801) is the ninth convoy starting from Piscinola, only this sequence assures the feasibility of the strategy. As can be seen, only if the time required to repair the train is lower than 10 minutes is it worth waiting at the station. In all other cases, due to the great propagation of delays on following runs, it is best to remove the faulty train from service and request a replacement convoy.

6.5 The robustness assessment of the optimal intervention strategies

The results shown in previous applications were based on the deterministic simulation of the network. Obviously, this kind of approach does not allow the robustness evaluation of the proposed strategies. Hence, a more complex procedure, which involves two phases (see paragraph 4.1 for more detail), has to be performed: the first one adopts deterministic simulations which enable the evaluation of optimal and near optimal strategies; the second one, which is based on stochastic simulations, assesses the effectiveness of the strategies selected in the previous step.

Obviously, as shown by relation (4.8), all elements of the network can be affected by uncertainty. Nevertheless, the main factors which are worth analysing concern the service performances (e.g. speed and acceleration variations), the planned timetable (arrival and departure delays, dwell time variations) and the travel demand levels at stations.

Two different analyses based on the current Line 1 infrastructure (i.e. from Piscinola to Garibaldi) have been carried out to investigate the effects of stochasticity on recovery solutions. In particular, the simulated service concerns the daily timetable performed by METRONAPOLI during a weekday, that is:

- a train every 8 minutes from 6:00 am to 09:00 pm;
- a train every 14 minutes from 09:00 pm to midnight.

Furthermore, the rail operator considers the dwell time as a constant value for all the stations, amounting to 20 seconds. All trains are composed of two traction units (i.e. 4 coaches) , for a total capacity of 864 passengers.

The first application focused on the variability of travel demand, the randomness of breakdown occurrence and above all, the stochastic effect of dwell times on the service. Indeed, especially in metro-rail contexts, due to the short distance between stations, running times can be considered as constant values while passenger flows are one of the main disturb effects of the service and cannot be neglected. This implies the resolution of the dwell time estimation problem, (see paragraph 4.5.2) that is a particular fixed point problem which does not fulfil completely the hypotheses of the Brouwer's theorem. As a consequence, although there is a solution, it is necessary to look for numerical evidence that this solution could be also unique. To this purpose, the procedure previously illustrated (using both the iterative and the MSA algorithms) is applied to the morning rush hour (i.e. between 7 am and 9 am) 10 times starting from different random enumerations of dwell times. In all 10 cases, both algorithms converge to the same configuration of dwell time meaning that, although the uniqueness can never be demonstrated mathematically, there is a numerical evidence that the solution could be unique.

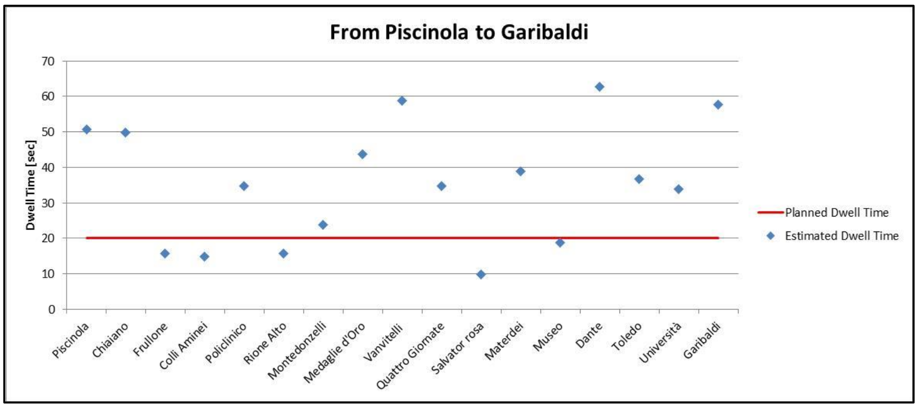


Figure 6.8 Comparison of the estimated dwell times and the planned dwell times (source: Placido et al., 2015a).

Figure 6.8 shows a comparison of the estimated dwell time values and the planned dwell times. It is worth highlighting that due to the high headway between two consecutive runs (i.e. 8 minutes), the snowball effect during ordinary conditions is not evident and therefore, there is no propagation of delay from one train to the following one. Figure 6.9 instead represents the number of estimated passenger for each coach although, in this application, the variability of discomfort perceived by passengers on board the train due to crowding levels has been neglected.

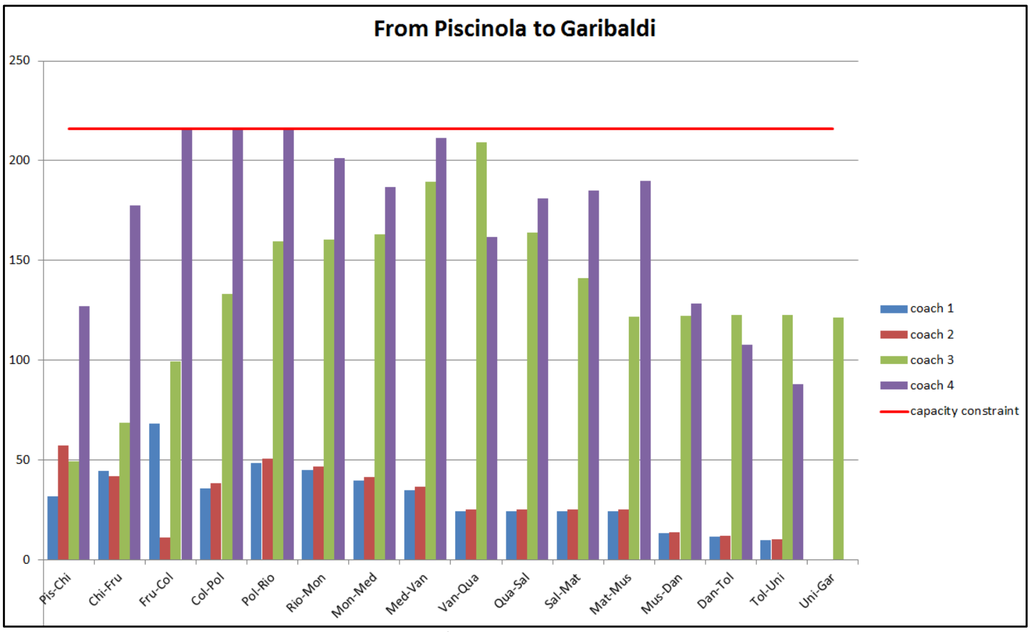


Figure 6.9 Load diagram of a general train during the outward trip.

However, the dwell time is a random variable. The value obtained by means of the abovementioned algorithms can be viewed as the expected value of the dwell time necessary to complete the boarding/alighting process. Nevertheless, as confirmed by several authors (see for instance Goverde et al., 2001; Yuan, 2002), the distribution of free dwell times (i.e. the time necessary to let passengers alight/board the train) may vary from system to system and can depend on many factors such as the earliness of the train arrival. Therefore, this distribution has to be surveyed. Since in the following simulation no information regarding the probability density function of dwell times is available, it is assumed that the departure delay follows an exponential distribution whose mean value is equal to the difference between the estimated dwell time and the planned dwell time.

As far as the application is concerned, the malfunction of the on-board ATP system of a train performing its service during the morning rush hour is considered. This forces the train to keep a speed lower than 45 km/h causing a bottleneck for the following runs. In this case, dispatchers have to decide where it is convenient to recover the faulty train. However, due to the randomness of the event, it is not possible to establish in advance where the breakdown can occur. Thus, some recovery solutions cannot be implementable. For instance, if the train starts having problem between Piscinola and Colli Aminei stations, there are six available intervention strategies. It is possible to use the recovery track in Colli Aminei, Medaglie d'Oro and Garibaldi stations during the outward trip (i.e. from Piscinola to Garibaldi), or to stop the train in Medaglie d'Oro and Colli Aminei but during the return trip. Obviously, it is also possible to let the train continue its service up to Piscinola where it is driven to the depot. The same strategies are not all feasible if the failure event occurs between Garibaldi and Colli Aminei. Therefore, considering the infrastructure characteristics (i.e. the location of the points and recovery tracks) and the resulting recovery solutions, it is worth analysing 5 different failure scenarios, namely:

- The train breaks down during the outward trip (i.e. from Piscinola station to Garibaldi station) before arriving to Colli Aminei station. There are 6 implementable strategies;
- The train breaks down during the outward trip before arriving to Medaglie d'Oro station. There are 5 possible recovery solutions;
- The train breaks down during the outward trip before arriving to Garibaldi station. In this case there 4 implementable strategies;
- The train breaks down during the return trip (i.e. from Garibaldi station to Piscinola station) before arriving to Medaglie d'Oro station. Thus, 3 strategies are feasible;
- The train breaks down during the return trip before Colli Aminei station. Just 2 recovery solutions can be applied.

In any failure scenario, a spare train is introduced to the service for performing the following runs of the day in substitution of the faulty train. This assumption shows the importance of investing in rolling stock in order to increase the reliability and feasibility of the service especially in the case of degraded conditions.

In this application, the uncertainty of users' arrival on the platform is taken into account adopting two different arrival rates corresponding to average level of travel demand (50th percentile of the fitted distribution) and to an exceptionally high value of travel demand (95th percentile of the fitted distribution). Based on this data, the OPM1.0 model assigns passengers to each train according to the network performances. Thus the deterministic phase can be completed as shown by Tables 6.21–6.25. Bold values represent the optimal and the near-optimal solutions which are selected for the following stochastic analysis. In some cases, different travel demand levels lead to a different set $N(\hat{\boldsymbol{y}})$ (set of optimal strategies, see paragraph 4.1) which confirms the fact that travel demand influences the planning of recovery solutions and cannot be neglected. Furthermore, it is worth noting that in most cases, when the failure occurs during the outward trip (i.e. Table 6.21–6.23) the optimal strategies suggest recovering the train during the return journey. This is due to the fact that during the morning rush hours many passengers travel from the suburbs towards the city centre (i.e. Dante, Toledo, Garibaldi). Therefore, passengers prefer travelling by a slower train to waiting for more crowded convoys which run faster. The opposite direction is instead almost unloaded and the interruption of the run does not produce a great cost increase of.

However, when the passenger flow level is considerable (i.e. 95th percentile), for Failure 1 and 2, the model suggests recovering the train as soon as possible so as to reduce the propagation of delays on following runs. Indeed, in these cases, all trains are crowded and keeping the faulty train on operation does not produce great benefits to passengers.

Once established the set $N(\hat{\boldsymbol{y}})$ for each simulated scenario (i.e. bold values in the previous tables), the procedure requires the resolution of the dwell time estimation problem. Thus, the average dwell time is calculated for each station and it is compared to the planned value established by the rail operator (i.e. 20 seconds).

Failure 1	**Travel Demand**	
Strategy	**50th percentile**	**95th percentile**
Colli Aminei (outward trip)	€ 587,464	€ **2,351,480**
Medaglie d'Oro (outward trip)	€ 586,201	€ 2,430,870
Garibaldi	**€ 578,208**	€ 2,428,640
Medaglie d'Oro (return trip)	€ 578,361	€ 2,428,720
Colli Aminei (return trip)	**€ 578,001**	**€ 2,419,970**
Piscinola	**€ 577,305**	**€ 2,427,010**

Table 6.21 Cost of the strategy in the case of failure before Colli Aminei (source: Placido et al., 2015a)

Failure 2	**Travel Demand**	
Strategy	**50th percentile**	**95th percentile**
Medaglie d'Oro (outward trip)	€ 579,901	**€ 2,327,600**
Garibaldi	**€ 573,205**	€ 2,421,750
Medaglie d'Oro (return trip)	€ 572,364	€ 2,420,230
Colli Aminei (return trip)	**€ 572,309**	**€ 2,420,130**
Piscinola	**€ 572,208**	**€ 2,419,970**

Table 6.22 Cost of the strategy in the case of failure before Medaglie d'Oro (source: Placido et al., 2015a)

Failure 3	**Travel Demand**	
Strategy	**50th percentile**	**95th percentile**
Garibaldi	€ 573,134	€ 2,421,630
Medaglie d'Oro (return trip)	**€ 572,293**	**€ 2,420,110**
Colli Aminei (return trip)	**€ 572,238**	**€ 2,420,020**
Piscinola	**€ 572,137**	**€ 2,419,850**

Table 6.23 Cost of the strategy in the case of failure before Garibaldi (source: Placido et al., 2015a)

The difference corresponds to the mean delay whose distribution follows a negative exponential random variable. Obviously, in case the estimated dwell time is lower than the planned one, the mean delay is considered equal to zero. Based on this assumption, 50 stochastic simulations are set up and, for each selected scenario, the objecting function (4.3) is evaluated. Table 6.26 shows the results of the stochastic approach. Basically, the DSS provides the robustness of the optimal solutions specifying the number of times they are better than the other strategies (i.e. probability of success).

As can be seen, the optimal solutions evaluated in the deterministic phase still remains preferable also adopting a stochastic approach.

In particular, the more the difference in terms of user generalised cost is, the higher is the robustness of the optimal solution identified in the deterministic phase.

Failure 4	**Travel Demand**	
Strategy	**50th percentile**	**95th percentile**
Medaglie d'Oro (return trip)	**€ 572,256**	**€ 2,420,070**
Colli Aminei (return trip)	**€ 572,195**	**€ 2,419,960**
Piscinola	**€ 572,077**	**€ 2,419,770**

Table 6.24 Cost of the strategy in the case of failure before before Medaglie d'Oro (return trip) (source: Placido et al., 2015a)

Failure 5	**Travel Demand**	
Strategy	**50th percentile**	**95th percentile**
Colli Aminei (return trip)	**€ 572,364**	**€ 2,420,230**
Piscinola	**€ 572,211**	**€ 2,419,980**

Table 6.25 Cost of the strategy in the case of failure before Colli Aminei (return trip) (source: Placido et al., 2015a).

	Travel Demand	
Failure 1	**50th percentile**	**95th percentile**
Colli Aminei (outward trip)	-	100%
Garibaldi	0%	-
Colli Aminei (return trip)	0%	0%
Piscinola	100%	0%
Failure 2		
Medaglie d'Oro (outward trip)	-	100%
Medaglie d'Oro (return trip)	16%	-
Colli Aminei (return trip)	26%	0%
Piscinola	58%	0%
Failure 3		
Medaglie d'Oro (return trip)	14%	20%
Colli Aminei (return trip)	26%	28%
Piscinola	60%	52%
Failure 4		
Medaglie d'Oro (return trip)	8%	4%
Colli Aminei (return trip)	28%	24%
Piscinola	64%	72%
Failure 5		
Colli Aminei (return trip)	24%	14%
Piscinola	76%	86%

Table 6.26 Results of the stochastic phase. (source: Placido et al., 2015a).

In some cases, this solution is totally confirmed even by the stochastic analysis. This is due to the fact that service frequency is not so high and, therefore, the influence of passengers on the service (i.e. increase in dwell times) is not noteworthy. Indeed, the propagation of delays and the resulting conflicts between trains (i.e. snowball effect) are limited. However, the sensitivity analysis gives important indications about the error degree when only the deterministic approach is implemented.

The second application investigates the effects of the same failure scenario introducing some changes on recovery solutions and stochastic parameters. This time, the dwell time, although it is still evaluated as a random variable, is not considered as flow dependent. As a result, the complex procedure related to the dwell time estimation problem is neglected. However, contrary to what previously done, the departure delay of each train at each station simply follows a negative exponential random variable whose average is 10 seconds. In addition, the variability of train performance (i.e. acceleration and speed) is defined according to a piecewise linear distribution function where 33% of the trains are supposed to perform at 85%–90%, 33% at 90%–95%, and 34% at 95%–100%. In fact, also the variability in acceleration and maximum speed could strongly influence the service.
Furthermore, more recovery strategies are evaluated so as to take completely advantage of any available point and/or recovery track of the line, which results in the following list of strategies:

1. The faulty train continues to perform its service all day;
2. the train stops at Colli Aminei during its outward trip (i.e. from Piscinola to Garibaldi) and is then driven onto the recovery track. No spare trains are considered;
3. the train stops at Medaglie d'Oro during its outward trip and is then driven onto the recovery track. No spare trains are considered;
4. the train stops at Garibaldi at the end of its outward trip and is then driven onto the recovery track. No spare trains are considered;
5. the train completes the outward trip and starts the return trip (i.e. from Garibaldi to Piscinola) up to Medaglie d'Oro where it is driven onto the maintenance track. No spare trains are considered;

6. the train completes the outward trip and starts the return trip up to Colli Aminei where it is driven onto the maintenance track. No spare trains are considered;
7. the train completes the outward trip and starts the return trip up to Piscinola where is driven to the depot. No spare trains are considered;
8. the train stops at Colli Aminei during its outward trip and is then driven onto the recovery track. A spare train starts from Piscinola to replace the faulty rolling stock for the rest of the daily operation;
9. the train stops at Medaglie d'Oro during its outward trip and is then driven onto the recovery track. A spare train starts from Piscinola to replace the faulty rolling stock for the rest of the daily operation;
10. the train stops at Garibaldi at the end of its outward trip and is then driven onto the recovery track. A spare train starts from Piscinola to replace the faulty rolling stock for the rest of the daily operation;
11. the train completes the outward trip and starts the return trip up to Medaglie d'Oro where it is driven onto the maintenance track. A spare train starts from Piscinola to replace the faulty rolling stock for the rest of the daily operation;
12. the train completes the outward trip and starts the return trip up to Colli Aminei where it is driven onto the maintenance track. A spare train starts from Piscinola to replace the faulty rolling stock for the rest of the daily operation;
13. the train completes the outward trip and starts the return trip up to Piscinola where it is driven to the depot. A spare train starts from Piscinola to replace the faulty rolling stock for the rest of the daily operation;
14. the train stops its run at Dante and, after changing direction, is driven empty to the depot. A spare train starts from Piscinola to replace the faulty rolling stock for the rest of the daily operation;
15. the train stops its run at Dante and, after changing direction, is driven empty to the depot. No spare trains are considered;

16. the train stops its run at Vanvitelli and, after changing direction, is driven empty to the depot. No spare trains are considered;
17. the train stops its run at Vanvitelli and, after changing direction, is driven empty to the depot. A spare train starts from Piscinola to replace the faulty rolling stock for the rest of the daily operation;
18. the train stops its run at Medaglie d'Oro and, after changing direction, is driven empty to the depot. A spare train starts from Piscinola to replace the faulty rolling stock for the rest of the daily operation;
19. the train stops its run at Medaglie d'Oro and, after changing direction, is driven empty to the depot. No spare trains are considered;
20. the train stops its run at Coli Aminei and, after changing direction, is driven empty to the depot. No spare trains are considered;
21. the train stops its run at Colli Aminei and, after changing direction, is driven empty to the depot. A spare train starts from Piscinola to replace the faulty rolling stock for the rest of the daily operation.

As it can be seen, for any feasible strategy, the advantage of introducing a spare train is evaluated in terms of user generalised cost.

The deterministic calculation of the optimal solutions provides the results showed in Table 6.27 and 6.28 and Figure 6.10. Similarly to the previous example, the randomness of travel demand is considered through the adoption of two different pattern profiles which now correspond to the average (50^{th} percentile) and the medium-high level (85^{th} percentile) of passenger on the platforms.

As it can be seen, the optimal solution which produces the lowest user generalised cost is that corresponding to '*strategy 13*'.

All strategies which consider the introduction of the spare train are evidently preferable inasmuch as they give the possibility to calm down the negative effects of the breakdown in less time than the other alternatives. In addition, these solutions enable to keep the planned frequency levels during the rest of the day. The neighbourhood of the optimal solution is constituted of strategy

12 and 10. All the three optimal recovery solutions suggests stopping the train after completing the outward trip.

Scenario	Average travel demand level User generalised cost	High travel demand level User generalised cost
1	€ 582,728	€ 1,679,960
2	€ 1,010,220	€ 1,991,610
3	€ 1,008,680	€ 2,021,960
4	€ 1,000,280	€ 1,982,610
5	€ 1,000,390	€ 1,982,780
6	€ 1,000,030	€ 1,982,290
7	€ 999,338	€ 1,981,330
8	€ 587,464	€ 1,682,420
9	€ 586,201	€ 1,700,280
10	**€ 578,208**	**€ 1,674,060**
11	€ 578,361	€ 1,674,280
12	**€ 578,001**	**€ 1,673,790**
13	**€ 577,305**	**€ 1,672,830**
14	€ 579,384	€ 1,676,050
15	€ 1,001,420	€ 1,984,550
16	€ 1,008,460	€ 2,011,190
17	€ 586,131	€ 1,694,210
18	€ 586,054	€ 1,700,090
19	€ 1,008,530	€ 2,021,760
20	€ 1,010,070	€ 1,991,410
21	€ 587,318	€ 1,682,220

Table 6.27 Objective function values for any feasible intervention strategy (source: Placido et al., 2015b)

This confirms the fact that the higher arrival rates in that section force the rail operator to keep the service also in degraded conditions so as to reduce the waiting time of a large number of customers.

Strategy 10, 12 and 13 are then simulated again in the following stochastic phase. For this purpose, 100 stochastic scenarios were constructed based on previous assumptions on acceleration, speed and departure delays.

	Strategy 10	**Strategy 12**	**Strategy 13**
average travel demand	27%	32%	41%
high travel demand	34%	33%	33%

Table 6.28 Number of times (%) of optimality for each strategy (stochastic analysis) (source: Placido et al., 2015b).

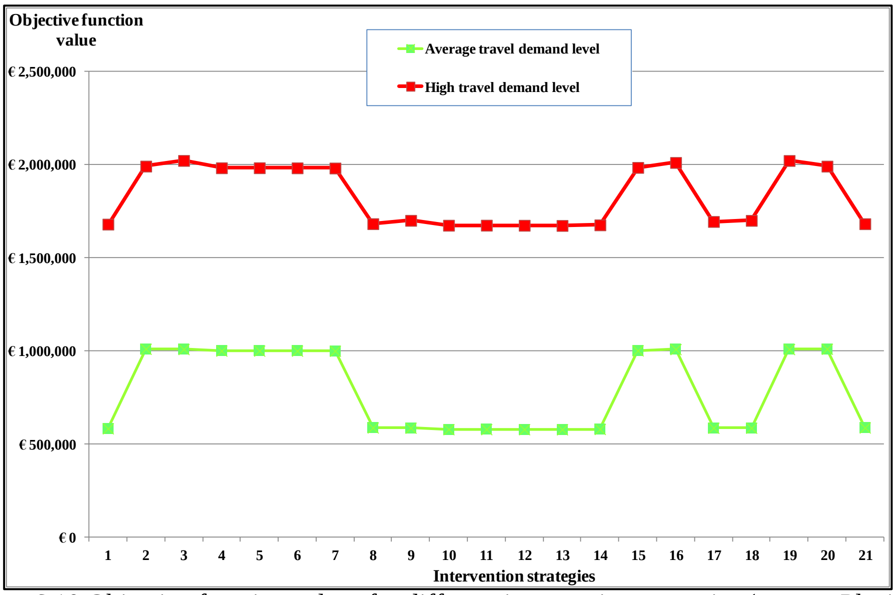


Figure 6.10 Objective function values for different intervention strategies (source: Placido et al., 2015b).

As regards the computational results (Table 6.28), given average conditions, strategy 13 is the one which guarantees the highest level of robustness. It has to be preferred since it proves the optimal solution in most cases. Moreover, from an operational point of view, this alternative is better than the other two: after running at low speed up to the terminus (i.e. Piscinola), the train is driven directly to the depot and does not need to be hauled from a recovery track at the end of the daily operations. In the case of crowded days, by contrast, all the three alternatives are equivalent and there is no evident advantage to choosing one over another. For this reason, strategy 13 is still the best since it guarantees fewer movements without passengers. However, even if this result may seem the same as that obtained without considering parameter variability, simulations show that only in 41% of cases does the deterministic approach correspond to a 'real condition' scenario, while in the other 59% of cases the deterministic approach may provide a non-optimal intervention strategy. Therefore, the addition of the stochastic procedure gives important indications about the error degree when only the deterministic approach is implemented.

6.6 The definition of a new objective function for increasing effectiveness and efficiency of the rail system

Due to the introduction of liberalisation and competition within the rail system (see paragraph 2.1), public authorities now try to pursue the difficult task of considering both public interests (sustainability, accessibility, employment, etc.) and commercial interests (profit, return on investment, growth). In this context, it is necessary to carry out analyses about the rail service so as to gain useful information for increasing efficiency, effectiveness and productivity. In many cases, great importance has been attached to Key Performance Indicators (KPIs) which do not usually consider customer needs (Lan and Lin, 2006; Smith, 2012; Hansen et al., 2013). The latter have been considered in this thesis as the most important, highlighting the fact that although accurate financial management is required, the main task is to 'capture' the highest number of passengers and satisfy their requirements. However, especially in the case of concession regimes, rail service must assure good levels of effectiveness, efficiency and productivity. Adding proper improvements to the proposed method (in particular to the On-Platform Model), the DSS can be useful to define performance criteria taking into account both user satisfaction and operator interests. In this case, the objective function 4.3 becomes:

$$Z(\boldsymbol{y}, \boldsymbol{fc}, \boldsymbol{np}, \boldsymbol{unf}, \boldsymbol{rc}) = \frac{ugc}{\beta_{ugc}} + \frac{pen}{\beta_{pen}} + \frac{1}{\eta} + \frac{oc}{\beta_{oc}} \tag{6.1}$$

where:

- ugc is the already mentioned user generalised cost (objective function 4.3). In this context, the dependence of $\beta_{on-board}$ weight on rail crowding is not neglected;
- pen represents the extra-cost perceived by passengers who are forced to leave the system. Indeed, in order to perform a better simulation of travel demand, the model assumes that passengers leave the system after waiting for more than two trains without managing to board. The

same event happens in case no train is coming after a certain time period (for Line 1 test case this time is considered equal to 30 minutes) which is provided as input data. In fact, the time passengers are willing to wait for a train on the platform is strongly correlated to the characteristics of the public transport system and therefore, it depends on the analysed network. In this way, the OPM 1.0 tool can simulate realistically passenger behaviour in case of failure scenario when it is likely that the service can be interrupted for a period:

$$pen = \sum_{s=station} \cdot \sum_{p=platform} \cdot \sum_{r=run} \cdot pl^{r}_{s,p} \cdot (optw_{s,p} + tls) \cdot vot \tag{6.2}$$

Where $pl_{s,p}$ is the number of passenger leaving the system at station *s* and on platform *p* between run (*r-1*) and run *r*, $optw_{s,p}$ is the time these passengers have waited before leaving; tls is the time necessary to leave the system, namely the time passengers need to change public transport (for Line 1 test case this time is considered as 15 minutes);

- η is the attractiveness of the rail service and it is calculated as follows:

$$\eta = \sum_{r=run} \frac{pax - km_r}{seat - km_r} 100\% \tag{6.3}$$

where $pax - km_r$ is the number of passenger per kilometres of run *r* while $seat - km_r$ is the number of seat per kilometres of run *r*. This indicator is extremely important for train operating companies. In fact, as already demonstrated (see for instance Albrecht, 2009), the operating cost is mainly proportional to the operational effort which is generally indicated as offered seat per kilometres. Obviously, train operating companies would like to carry a given demand with minimal cost which results in operational efficiency values as much as possible close to 1. Therefore, it is worth underlining that the lower is the operational efforts, the bigger is the occupation rate of the trains and the disutility perceived by passengers on board;

- oc is the operational cost, namely the cost rail operators have to spend for each train performing the service:

$$oc = \sum_{r=run} length_r \cdot \frac{\cos t}{km}_r \cdot nudt_r \tag{6.4}$$

where $length_r$ is the length of the path performed by run *r* expressed in kilometres, $\frac{cost}{km}_r$ is the cost per kilometre, $nudt_r$ is the number of traction units composing the run *r* (for Line 1 test case, the $\frac{cost}{km}_r$ is equal to 18.17 €/km for each traction units composing the train);

- β_{ugc}, β_{pen} and β_{oc} are homogeneity coefficients which are necessary to homogenise the different values.

The new methodology has been tested on the new Line 1 service (i.e. from Piscinola to Garibaldi). In particular, 6 different scenarios in terms of fleet composition have been considered, namely:

- Scenario 1: ordinary service performed by METRONAPOLI, consisting of a fleet of 10 train which are composed by two traction units. Each traction unit can carry 432 passengers and so the maximum number of passenger per train is 864;
- Scenario 2: ordinary service performed until 2 pm, afterwards all the train are decoupled and continue their service with just one traction unit (i.e. 432 passengers per train);
- Scenario 3: just 34 of the daily 242 runs are performed by double traction unit trains, mostly running during rush hours. For the rest of the day, single traction unit trains are provided. This strategy is implemented so as to increase the operational efficiency index every time its value is below 0.5;
- Scenario 4: all the service is performed by triple traction unit trains (1296 passengers per train) which is the maximum train

length that Line 1 stations can host. This strategy increases the service quality perceived by the passengers;

- Scenario 5: the whole fleet is constituted by single traction unit trains so as to increase as much as possible the operational efficiency;
- Scenario 6: ordinary service performed until 7 pm, afterwards, just single traction unit trains are considered.

The aim is to look for the optimal fleet composition which maximises the benefits for customers, rail operators and community.

	USER COST	PENALTY	SERVICE COST	SEAT-KM	PAX-KM	η
scenario 1	€ 493,572.48	€ 5,074.52	€ 145,159.40	3,451,230	1,493,940	0.43
scenario 2	€ 463,969.90	€ 52,521.10	€ 109,019.27	2,591,980	1,308,500	0.50
scenario 3	€ 469,829.10	€ 85,241.90	€ 82,516.51	1,961,870	1,212,570	0.62
scenario 4	€ 407,347.00	-	€ 217,739.10	5,176,850	1,509,220	0.29
scenario 5	€ 455,120.00	€ 105,044.00	€ 72,579.70	1,725,610	1,131,730	0.66
scenario 6	€ 500,052.50	€ 11,231.50	€ 132,510.36	3,150,490	1,472,740	0.47

Table 6.29 Results of the implemented scenarios.

Table 6.29 shows the results of the implemented scenarios. As it can be seen, user generalised cost is minimum in the case of scenario 4, since no passenger is forced to leave the system and manages to board the first approaching train. As expected, crowding levels are lower than those of the other scenarios and therefore, customers feel comfortable on-board the train because of the increased space. Scenario 5 by contrast, fulfils just the needs of train operating companies. In fact, the operational efficiency is the highest possible (η=0.66), but passengers are very disadvantaged and in most cases leave the system (Penalty=105,044.00 €). The other scenarios underline how the increase in operational efficiency entails a reduction of service quality perceived by passengers. In order to find the optimal solution, Table 3 shows the objective function values of (6.1).

It is worth noting that scenario 4 is the ideal fleet sequence for customers' viewpoint. In fact, it provides the best level of service quality and, due to the high level of travel demand served, it keeps operational efficiency (η=0.29) close to the average level generally achieved by metro lines (Albrecht, 2009). On the other side, service costs are considerably higher. Unfortunately, this strategy cannot be implemented in the reality since METRONAPOLI does not have enough rolling stock to perform this kind of service. Surprisingly, the current fleet composition provides the optimal objective function value limiting the user generalised cost and providing good level of service efficiency.

As shown by Figure 6.11, the other solutions increase the operational efficiency at the expense of the service quality perceived by passengers.

In conclusion, it is worth adopting a DSS for planning the service considering the needs of the many participants involved in the rail system. This is extremely important in the case of concession regimes which aim to keeping high level of service quality maximising efficiency and productivity. In particular, operational efficiency has to be taken into account since it could provide also important indications about how to direct future investments.

	USER COST/β_{ugc}	PENALTY/β_{pen}	$1/\eta$	SC/ β_{pen}	Z
scenario 1	**4.936**	**0.507**	**2.31**	**1.45**	**9.20**
scenario 2	4.640	5.252	1.98	1.09	12.96
scenario 3	4.698	8.524	1.62	0.83	15.67
scenario 4	4.073	0	3.43	2.18	9.68
scenario 5	4.551	10.504	1.52	0.73	17.31
scenario 6	5.001	1.123	2.14	1.33	9.59

Table 6.30 Objective function values of the implemented scenarios.

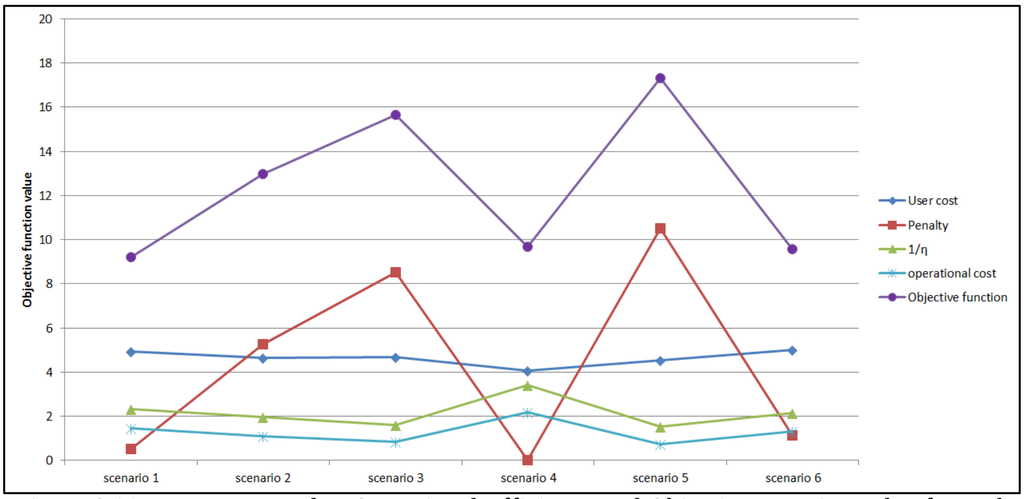


Figure 6.11 User cost, Penalty, Operational Efficiency and Objective Function value for each simulated scenario.

6.7 Benefits of the combination between macroscopic and microscopic approach.

The model described in this thesis is quite complex and requires long computational time to be implemented. Furthermore, when the set of scenarios is huge an exhaustive approach cannot be put into practice. For instance, this technique could not be convenient when, after failure events, it is necessary to plan a completely new timetable which satisfies new travel demand profiles.

Therefore, the proposed microscopic approach has been combined with the macro-optimisation model presented in Cadarso et al. (2013).

This model, known as INtegrated TImetable and ROlling Stock Rescheduling Model (INTIROSRM), aims at computing the timetable and the rolling stock schedule for a disrupted metro network accounting for passengers flows. The INTIROSRM is based on a multi-objective function which minimises the incurred system costs and the passenger inconvenience. The latter is expressed through the number of '*denied passengers*', namely the number of users who do not manage to board a train in less than 10 minutes and decide to leave the system.

As regards the infrastructure, the whole network is represented through a simplified graph where nodes are the stations while links are the line sections connecting them. Running times are considered constant and headways are deterministic values imposed as input data. Fleet size, rolling stock characteristics and train composition constraints ensure the train units' flow balance. Basically, the model, according to passenger flow at each station, provides a new timetable which can include or cancel runs or modify frequency and fleet composition. However, only the microscopic approach can check if the solution provided by the macro-optimisation model is feasible or not from an operational point of view. Furthermore, the macro model does not perform a real travel demand assignment. In fact, it treats the demand heuristically inasmuch as it is unable to trace individual passengers. Hence, demand on each arc (i.e. between successive stations) is not linked with the one of successive arcs. As a consequence, a denied passenger still shows up in the demand of later arcs. Although in Cadarso et al. (2013) it is demonstrated that the denied demand is very well approached whenever the passenger costs are part of the objective, the flow assignment performed by the On-Platform Model can give more precise indications about the planned solutions. Therefore, the micro-simulation approach is adopted to evaluate the feasibility of the proposed solutions not only in terms of operating service (i.e., adoption of recovery tracks, precise time requirements for shunting operations, signalling system constrains) but also in terms of user inconvenience (i.e. actual evaluation of passenger who cannot travel at their desired time).

On the other hand, the great advantage provided by the macroscopic approach is the possibility to design a new optimal timetable in few seconds (see the table below to have an idea about the computational time required by the macro model).

The two models have been indeed joined by means of an iterative process. At each iteration, the macro-optimisation procedure computes a strategy (i.e. new timetable). This strategy is then assessed by the micro-one providing feedback to the macro-optimisation model in terms of penalties. These penalties affect either the schedule (if the strategy turns out to be infeasible as evaluated by the micro-simulation approach) and/or the passenger inconvenience (in order to minimise the number of denied passengers). The iterative process stops when the convergence criterion is reached, that is when the Mean Absolute Percentage Error (MAPE) in the number of denied passengers between two iterations is lower than 5%:

$$MAPE_{pax_\nu} = \frac{\sum_{s \in S} \left| dp_s^{\nu} - dp_s^{\nu-1} \right|}{\sum_{s \in S} dp_s^{\nu-1}} < 0.05 \tag{6.5}$$

where dp_s is the number of denied passengers at station s and ν is the iteration.

The proposed methodology has been applied to Line 1 system. In particular, the following disrupted scenario is simulated: the rolling stock material performing run 801 (i.e. the run which starts from Dante station at 7:16 a.m.) breaks down at Rione Alto station. Hence, all passengers are forced to alight the train and the faulty train is driven to the closest recovery track, which is located in Colli Aminei station (so as not to disrupt the rest of the train service). It is assumed that the time needed to fix the faulty train is 1 hour. Obviously, this train can no longer follow the planned timetable. Therefore, the aim is to determine the actions that should the operator take in order to reschedule the system as soon as possible considering both operator and passenger costs.

Table 6.31 highlights the solutions provided by the macro-optimisation and microscopic approach. The first column is the iteration number. The rest of the columns show different characteristics of the given solution. Column TU gives the number of train units used by the solution. Columns TSOC and EMOC give the total operational costs for passenger train services and empty movements, respectively. Column DP shows the number of denied passengers as estimated by the macro-optimisation model. Column ST gives the solution time in seconds of the macro optimisation model. The computational time of the microscopic approach is not reported since it is constant and equal to about 500 seconds (as it can be seen the time required by the micro approach is significant). The last column shows the number of waiting passengers (WP_m) as calculated by the microscopic approach. Basically, this value represents the users who are not able to board the first arriving train because there is not enough capacity. However, since during the simulation there is no interruption higher than 30 minutes, it is supposed that in the microscopic approach no passengers leave the system.

Iteration	TU	TSOC	EMOC	#DP	ST	$\#WP_m$
0	35	54945.28	1079.20	3594.62	3.31	21919.37
1	35	61813.44	1884.80	1980.42	2.38	6586.95
2	35	61813.44	2067.20	1705.58	2.80	5479.84
3	35	60756.80	1915.20	1792.95	2.03	5399.76
4	35	59964.32	1854.40	1758.72	2.06	5857.25

Table 6.31 Solutions of the macro-optimisation and the microscopic approaches.

As it can be seen, the macro-optimisation approach (accounting for the feedback provided by the micro-simulation approach at each iteration) determines a recovery strategy which reduces the number of denied passengers. In particular, WP_m decreases from 21919.37 to 5399.76 which corresponds to the optimal solution in terms of 'waiting passengers'. However, it is worth noting that reducing the number of denied passengers (that is people waiting more than 10 minutes on the platform and leave the system) could not correspond to a decrease in the number of users who manage to board the first arriving train. Indeed, increase the frequency with a given fleet of train (the number of traction units is fixed and equal to 35) means to have more rail convoys running within the network with less capacity since they are decoupled. That is why, the trend of DP and WP_m does not coincide at last iteration.

In conclusion, the combination of two approaches is extremely advantageous when it is necessary to plan a new timetable after the occurrence of critical events. The number of intervention strategies to be tested is huge and this prevents exhaustive procedures based on a micro model from being implemented. The macro model in fact, through an optimisation method can provide new timetables in short time and reduces the number of microscopic simulations. Nevertheless, in order to verify the feasibility of the planned solutions, neither of the two approaches can be neglected. In particular, the application has demonstrated that feedbacks provided by the micro model at each iteration brings the macro-optimisation model to converge to a solution in few steps.

6.8 Application of the proposed model in the case of conventional rail lines.

Previous paragraphs have demonstrated the benefits of this new approach for metropolitan lines which are characterised by high frequency and travel demand levels. However, the same methodology can be applied also in the case of conventional rail lines by introducing some changes to the On-Platform Model.

To this aim, further applications have been performed on the regional railway line '*Formia - Napoli Centrale*' and its branch '*Villa Literno - Napoli Gianturco*' which is also known as the Line 2 of Naples metro system (Figure 6.12).

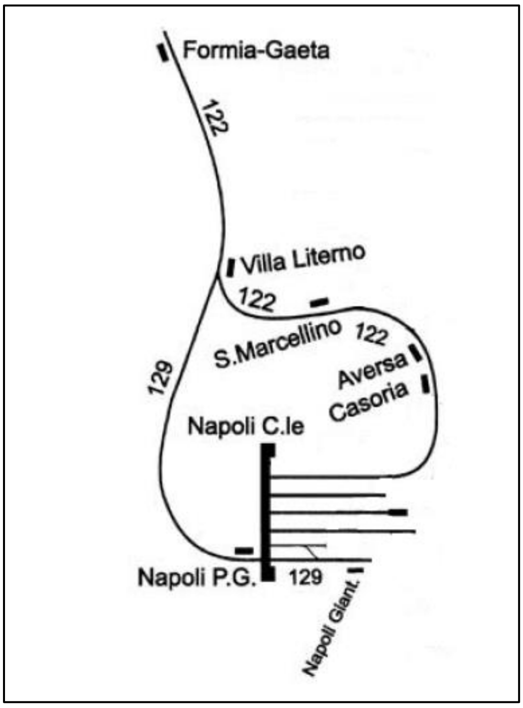


Figure 6.12 Graphic representation of the 'Formia – Napoli Centrale' (i.e. n.122) with its branch 'Villa Literno – Napoli Gianturco' (i.e. n.129).

The Line, consisting of 26 stations for a total length of 122,34 km, is extremely important for the regional traffic of Campania region. Indeed, it is a vital connection between the city of Naples and several high density villages spread along the coast, including those of the southern part of Lazio region. It connects also Naples with Rome and it is interested by a highly heterogeneous rail traffic (intercity, regional, metropolitan and freight trains). High speed trains do not use this infrastructure since there is another line which is exclusive for this kind of rail convoys.

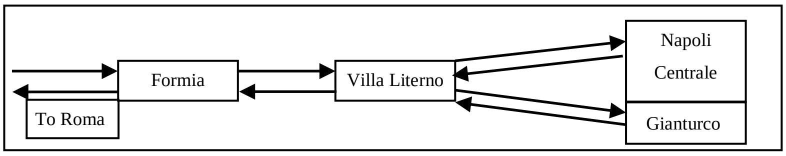


Figure 6.13 Simplified representation of the line showing the connection with Rome.

The first step for implementing the proposed procedure is the microscopic reproduction of the line by means of OpenTrack. Even in this case, the infrastructure has been represented with the maximum level of details in terms of both distances and signalling system. In particular, the latter is based on the BACC train spacing system (see paragraph 2.4.9) with the addition of the SCMT (*Sistema di Controllo Marcia Treno*) which is the Italian standard for the ETCS level 1 (paragraph 2.4.10.2). Furthermore, the section '*Villa Literno – Napoli Gianturco*' is trivialised meaning that trains can be driven indifferently on both even and odd tracks.

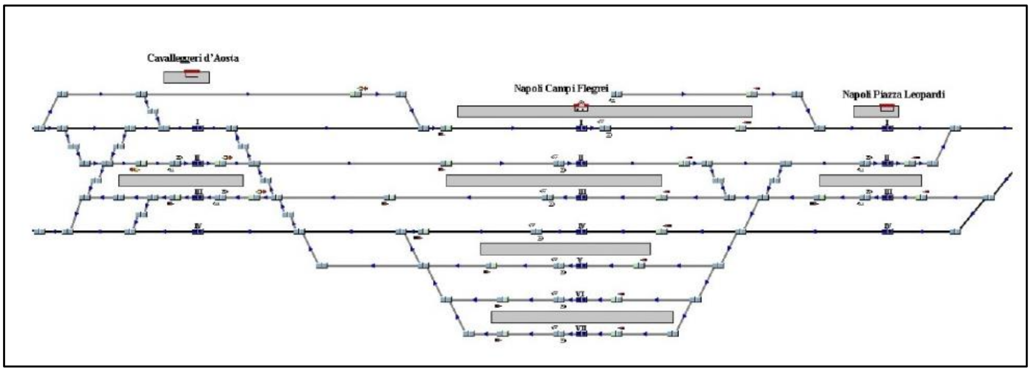


Figure 6.14 Representation of Napoli Campi Flegrei station in Opentrack.

The two sections starting from 'Villa Literno' station are connected in 'Napoli Centrale' (i.e. the main station of the city) through 'Napoli Gianturco' station.

As regards the trains, the whole fleet running on this line has been reproduced in OpenTrack. Basically, the main features which are important for the ***SeSM*** and the ***OPM*** are summarised in Table 6.32.

	Capacity [pass.]	Max speed [Km/h]	Lenght [m]
TAF Train (regional double-decker train)	841	140	104
Minuetto Train (regional train)	284	160	52
Regional Train	452	140	132
Inter-regional Train	520	160	185
Metropolitan Train	328	140	102
ETR 450 (i.e. Intercity high speed)	390	250	234
Intercity Train (8 coaches)	504	200	228

Table 6.32 Characteristics of the rolling stock.

The service simulation has concerned the ordinary timetable of a weekday between 6.00 and 13.00. This time period covers the peak hours during the morning where trains are full of commuters. Basically, there are 169 trains divided as follows:

- 12 intercity trains;
- 59 regional/inter-regional trains;
- 98 metropolitan trains.

Figure 6.15 Different trains running on the Line '*Formia – Napoli Centrale*'.

As far as the Travel Demand Model is concerned, conventional rail lines are based on a completely different service (i.e. low frequency and long distance) from the metropolitan one and therefore, some new assumptions have to be considered. First of all, the arrival rate at station is not significant anymore. Indeed, generally passengers know the timetable and go to the station some minutes before the departure of the train. Hence, what they consider as a disutility is the delay of the train respect to the planned departure time. For the same reason, the Pre-Platform Model, instead of estimating the arrival rate at station, evaluates in this case the total number of users willing to board the train. In addition, except from particular situations where passengers do not have any alternative, since the service frequency is very low, it is likely that no one would remain on the platform waiting for a following convoy in case the capacity of the train is reached (very seldom event) or in case the line is interrupted. According to these hypotheses, the objective function (4.3) is modified in the following:

$$Z(\boldsymbol{y}, \boldsymbol{fc}, \boldsymbol{tnp}, \boldsymbol{rnp}, \boldsymbol{td}) =$$
$$= \sum_i \beta^i_{VOT} \cdot \left(\beta^i_{waiting} \cdot \sum_s \sum_p \sum_r tdel^{i,r}_{s,p}(\boldsymbol{y}, \boldsymbol{fc}, \boldsymbol{tnp}, \boldsymbol{rnp}) \cdot fw^{i,r}_{s,p}(\boldsymbol{td}) + \right. \tag{6.6}$$
$$\left. + \beta^i_{on-board} \cdot \sum_l \sum_r tb^{i,r}_l(\boldsymbol{y}, \boldsymbol{fc}, \boldsymbol{tnp}, \boldsymbol{rnp}) \cdot fb^{i,r}_l(\boldsymbol{td}) \right)$$

where the new term $tdel^{i,r}_{s,p}$ is the time delay of run *r* experienced by users of category *i* at station *s* and on platform *p*.

For this specific application, information about the travel demand (i.e. ***PPM*** outputs) have been drawn from the population census data provided by the ISTAT (‘*Istituto nazionale di STATistica*’). In this way, for each train category, four different matrices have been obtained corresponding to different time intervals (before 7.15, 7.15 – 8.15, 8.15 – 9.15, after 9.15). Since it is not possible to carry out detailed mobility information after 9:15, it is assumed that all trains performing the service between 9:15 and 13:00 have the same matrices.

In Table 6.33, the origin destination matrix of commuters travelling before 7:15 is shown. The procedure adopted to obtain these data can be mainly divided in two steps:

1. For each station, the catchment area is determined. Thus, all municipalities served by that station are identified;
2. for each municipality, the number of people, who takes the train for reaching one of the other municipality served by the other stations of the line, can be extracted from the ISTAT database. Information about the four time intervals is also provided.

Obviously, this process has to be repeated for each station of the whole line.

STAZIONI	Formia	Minturno	Sessa Aurunca	Falciano	Cancello Arnone	Villa Literno	Albanova	San Marcellino	Aversa	Sant'Antimo	Frattamaggiore	Casoria	Napoli	Giugliano	Quarto	Pozzuoli
Formia	0	1	3	1	0	1	1	0	11	1	2	1	211	1	0	3
Minturno	180	0	0	0	0	1	0	0	6	0	2	1	65	0	0	0
Sessa Aurunca	61	0	0	0	0	0	0	0	10	0	0	0	96	0	0	1
Falciano	36	0	0	0	0	0	0	0	13	0	0	0	106	0	0	0
Cancello Arnone	1	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0
Villa Literno	2	0	0	0	0	0	0	0	7	0	0	0	43	0	0	0
Albanova	3	0	0	1	0	0	0	0	17	0	0	1	37	0	0	0
San Marcellino	8	0	0	0	0	0	0	0	4	0	0	0	40	0	0	0
Aversa	147	0	1	3	0	0	0	0	0	0	1	4	945	0	0	3
Sant'Antimo	14	1	0	0	0	0	0	2	64	0	3	1	211	1	0	2
Frattamaggiore	21	1	0	0	0	0	0	0	64	0	0	6	322	1	0	2
Casoria	47	2	0	1	0	0	0	0	52	0	1	0	239	0	0	2
Napoli	194	2	1	0	0	1	0	0	48	1	6	4	6090	7	33	99
Giugliano	17	0	0	1	0	0	0	0	0	0	0	0	86	0	0	1
Quarto	11	0	0	0	0	1	0	0	0	0	0	0	602	0	0	14
Pozzuoli	11	0	0	0	0	1	0	0	6	0	0	0	666	0	6	0

Table 6.33 Estimated OD matrix for travellers before 7:15.

As it can be seen from Table 6.33, Line 2 stations are not included in the OD matrix. This is due to the fact that this procedure can provide train passenger movements among different municipalities but it cannot produce detailed information of movements internal to each district. Therefore, the total number of users travelling within Naples (drawn from the ISTAT database) is equally divided by all metro lines of the city. Furthermore, just three main stations of

Line 2 are considered, that is: Napoli Piazza Garibaldi, Napoli Mergellina and Napoli Campi Flegrei.

By means of the ***OPM***, these matrices have been assigned to the ordinary rail service obtaining the user generalised costs illustrated in Table 6.34. The parameter values of relation 6.6 are the same adopted in previoius applications, that is:

$$\beta_{waiting} = 2.5\,;\ \beta_{on\text{-}board} = 1\,;\ \beta_{VOT} = 5\,euro\,/\,hour$$

The great difference in terms of total cost between the two travel directions is due to the higher attractiveness of the city of Naples which hosts universities, several hospitals, offices, banks and it is therefore interested by a great amount of commuter flow.

	Class of train			
Direction of travel	Intercity	Regional	Metropolitan	Total
Formia-Napoli	€ 1,396.85	€ 11,329.06	€ 18,243.22	€ 30,969.13
Napoli-Formia	€ 1,878.40	€ 2,490.36	€ 13,201.20	€ 17,786.10
				€ 48,539.08

Table 6.34 User generalised costs of the ordinary service.

The time diagrams provided by OpenTrack are illustrated in Figures 6.16 and 6.17.

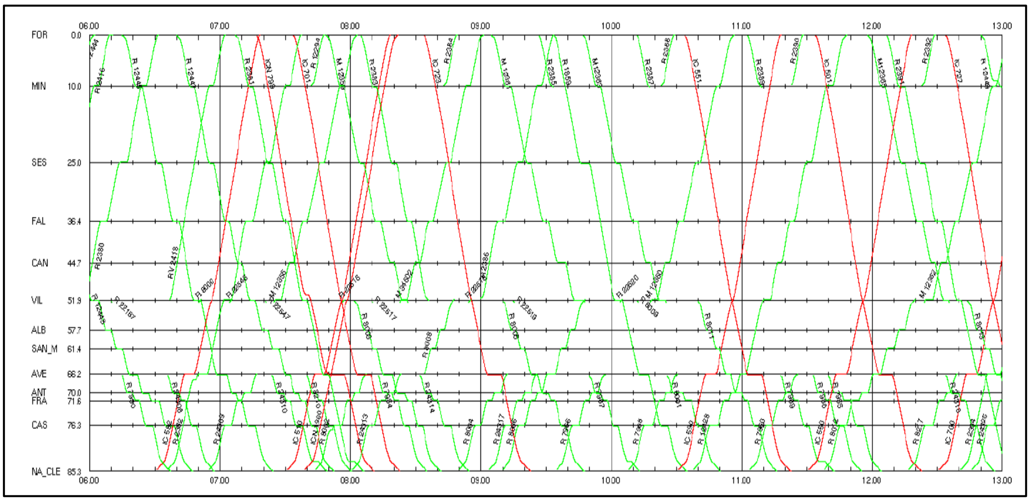


Figure 6.16 Time diagram between ‘Formia – Napoli Centrale’, simulation of the ordinary service condition.

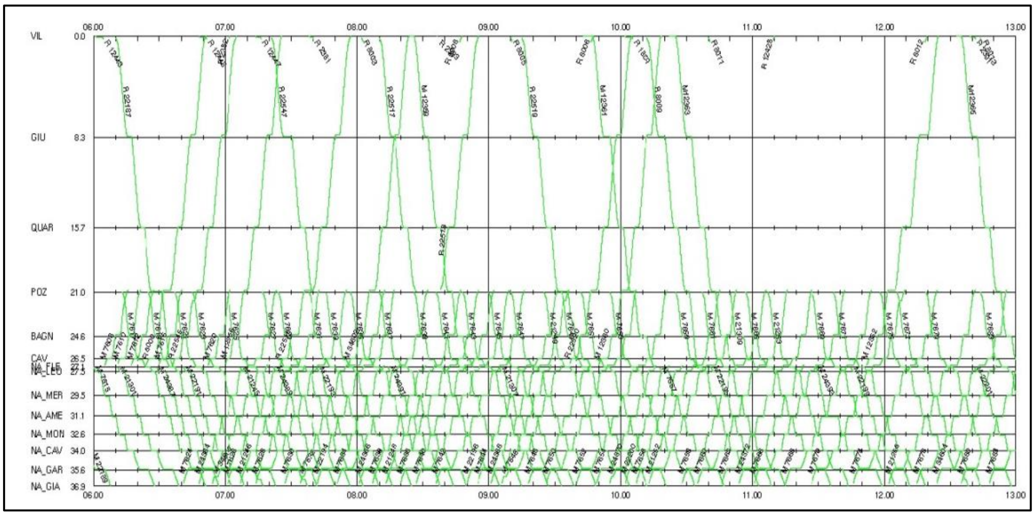


Figure 6.17 Time diagram between ‘Villa literno – Napoli Gianturco’, simulation of the ordinary service condition.

However, Intercity trains in the opposite direction (i.e. Napoli-Formia toward Roma) are also crowded due to the presence of workers and students travelling in direction of the capital.

Four different disrupted scenarios have been simulated. For each disrupted scenario, the following test cases are assessed:

- the non-intervention strategy, namely the dispatcher waits for the end of the disruption without modifying the planned timetable;
- intervention strategy applied after 30 minutes from the occurrence of the failure event;
- intervention strategy applied immediately after the occurrence of the failure event.

The application wants to demonstrate the importance of reacting as rapidly as possible to keep high levels of service quality even in degraded operation regimes. Indeed, since generally no emergency timetables are planned by rail operators, dispatchers have to make decisions they will be responsible for based on their personal experience and sometimes, they prefer to leave the system without any kind of intervention. Adopting a decision support system as the one presented in this paper, it is thus possible to determine feasible solutions which can be standardised and implemented in real time in case of necessity.

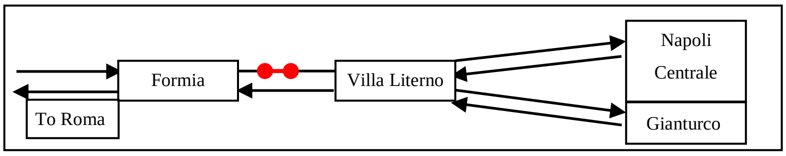


Figure 6.18 Graphic representation of the failure of the Scenario 1.

The first simulated scenario involves a service interruption of two hours (i.e. from 7.00 to 9.00) to the odd track (i.e. direction Formia-Napoli) between the stations of Minturno and Sessa Aurunca (see Figure 6.18), namely before the node of Villa Literno.

In case of non-intervention, all trains on the odd track are forced to wait for the end of the disruption to start again their runs. This causes a propagation of delay which affects 8 consecutive trains (Figure 6.19). As a consequence, the user generalised cost of the whole scenario largely increases (+39.2%) and it is now equal to 76,559.58 € (Table 6.35).

	Class of train			
Direction of travel	Intercity	Regional	Metropolitan	Total
Formia-Napoli	€ 3,593.59	€ 20,292.75	€ 25,887.24	€ 49,773.58
Napoli-Formia	€ 1,878.40	€ 2,706.51	€ 13,201.20	€ 17,786.10
				€ 67,559.68

Table 6.35 User generalised cost scenario 1, without intervention.

As can be seen, trains on the even track are not influenced by the disruption since they can continue to run without any interruption.

Therefore, the first intervention strategy is to enable the trains on the odd track to use alternately the even track. In particular, since this section of the line is not trivialised, the trains directed to Napoli have a speed restriction of 90 km/h.

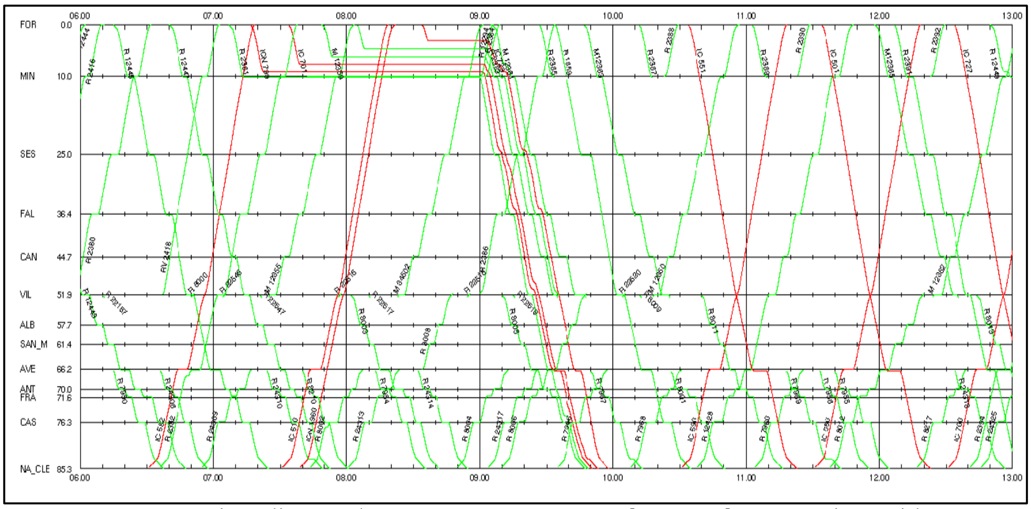


Figure 6.19 Time diagram between '*Formia – Napoli Centrale*', scenario 1 without intervention.

Furthermore, it is assumed that dispatchers wait 30 minutes before giving this instruction.

Analysing the results of the simulation, it comes to light that, due to the alternate circulation, both directions are affected by delay (Figure 6.20). However, delays are notably lower than the previous case and this is confirmed by the user generalised cost of the strategy (Table 6.36).

Indeed, passengers travelling on the odd track experience a lower disutility while users on the opposite direction undergo a slight increase of cost. Nevertheless, the strategy reduces the total cost of 22.2% since it passes from 67,559.68 € (without intervention) to 52,522.33 €.

	Class of train			
Direction of travel	Intercity	Regional	Metropolitan	Total
Formia-Napoli	€ 1,778.38	€ 12,684.36	€ 19,044.51	€ 33,507.26
Napoli-Formia	€ 2,475.00	€ 2,509.45	€ 14,060.62	€ 19,045.07
				€ 52,522.33

Table 6.36 User generalised cost scenario 1, first intervention strategy.

The last strategy of Scenario 1 consists in enabling the alternate circulation immediately after the breakdown occurrence. As expected, delay propagation in both directions is lower than the previous cases (Figure 6.21) which results in calming down more rapidly the negative effects of the disruption.

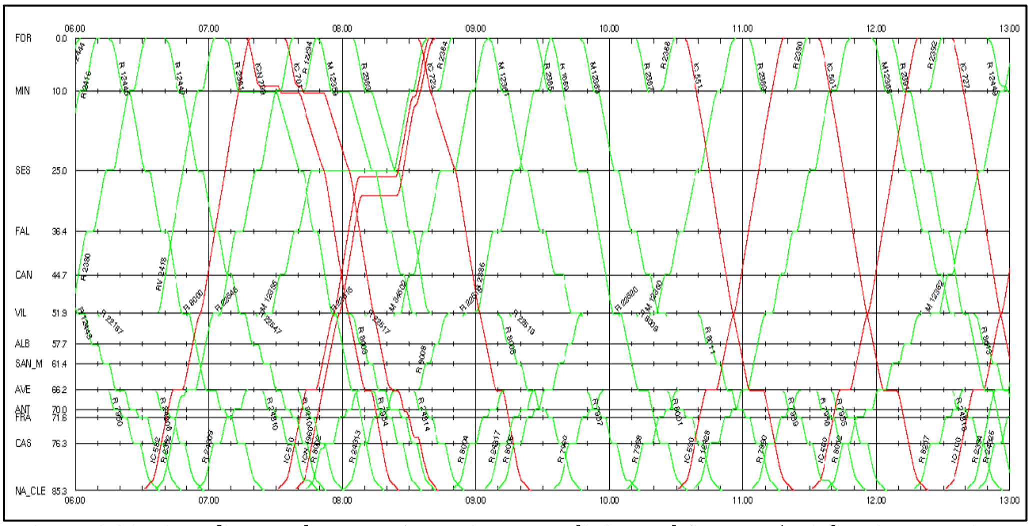


Figure 6.20 Time diagram between '*Formia – Napoli Centrale*', scenario 1 first intervention strategy.

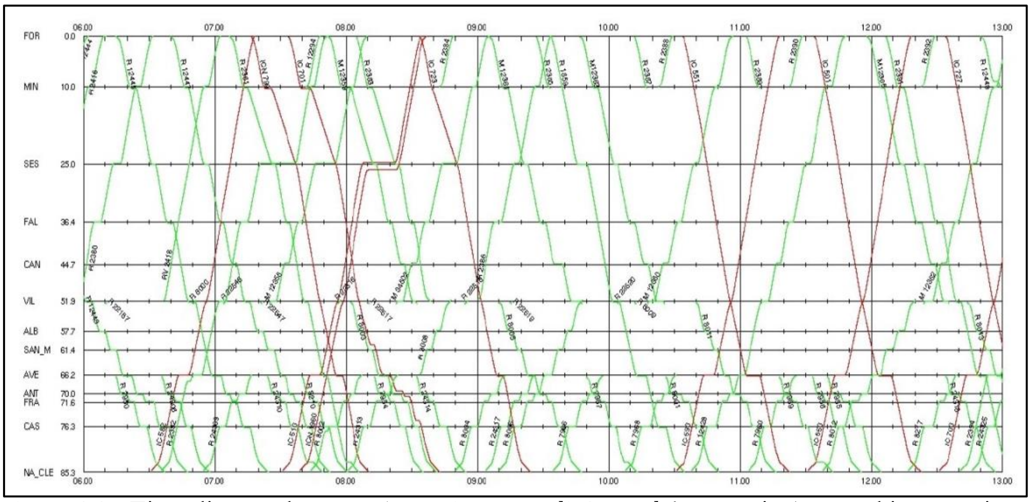


Figure 6.21 Time diagram between '*Formia – Napoli Centrale*', scenario 1 second intervention strategy.

The user generalised cost is in this case equal to 50,543.25 € (Table 6.37), just 4,1% higher than the ordinary service conditions.

	Class of train			
Direction of travel	Intercity	Regional	Metropolitan	Total
Formia-Napoli	€ 1,651.95	€ 11,444.52	€ 18,780.51	€ 31,876.98
Napoli-Formia	€ 2,345.16	€ 2,908.74	€ 13,412.36	€ 18,666.27
				€ 50,543.25

Table 6.37 User generalised cost scenario 1, second intervention strategy.

Scenario 2 is still based on a service interruption of two hours (from 7.00 to 9.00) to the odd track (direction Formia – Napoli) between Casoria and Napoli Centrale, that is the last part of the line before reaching the main station of the city of Naples.

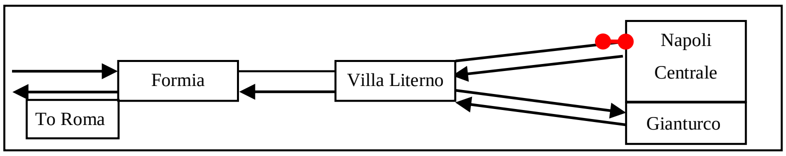


Figure 6.22 Graphic representation of the failure of the Scenario 2.

The non-intervention strategy produces large delays which, this time, affect both the even and the odd direction. This is due to the fact that, the faulty section is near the Napoli Centrale station and therefore, because of the interlocking system, some paths are blocked and thus not available (Figure 6.23).

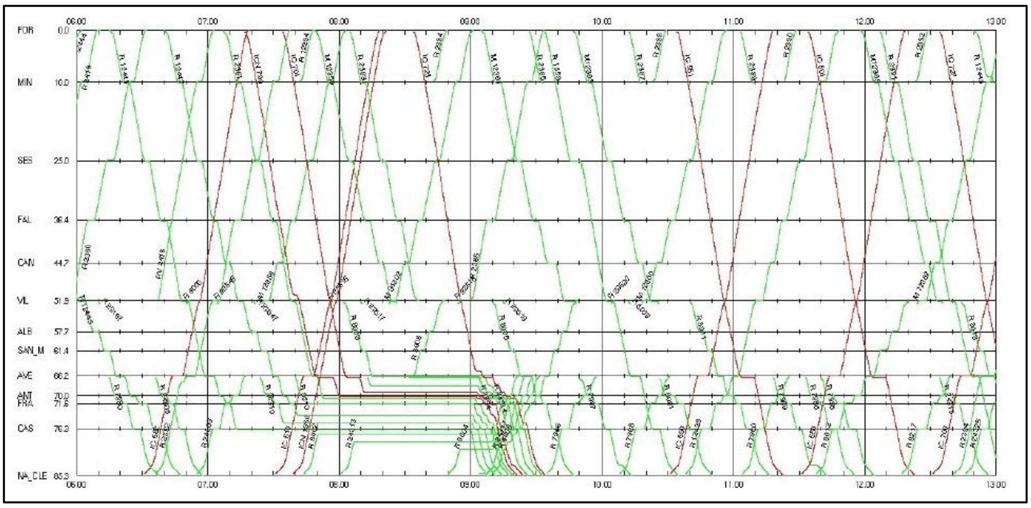


Figure 6.23 Time diagram between '*Formia – Napoli Centrale*', scenario 2 without intervention.

As regards the user generalised cost, Table 6.38 shows the increase of disutility perceived by passengers (+74.4% respect to the ordinary service). Moreover, it is worth noting that metropolitan trains are not delayed inasmuch as they just run on the 'Formia – Gianturco' line, which is not affected.

	Class of train			
Direction of travel	Intercity	Regional	Metropolitan	Total
Formia-Napoli	€ 2,753.00	€ 42,849.01	€ 18,423.22	€ 64,025.23
Napoli-Formia	€ 1,878.40	€ 2,908.74	€ 13,201.20	€ 23,061.50
				€ 87,086.73

Table 6.38 User generalised cost scenario 2, without intervention.

By enabling the alternate circulation after 30 minutes from the start of the disruption (i.e. first intervention strategy), the benefits to the service are remarkable (Table 6.39). User generalised cost is 57,228.27 € with an increase of just 17.9% and a reduction in comparison with the non-intervention strategy of 34.2%.

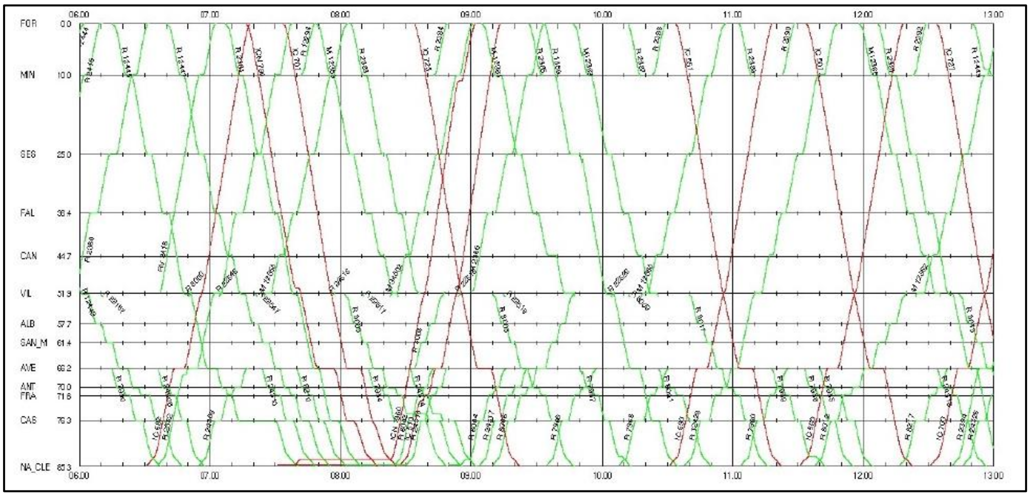


Figure 6.24 Time diagram between '*Formia – Napoli Centrale*', scenario 2 first intervention strategy.

	Class of train			
Direction of travel	Intercity	Regional	Metropolitan	Total
Formia-Napoli	€ 1,447.25	€ 13,611.27	€ 18,423.22	€ 33,481.74
Napoli-Formia	€ 5,358.79	€ 5,186.53	€ 13,201.20	€ 23,746.52
				€ 57,228.27

Table 6.39 User generalised cost scenario 2, first intervention strategy.

Obviously, the situation further improves if the alternate circulation is imposed as rapidly as possible (second intervention strategy). In fact both even and odd trains collect fewer delay and this results in a higher service quality level during the degraded regime (Table 6.40).

	Class of train			
Direction of travel	Intercity	Regional	Metropolitan	Total
Formia-Napoli	€ 1,446.33	€ 13,014.24	€ 18,423.22	€ 32,883.79
Napoli-Formia	€ 2,163.92	€ 5,005.32	€ 13,201.20	€ 20,370.45
				€ 53,245.23

Table 6.40 User generalised cost scenario 2, second intervention strategy.

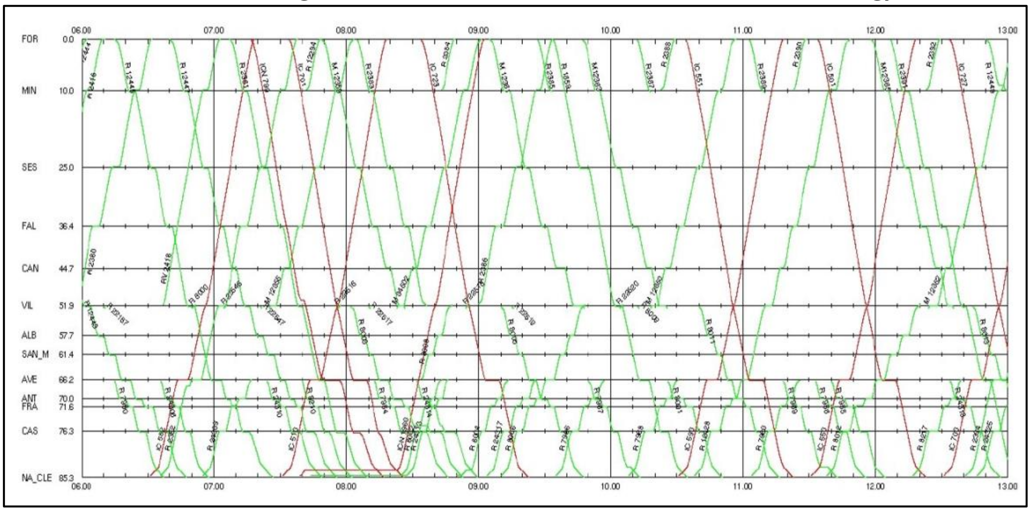


Figure 6.25 Time diagram between '*Formia – Napoli Centrale*', scenario 2 second intervention strategy.

Scenario 3, by contrast, considers a service interruption of two hours to the odd track (direction Formia – Napoli Gianturco) between the stations of Pozzuoli and Bagnoli, namely a small section of the metropolitan branch which crosses the city of Naples (Figure 6.26).

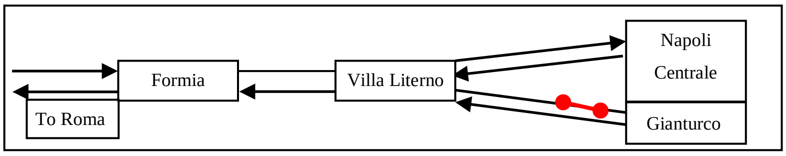


Figure 6.26 Graphic representation of the failure of the Scenario 3.

Actually, due to the high frequency, the number of trains affected by the disruption is higher than the previous two scenarios (Figure 6.27). In addition, metropolitan trains are more crowded and this is the reason why the user generalised cost is more than three times higher than the ordinary service (Table 6.41).

	Class of train			
Direction of travel	Intercity	Regional	Metropolitan	Total
Formia-Napoli	€ 1,396.85	€ 11,329.06	€ 99,328.92	€ 112,054.83
Napoli-Formia	€ 1,878.40	€ 2,490.36	€ 20,054.60	€ 24,423.35
				€ 136,478.18

Table 6.41 User generalised cost scenario 3, without intervention strategy.

As expected, passengers of intercity and regional trains do not perceive any increase in delay and travel time.

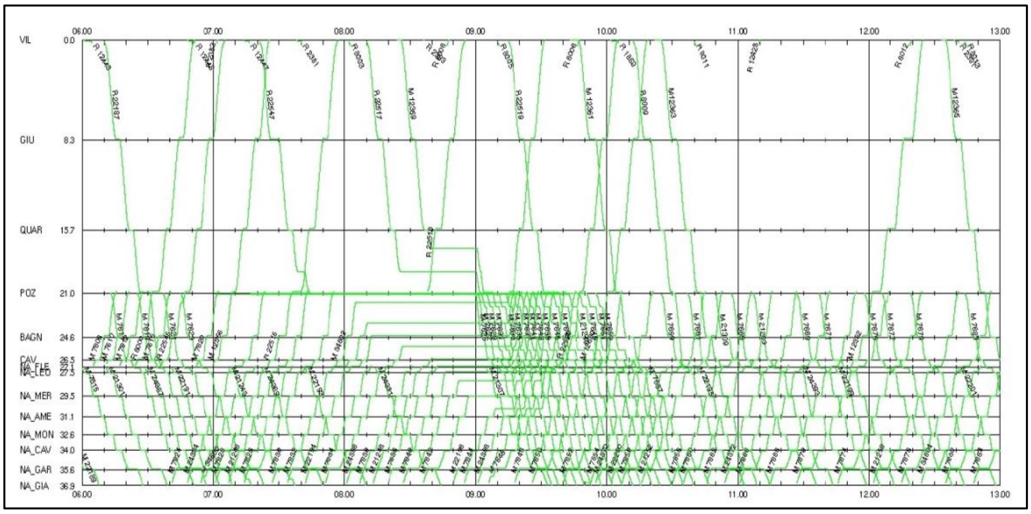


Figure 6.27 Time diagram between '*Villa Literno – Napoli Gianturco*', scenario 3 without intervention.

Similar to what done with the other scenarios, the alternate circulation regime is introduced after 30 minutes from the section closing (first intervention strategy). However, in this case, since the branch 'Villa Literno – Gianturco' is totally trivialised, there are no speed restrictions and trains can run to the best of their performance. As a consequence, running times remain the same in both directions.

Table 6.42 summarises the outputs of the assignment process. Basically, the cost increases of 60.2% (77,822.52 €) which is quite far from the value of the non-intervention strategy (136,478.18 €).

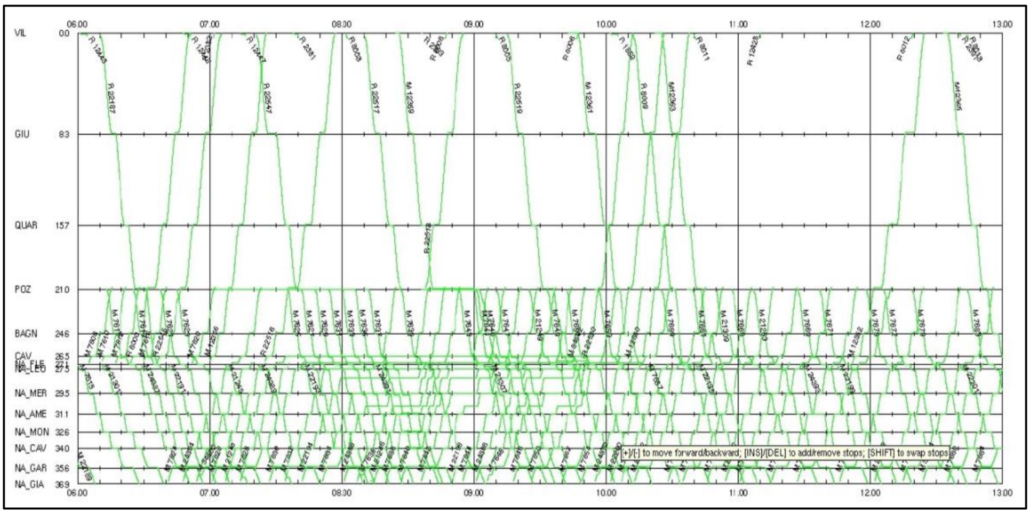


Figure 1) Time diagram between ‘Villa Literno – Napoli Gianturco’, scenario 3 first intervention strategy.

	Class of train			
Direction of travel	Intercity	Regional	Metropolitan	Total
Formia-Napoli	€ 1,396.85	€ 11,329.06	€ 31,199.71	€ 43,925.62
Napoli-Formia	€ 1,878.40	€ 2,490.36	€ 29,528.15	€ 33,896.90
				€ 77,822.52

Table 6.42 User generalised cost scenario 3, first intervention strategy.

Once again, a prompt reaction (i.e. second intervention strategy) reduces the disturbance effects especially in case of high frequency service where conflicts between trains arise very rapidly (Figure 6.28). Indeed, as shown by Table 6.43, user generalised cost is equal to 54,396.37 €, that is 12.1% higher than the ordinary case.

	Class of train			
Direction of travel	Intercity	Regional	Metropolitan	Total
Formia-Napoli	€ 1,396.85	€ 11,329.06	€ 22,720.69	€ 35,446.60
Napoli-Formia	€ 1,878.40	€ 2,490.36	€ 14,581.02	€ 18,949.77
				€ 54,396.37

Table 6.43 User generalised cost scenario 3, second intervention strategy.

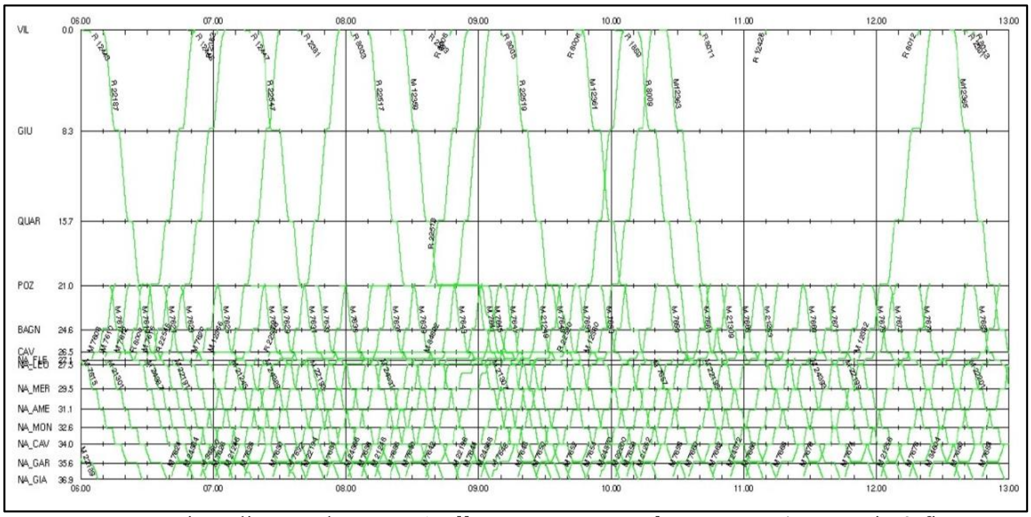


Figure 6.28 Time diagram between '*Villa Literno – Napoli Gianturco*', scenario 3 first intervention strategy.

Completely different is the disruption simulated in scenario 4 where, because of an accident involving a person, the interruption of 4 hours (i.e. from 7.00 to 11.00) of both tracks near Albanova station is considered. This forces the infrastructure manager to interrupt the service between Villa Literno and Napoli Centrale (Figure 6.29).

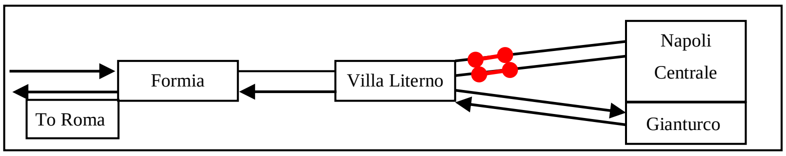


Figure 6.29 Graphic representation of the failure of the Scenario 3.

Therefore, all intercity and regional trains running through this section are forced to stop and experience a great delay. This inconvenience is highlighted by the objective function values in Table 6.44.

	Class of train			
Direction of travel	Intercity	Regional	Metropolitan	Total
Formia-Napoli	€ 6,746.84	€ 101,735.04	€ 18,243.22	€ 126,725.10
Napoli-Formia	€ 7,716.80	€ 12,207.33	€ 13,201.20	€ 33,125.33
				€ 159,850.43

Table 6.44 User generalised cost scenario 4, without intervention.

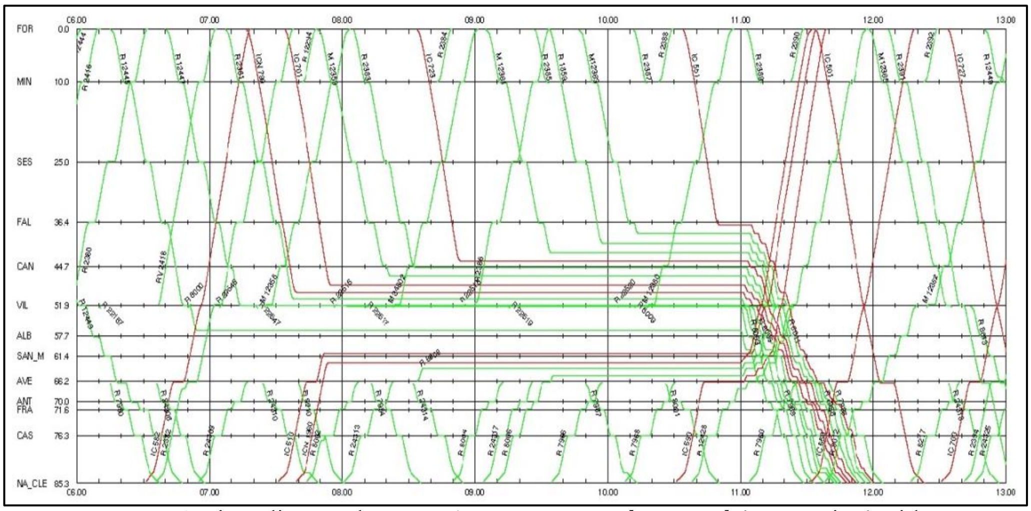


Figure 6.30 Time diagram between '*Formia – Napoli Centrale*', scenario 4 without intervention.

As first intervention strategy, after 30 minutes from the failure event, it is supposed to use the node in Gianturco station to keep the connection between Formia (or Rome) and Naples via Villa Literno, so as to satisfy a great part of passenger flow. This means that, within the branch 'Villa Literno – Gianturco', intercity, regional and metropolitan trains run all together. In addition, the trains within the section Villa Literno – Napoli Centrale' during the breakdown occurrence, move toward Napoli Centrale enabling the connection with new runs travelling on the metropolitan line. Figure 6.32 and 6.31 show the time diagrams of both lines while user generalised costs are provided in Table 6.45.

	Class of train			
Direction of travel	Intercity	Regional	Metropolitan	Total
Formia-Napoli	€ 2,033.32	€ 45,830.53	€ 18,454.87	66,318.72
Napoli-Formia	€ 4,270.18	€ 12,135.93	€ 13,132.09	€ 29,538.20
				€ 95,856.93

Table 6.45 User generalised cost scenario 4, first intervention strategy.

As it can be seen, the higher discomfort experienced by metropolitan passengers is compensated by the great reduction in that of intercity and regional train users. However, total cost is still high since many commuters remain un-served for a long period especially within the section 'Villa Literno – Napoli Centrale'.

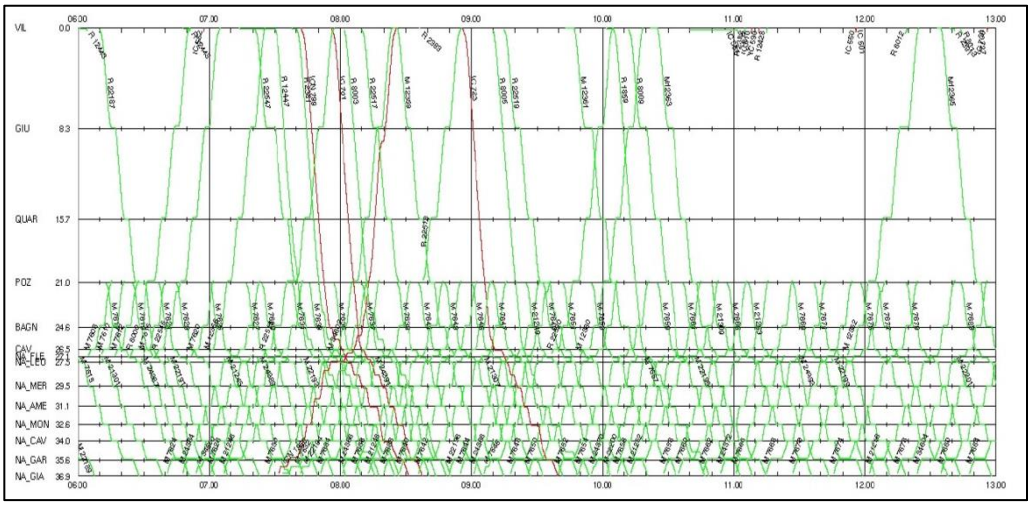


Figure 6.31 Time diagram between '*Villa Literno – Napoli Gianturco*', scenario 4 first intervention strategy.

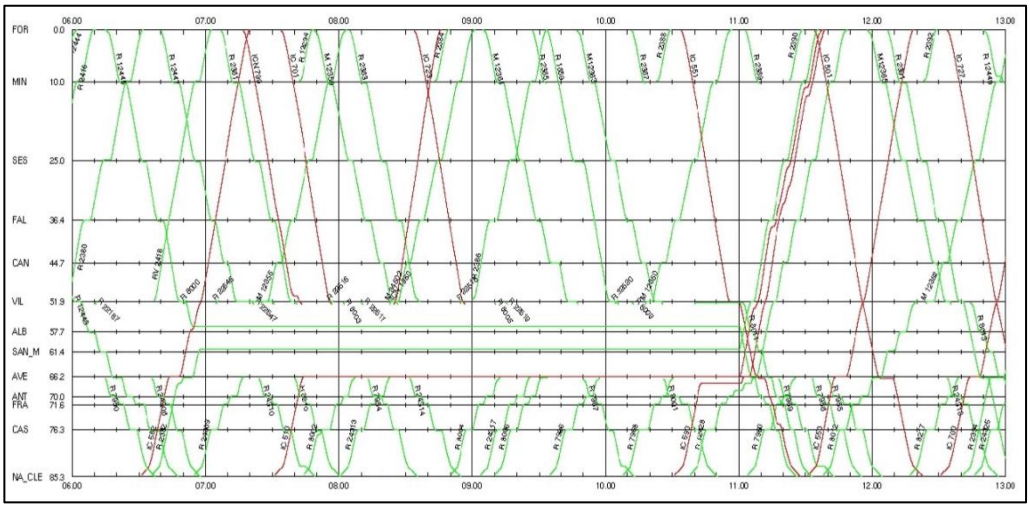


Figure 6.32 Time diagram between '*Formia – Napoli Centrale*', scenario 4 first intervention strategy.

Likewise, the second intervention strategy concerns the same rescheduling solutions which are applied immediately after the section closing. In this case, the benefits are not so remarkable (see Table 6.46). Indeed, the total cost (94,093.60 €) is just 1.6% lower than the one of the previous recovery plan (95,856.93 €). However, both values are much better than the non-intervention solution which, although is the optimal one from an operational point of view, does not take customer needs into account.

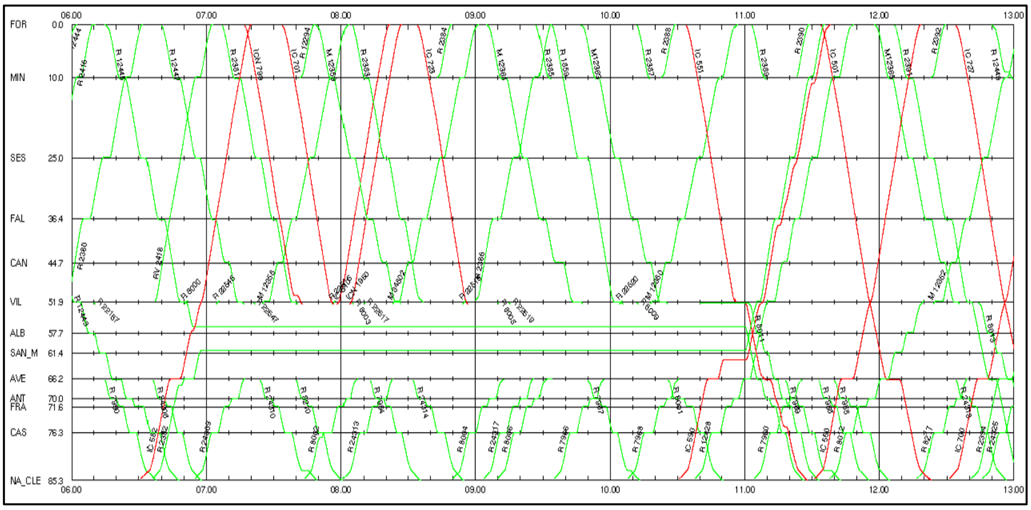


Figure 6.33 Time diagram between '*Formia – Napoli Centrale*', scenario 4 first intervention strategy.

	Class of train			
Direction of travel	Intercity	Regional	Metropolitan	Total
Formia-Napoli	€ 2,033.32	€ 45,614.72	€ 18,483.57	66,131.61
Napoli-Formia	€ 2,363.95	€ 12,066.19	€ 13.481.86	€ 27,911.99
				€ 94,093.60

Table 6.46 User generalised cost scenario 4, second intervention strategy.

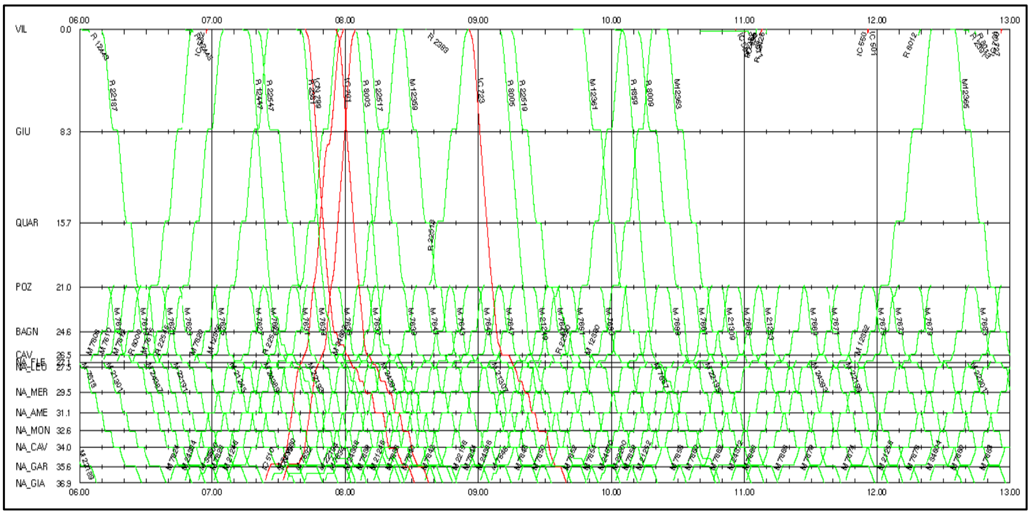


Figure 6.34 Time diagram between '*Villa Literno – Napoli Gianturco*', scenario 4 first intervention strategy.

In conclusion, this test case on a conventional rail line in the south of Italy has proved the benefits of performing an off-line assessment of intervention strategies which should be proposed during disruptions. Basically, dispatchers, supported by this kind of analysis, could make rapid decisions which are out of their responsibility.

CHAPTER 7: CONCLUSION

7.1 Resume of the main achievements.

This thesis proposed a decision support system for managing railway networks in any kind of service conditions with particular attention to the analysis of recovery solutions after the occurrence of breakdowns or failures. The model is based on an off-line procedure which simulates rail operations and assigns passenger flows to the service. This approach enables to determine intervention strategies which maximise service quality perceived by customers.

The idea is to introduce a new methodology which should not substitute dispatchers but provides them feedbacks about decisions that have to be made during critical failure events. To this purpose, the procedure is presented as a multidimensional optimisation problem whose objective function expresses the generalised cost of users travelling within the network during the simulation. It is composed of four models which guarantee the complete analysis of the rail network, namely:

- a '*Failure Model*' whose aim is to assess the breakdown contexts with the higher probability of occurrence. In this way, the failure scenarios which are worth analysing can be selected;
- a '*Service Simulation Model*' which simulates microscopically rail operations and performances through the adoption of a synchronous microscopic simulation software performing both deterministic and stochastic simulations.
- a '*Supply Model*' for the definition of the characteristics and performances of all public transport systems within the study area. In this way, split demand among transport modes can be taken into account;
- a '*Travel Demand Model*' consisting of two sub-models, the Pre-Platform Model and the On-Platform Model. The first one estimates the number of passenger arriving at stations according to the characteristics of all transport modes (i.e. interaction with the Supply Model which

generates a fixed point problem). The second one enables the dynamic assignment of passengers to the rail service. As a result, users' behaviour on the platform is simulated taking into account the maximum capacity of each train and estimating the dwell time necessary to complete the boarding/alighting process. The flow assignment produces a new fixed point problem inasmuch as travel demand is influenced by rail performances which, in turn, are affected by passenger flow levels on the platform.

Therefore, the model simulates rail operations during failure scenarios and possible intervention strategies considering also travel demand. Each solution is evaluated in order to select at the end of the process, the optimal intervention strategy from passengers' standpoint.
The application of the procedure to the main failure contexts determined by the Failure Model can bring to the determination of a database with suggests the decisions to make for any kind of events.

However, the contributions provided by this research work are numerous and are not limited to the simple definition of intervention strategies after disruptions.
Microscopic simulation models generally focus just on the simulation of the operational service but neglect passengers flow. For this reason, a tool for the assignment of passenger flows to the rail service working in combination with a microscopic simulation software is proposed, whose name is OPM 1.0. This program, after the simulation of the scenario, is able to determine load diagrams, passenger trip information and platform congestion. As a consequence, it is thus possible to assess demand peaks, temporary capacity variations, temporary over-saturation of supply elements, and formation and dispersion of queues.

As already said, the interaction between passengers on the platform and rail service results in a fixed point problem whose resolution provides the dwell times at each station. A specific module called DwTE 1.0 is dedicated to the estimation of dwell times as flow dependent and gives information about crowding levels within each coach.

The stability and the robustness of intervention strategies is not neglected. In particular, implementing a two-step procedure, the proposed approach is first adopted to select a set of optimal intervention strategies under deterministic assumptions; then, the optimal solutions are assessed by means of several stochastic simulations, varying kinematic parameters, travel demand levels and dwell times at stations.

Numerous applications on the Line 1 of Naples metro system are carried out. Outputs demonstrate the benefits of this model in providing intervention strategies which fulfil customers' needs and keep high levels of service quality even during degraded service regimes. Further analyses show that breakdown severity and travel demand levels have to be taken into account since they can affect recovery solutions.

Sensitivity analyses on recovery strategies are also applied to Line 1 metro system. They highlight that the optimal solution found in the deterministic phase still performs well also in the stochastic assessment. However, this study gives important indications about the error degree when only the deterministic approach is implemented.

Other extensive computational experiments show the possibility to apply the proposed model also in different contexts. A more complex objective function indeed proves how to consider effectiveness and efficiency indexes together with the user generalised cost. In this way, the procedure provides complete information about service conditions achieving benefits on behalf of customers, train operating companies as well as community.

A test case on the regional line 'Formia – Napoli' exposes also the feasibility of the model for the analysis of conventional rail systems. In particular, introducing some changes to the On-Platform Model, numerical results highlight the importance of reacting as rapidly as possible so as to reduce users' discomfort. Obviously, this can be done only if off-line procedure as the one presented in this thesis are implemented prior the occurrence of the failure incident. Dispatchers can thus make decisions out of their responsibility following planned instructions which have already been tested.

Finally, the combination with a macro-optimisation model enhances the potentiality of the proposed model. In fact, when dealing with the design of new timetables and rolling stock schedules, the implementation of an iterative approach is able to determine feasible strategies in short time. In particular, the macro model computes a new timetable considering service costs and passengers' inconvenience; the micro model instead checks whether the strategy is feasible and provides feedbacks which make the macro procedure improve the solution. An experimental test on Line 1 proves the benefits of this new methodology when it is necessary to plan emergency timetables.

7.2 Future research prospects.

Due to the extent of the proposed approach, research prospects are numerous. First of all, it is undoubtedly that it is necessary to investigate the feasibility of the model under a larger number of disruption events and to test it also on more complex networks.

The tools provided in this thesis, that is OPM 1.0 and DwTE 1.0, should be integrated within a microscopic simulation software. This would speed up the assignment process and reduce the number of input files for running the programs. In addition, it would enable the interaction between travel demand and rail service on-line. As a result, instead of solving a fixed point problem, during the simulation when a train is approaching a station, the software could ask DwTE 1.0 the dwell time which is required in proportion to travellers on the platform. In addition, knowing the number of on-board passengers within each coach, ITS systems could also be designed to inform customers about the crowding levels of the approaching train and to suggest them how to place along the platform or what coach of the train should be preferred.

Dwell time values are also vital for the definition of energy saving strategies. Indeed, they are part of a time window which affects driving behaviour. Including also this kind of analysis in the proposed model could bring to the evaluation of a trade-off between the saving of energy and passenger increase of cost.

Another important development should concern the definition of new dynamic OD matrix correction procedures. Indeed, during the course of this thesis, the estimation of travel demand has not been examined in depth since the objectives of the research work were different. However, important contributions could be carried out on this topic using OPM 1.0 as passenger assignment tool.

Finally, the combination with macro-optimisation models should be enhanced and improved. Especially when the set of alternatives which has to be evaluated is numerous, exhaustive approaches are not advantageous to find solutions in reasonable time even for an off-line procedure. Therefore, developing new optimisation techniques could help to reduce the number of microscopic simulations which results in a drastically decrease of computational time.

REFERENCES

Scientific references:

Acuna-Agost, R., Michelon, P., Feillet, D., & Gueye, S. (2011a). A MIP-based local search method for the railway rescheduling problem. Networks 57 (1), 69–86.

Acuna-Agost, R., Michelon, P., Feillet, D., & Gueye, S. (2011b). SAPI: statistical analysis of propagation of incidents. A new approach for rescheduling trains after disruptions. European Journal of Operational Research 215 (1), 227–243.

Albrecht, A.R., Panton, D.M., & Lee, D.H. (2013). Rescheduling rail networks with maintenance disruptions using problem space search. Computers & Operations Research 40, 703–712.

Albrecht, T. (2009). Automated timetable for demand-oriented service on suburban railways. Public Transp, 1(1), 5-20.

Albrecht, T., Binder, A., & Gassel, C. (2013). Applications of real-time speed control in rail-bound public transportation systems. IET Intelligent Transport Systems, 7(3), 305-314.

Barabino, B., Deiana, E., & Mozzoni, S. (2013). The quality of public transport service: The 13816 standard and a methodological approach to an Italian case. Ingegneria Ferroviaria, Volume 68, Issue 5, 475-499.

Bellei, G., Gentile, G., & Papola, N. (2005). A within-day dynamic traffic assignment model for urban road networks. Transportation Research Part B: Methodological, 39(1), 1-29.

Blum, J.R. (1954). Multidimensional stochastic approximation methods. Ann. Math. Stat., 25, 737-744.

Buchmueller, S., Weidmann, U., & Nash, A. (2010). Development of a dwell time calculation model for timetable planning. Timetable Planning and Information Quality, 105.

Butcher J.C. (1987). The numerical analysis of ordinary differential equations: Runge-Kutta and general linear methods. Wiley Interscience, New York (NY), USA.

Cacchiani, V., Huisman, D., Kidd, M., Kroon, L., Toth, P., Veelenturf, L., & Wagenaar, J. (2014). An overview of recovery models and algorithms for real-time railway rescheduling. Transportation Research Part B: Methodological, 63, 15-37.

Cadarso, L., Marín, Á., & Maróti, G. (2013). Recovery of disruptions in rapid transit networks. Transportation Research Part E 53, 15–33.

Caimi, G., Fuchsberger, M., Laumanns, M., & Lüthi, M. (2012). A model predictive control approach for discrete-time rescheduling in complex central railway station areas. Computers & Operations Research 39, 2578–2593.

Campora, S., Botte, M., D'Acierno, L., Placido, A., & Montella, B. (2014). The dynamic interaction between passenger flows and rail service. Abstract book of the 17th Meeting of the Euro Working Group on Transportation, pp. 44-45, Seville, Spain.

Canca, D., Barrena, E., Laporte, G., & Ortega, F.A. (2014). A short-turning policy for the management of demand disruptions in rapid transit systems. Annals of Operations Research, . Article in Press.

Cantarella, G.E. (1997). A general fixed-point approach to multi-mode multi-user equilibrium assignment with elastic demand. Transportation Science, 31, 107-128.

Cascetta, E, Papola, A., Marzano, V., Simonelli, F., & Vitiello, I. (2013a). Quasi-dynamic estimation of o-d flows from traffic counts: formulation, statistical validation and performance analysis on real data. Transportation research Part B, 55, 171-187.

Cascetta, E. (2009). Transportation systems analysis: Models and applications. Springer: New York (NY), USA.

Cascetta, E., & Cartenì, A. (2014). The hedonic value of railways terminals. A quantitative analysis of the impact of stations quality on travellers behaviour. Transportation Research Part A: Policy and Practice, 61, 41-52.

Cascetta, E., Cartenì, A., & Carbone, A. (2013b). Stations quality, aesthetics and attractiveness of rail transport: Empirical evidence and mathematical models. Ingegneria Ferroviaria, 68 (3), 241-261.

Cascetta, E., Cartenì, A., & Henke, I. (2014). The quality in public transportation. the campania regional metro system. Ingegneria Ferroviaria, 69 (4), 307-324.

Cipriani, E., Florian, M., Mahut, M. & Nigro, M. (2011). A gradient approximation approach for adjusting temporal origin-destination matrices. Transportation Research Part C, 19(2), 270–282.

Corman, F., D'Ariano, A., & Hansen, I.A. (2010a). Disruption handling in large railway networks. In: Proceedings of Computers in Railways, Beijing, China, 629–640.

Corman, F., D'Ariano, A., Hansen, I.A., Pacciarelli, D., & Pranzo, M. (2011a). Dispatching trains during seriously disrupted traffic situations. In: Proceedings of the IEEE International Conference on Networking, Sensing and Control, 323–328.

Corman, F., D'Ariano, A., Pacciarelli, D., & Pranzo, M. (2009). Evaluation of a green wave policy in real-time railway traffic management. Transportation Research Part C 17 (6), 607–616.

Corman, F., D'Ariano, A., Pacciarelli, D., & Pranzo, M. (2010b). Centralized versus distributed systems to reschedule trains in two dispatching areas. Public Transport 2 (3), 219–247.

Corman, F., D'Ariano, A., Pacciarelli, D., & Pranzo, M. (2012). Bi-objective conflict detection and resolution in railway traffic management. Transportation Research Part C 20 (1), 79–94.

Corman, F., D’Ariano, A., Pacciarelli, D., & Pranzo, M., (2010c). A tabu search algorithm for rerouting trains during rail operations. Transportation Research Part B 44 (1), 175–192.

Corman, F., D’Ariano, A., Pranzo, M., & Hansen, I.A. (2011b). Effectiveness of dynamic reordering and rerouting of trains in a complicated and densely occupied station area. Transportation Planning and Technology 34 (4), 341–362.

Curtius, E.W. & Kniffler, A. (1950). Neue Erkenntnisse uber die Haftung zwischen trelbrad und Schiene, Elektrische Bahnen 21 No. 9, 201-210

D’Acierno, L., Gallo, M., Montella, B., & Placido, A. (2012). Analysis of the interaction between travel demand and rail capacity constraints. WIT Transactions on the Built Environment, 128, 197-207.

D’Acierno, L., Gallo, M., Montella, B., & Placido, A. (2013a). Evaluation of travel demand impacts in the case of rail system failures. Procedia: social & behavioral sciences, vol. 87, 75-84.

D’Acierno, L., Gallo, M., Montella, B., & Placido, A. (2013b). The definition of a model framework for managing rail systems in the case of breakdowns. Proceedings of IEEE ITSC 2013 – 16th International IEEE Annual Conference on Intelligent Transportation Systems, pp. 1059-1064, The Hague, The Netherlands.

D’Ariano, A. (2008). Improving Real-Time Train Dispatching: Models, Algorithms and Applications. PhD thesis, TRAIL Thesis Series T2008/6, The Netherlands.

D’Ariano, A., & Pranzo, M. (2009). An advanced real-time train dispatching system for minimizing the propagation of delays in a dispatching area under severe disturbances. Networks and Spatial Economics 9 (1), 63–84.

D’Ariano, A., Corman, F., Pacciarelli, D., & Pranzo, M. (2008a). Reordering and local rerouting strategies to manage train traffic in real time. Transportation Science 42 (4), 405–419.

D’Ariano, A., Corman, F., Pacciarelli, D., & Pranzo, M. (2008a). Reordering and local rerouting strategies to manage train traffic in real time. Transportation Science 42 (4), 405–419.

D’Ariano, A., Pacciarelli, D., & Pranzo, M. (2007a). A branch and bound algorithm for scheduling trains on a railway network. European Journal of Operational Research 183 (2), 643–657.

D’Ariano, A., Pacciarelli, D., & Pranzo, M. (2008b). Assessment of flexible timetables in real-time traffic management of a railway bottleneck. Transportation Research Part C 16 (2), 232–245.

D’Ariano, A., Pranzo, M., Hansen, I.A., (2007b). Conflict resolution and train speed co-ordination for solving real-time timetable perturbations. IEEE Transactions on Intelligent Transportation Systems 8 (2), 208–222.

Dijkstra, E.W. (1959). A Note On Two Problems In Connexion With Graphs, Numerische Mathematik, 1, 269-271.

Dollevoet, T., Huisman, D., Schmidt, M., & Schöbel, A. (2012). Delay management with rerouting of passengers. Transportation Science 46, 74–89.

Ercolani, M., Placido, A., D’Acierno, L., & Montella, B. (2014). The use of microsimulation models for the planning and management of metro systems. WIT Transactions on the Built Environment 135, 509-521.

Esveld, C. (2001). Modern Railway Track. MRT-Productions, second edition.

Flamini, M., & Pacciarelli, D. (2008). Real time management of a metro rail terminus. European Journal of Operational Research 189 (3), 746–761.

Frederix, R., Viti, F., & Tampère, C.M.J. (2013). Dynamic origin-destination estimation in congested networks: Theoretical findings and implications in practice. Transportmetrica A: Transport Science, 9 (6), 494-513.

Goverde, R. M. (2007). Railway timetable stability analysis using max-plus system theory. Transportation Research Part B: Methodological, 41(2), 179-201.

Goverde, R. M. (2010). A delay propagation algorithm for large-scale railway traffic networks. Transportation Research Part C: Emerging Technologies,18(3), 269-287.

Goverde, R.M.P., Hooghiemstra, G., & Lopuhaä, H.P. (2001). Statistical analysis of train traffic. DUP Science, Delft, the Netherlands.

Gröger, T. (2002). Simulation Der Fahrplanerstellung Auf Der Basis Eines Hierarchischen Trassenmanagements Und Nachweis Der Stabilität Der Betriebsabwicklung. Dissertation, Veröffentlichungen des Verkehrswissenschaftlichen Institutes der RWTH Aachen, Vol. 60.

Hansen, I. A., Wiggenraad, P. B., & Wolff, J. W. (2013). Benchmark analysis of railway networks and undertakings. Proceedings of the 5th International Seminar on "Railway Operations Modelling and Analysis - RailCopenhagen, Copenhagen, Denmark.

Hansen, I.A. & Pachl. J. (2008). Railway timetable and traffic: analysis-modelling-simulation, Eurail Press.

Hauptmann, D. (2000). Automatiche und Diskriminierungsfreie Ermittlung Von Fahrplantrassen in Beliebig Großen Netzen Spurgefuhrter Verkehrssysteme (Automatic and Non-Discriminatory Train Path Allocation In Railway Networks Of Arbitrary Size), Dissertation am Institut fur Verkehrswesen, Eisenbahnbau und betrieb (Schriftenreihe No. 54), Universitat Hannover, Hestra Verlag.

Hirai, C., Kunimatsu, T., Tomii, N., Kondou, S., & Takaba, M. (2009). A train stop deployment planning algorithm using a petri-net-based modelling approach. Quarterly Report of RTRI 50 (1), 8–13.

Huerlimann, D. (2001). Objektorientierte Modellierung von Infrastrukturelementen und Betriebsvorgängen im Eisenbahnwesen. Dissertation am Institut für Verkehrsplanung und Transporttechnik (IVT) der ETH Zürich.

Jacobs, J. (2004). Reducing delays by means of computer-aided “on-the-spot” rescheduling. Computer in railways IX, 603-612.

Kanai, S., Shiina, K., Harada, S., & Tomii, N. (2011). An optimal delay management algorithm from passengers’ viewpoints considering the whole railway network. Journal of Rail Transport Planning & Management 1, 25–37.

Kecman, P., Corman, F., D’Ariano, A., & Goverde, R. M. (2013). Rescheduling models for railway traffic management in large-scale networks. Public Transport, 5(1-2), 95-123.

Kettner, M., Sewcyk, B., & Eickmann, C. (2003). Integrating microscopic and macroscopic models for railway network evaluation. In Proceedings of the European transport conference. Washington: Association for European Transport.

Kondo, R. (1980). Introduction of advanced type of automatic train stop system. In Japanese Railway Engineering 4/1980.

Kumazawa, K., Hara, K., & Koseki, T. (2010). A novel train rescheduling algorithm for correcting disrupted train operations in a dense urban environment. In: Hansen, I.A. (Ed.), Timetable Planning and Information Quality. Wit Press/Computational Mechanics.

Kunimatsu, T., Hirai, C., & Tomii, N. (2012). Train timetable evaluation from the viewpoint of passengers by microsimulation of train operation and passenger flow. Electrical Engineering in Japan, 181(4), 51-62.

Lamorgese, L., & Mannino, C. (2013). The track formulation for the train dispatching problem. Electronic Notes in Discrete Mathematics 41, 559–566.

Lan, L. W., & Lin, E. T. (2006). Performance measurement for railway transport: stochastic distance functions with inefficiency and ineffectiveness effects. Journal of Transport Economics and Policy, 383-408.

Lindner, T. (2011). Applicability of the analytical UIC Code 406 compression method for evaluating line and station capacity. Journal of Rail Transport Planning & Management, 1(1), 49-57.

Louwerse, I., & Huisman, D. (2014). Adjusting a railway timetable in case of partial or complete blockades. European Journal of Operational Research 235 (3), 583–593.

Mannino, C., & Mascis, A. (2009). Optimal real-time traffic control in metro stations. Operations Research 57 (4), 1026–1039.

Marinov, M., & Viegas, J. (2011). A mesoscopic simulation modelling methodology for analyzing and evaluating freight train operations in a rail network. Simulation Modelling Practice and Theory, 19(1), 516-539.

Mascis, A., & Pacciarelli, D. (2002). Job-shop scheduling with blocking and no-wait constraints. European Journal of Operational Research 143 (3), 498–517.

Mazzeo, A., Mazzocca, N., Nardone, R., D'Acierno, L., Montella, B., Punzo, V., Quaglietta, E., Lamberti, I., & Marmo, P. (2011). An integrated approach for availability and QoS evaluation in railway systems. Lecture Notes in Computer Science, 6894, 171-184.

Medeiros, G. (2008). RAMS analysis of railway track infrastructure. Dissertation for the qualification in the Msc. degree in 'Mestrado Integrado em Engenharia Civil'. Instituto Superior Técnico, Universidade de Lisboa, Portugal.

Middelkoop, D. & Bouwman, M. (2001). SIMONE: Large scale train network simulations. Proceedings of the 2001 Winter Simulation Conference, Institute of Electrical and Electronic Engineers, Piscataway, New Jersey, 1042-1047.

MVA Consultancy. (2008).Valuation of overcrowding on rail services. Prepared for Department for Transport.

Nakamura, T., Hirai, C., & Nishioka, Y. (2011). A practical train rescheduling algorithm using three predetermined factors. In: Proceedings of the 4th International Seminar on Railway Operations Modelling and Analysis – RailRome 2011, Rome, Italy, February 2011.

Nash, A., & Huerlimann, D. (2004). Railroad Simulation Using Open-Track. Computers in Railways IX, WIT Press, Southampton, UK.

Nash, A., Weidmann, U., Bollinger, S., Luethi, M., & Buchmueller, S. (2006). Increasing schedule reliability on the S-Bahn in Zurich, Switzerland: Computer analysis and simulation. In Transportation Research Record: Journal of the Transportation Research Board, No. 1955, Transportation Research Board of the National Academies, Washington, D.C, 17-25.

Pachl, J.. (2009). Railway Operation and Control. 2nd edition, VTD Rail Publishing, Mountlake Terrace (USA).

Pellegrini, P., Marlière, G., & Rodriguez, J. (2012). Real time railway traffic management modeling track-circuits. In: Delling, D., Liberti, L. (Eds.), Proceedings of the 12th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems, volume 25 of OpenAccess Series in Informatics (OASIcs), Dagstuhl, Germany. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 23-34.

Placido, A., Botte, M., D’Acierno, L., & Montella., B. (2015a). Effects of stochasticity on recovery solutions in the case of high density rail/metro networks. Rail Tokyo 2015. Tsudanuma Campus, Chiba Institute of Technology, Japan, March 23-26, 2015.

Placido, A., Cadarso, L., & D’Acierno, L. (2014a). Benefits of a combined micro-macro approach for managing rail systems in case of disruptions. Procedia: Social & Behavioral Sciences, 3, 195-204.

Placido, A., D’Acierno, L., Botte, M., Campora, S., & Montella, B. (2014b). The influence of dwell times on rail service planning. Proceedings of the 20th

Conference of the International Federation of Operational Societies (IFORS), p. 151, Barcelona, Spain.

Placido, A., D’Acierno, L., Botte, M., Gallo, M., & Montella., B. (2015b). A Sensitivity Analysis of Recovery Solutions in the Case of Rail Disruption Management. Proceedings of the Transportation Research Board – 94th meeting. Washington D.C., USA.

Placido, A., D’Acierno, L., Campora, S., & Botte, M. (2015). The dynamic interaction between passenger flows and rail service. Submitted to Transportation Research Part C: Emerging Technologies.

Placido, A., De Martinis, V., Montella, B., Gallo, M., & D’Acierno, L. (2014c). Effects of travel demand levels on optimal strategies for metro system management in failure contexts. Procedia: Social & Behavioral Sciences, vol. 111, 819-828.

Quaglietta, E. (2011). A Microscopic Simulation Model for Supporting the Design of Railway Systems: Development and Applications. PhD Thesis, University of Naples Federico II, Naples.

Quaglietta, E. Corman, F., & Goverde, R.M.P. (2013). Impact of a stochastic and dynamic setting on the stability of railway dispatching solutions. Proceedings of the International IEEE Conference on Intelligent Transportation Systems. The Hague, The Netherlands.

Quaglietta, E., & Punzo, V. (2013). Supporting the design of railway systems by means of a Sobol variance-based sensitivity analysis. Transportation Research Part C: Emerging Technologies, 34, 38-54.

Quaglietta, E., D’Acierno, L., Punzo, V., Nardone, R. & Mazzocca, N. (2011a). A simulation framework for supporting design and real-time decisional phases in railway systems. Proceedings of the 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), Washington (D.C.), USA, art. no. 6082913, 846-851.

Quaglietta, E., Punzo, V., Montella, B., Nardone, R., & Mazzocca, N. (2011b). Towards a hybrid mesoscopic-microscopic railway simulation model. 2nd International Conference on Models and Technologies for Intelligent Transportation Systems, Leuven, Belgium.

Radtke A. (2005) EDV-Verfahren zur Modellierung des Eisenbehnbetriebs (Software Tools to Model the Railway Operation), Habilitation am Institut fur Verkehrswesen, Eisenbahnbau und Betrieb (Schriftenreihe No. 64), Universitat Hannover, EurailPress.

Radtke, A. & Bendfeldt, J. (2001). Handling of railway operation problems with RailSys. Proceedings of the 5th World Congress on Rail Research, Cologne, Germany.

Radtke, A. & Watson R. (2007). Railway simulation in the United Kingdom, RTR No.1, 24-28.

Radtke, A. (2005). EDV-Verfahren zur Modellierung des Eisenbehnbetriebs (Software Tools to Model the Railway Operation). Habilitation am Institut fur Verkehrswesen, Eisenbahnbau und Betrieb (Schriftenreihe No. 64), Universitat Hannover, EurailPress.

Sauthoff, F. (1932). Die Bewegungswiderstaende der Eisenbahnwagen unter besonderer Berucksichtigung der neueren Versiche der Deutschen Reichsbahn, Ph.D. Thesis, Berlin, technical University (Technische Kochschule).

Schachtebeck, M., & Schöbel, A. (2010). To wait or not to wait and who goes first? Delay management with priority decisions. Transportation Science 44, 307-321.

Schlechte, T., Borndörfer, R., Erol, B., Graffagnino, T., & Swarat, E. (2011). Micro–macro transformation of railway networks. Journal of Rail Transport Planning & Management, 1(1), 38-48.

Schöbel, A. (2007). Integer programming approaches for solving the delay management problem. In: Geraets, F., Kroon, L.G., Schöbel, A., Wagner, D., Zaroliagis, C. (Eds.), Robust and Online Large-Scale Optimization, LNCS, vol. 4359. Springer, 145-170.

Schöbel, A. (2009). Capacity constraints in delay management. Public Transport 1 (2), 135-154.

Schumacher A. (2004) Ein Hybrides Verfahren zur Umlaufplanung von Fahrezeugen des Spurgefuhrten Verkehrs (A Hybrid Method for Vehicle Allocation in Rail Traffic), Dissertation am Institut fur Verkehrswesen, Eisenbahnbau und Betrieb (Schriftenreihe No. 60), Universitat Hannover, EurailPress.

Sewcyk, B. & Kettner, M. (2001). “Network Evaluation Model NEMO”. Proceedings of the World Congress on Railway Research.

Sewcyk, B. (2004). Makroskopische Abbildung des Eisenbahnbetriebs in Modellen Zur Langfristigen Infrastrukturplanung (Macroscopic display of Railway Operation in Models for Long-Term Infrastructure Planning). Dissertation am Institut fur Verkehrswesen, Eisenbahnbau und Betrieb (Schriftenreihe No. 61), Universitat Hannover, EurailPress.

Sewcyk, B., Radtke, A., & Wilfinger G. (2007). Combining Microscopic and Macroscopic Infrastructure Planning Models, Proceedings IAROR, Hannover.

Sheffi, Y., & Powell W.B. (1981). A comparison of stochastic and deterministic traffic assignment over congested networks. Transportation Research B, 15, 53-64.

Sheffi, Y., & Powell W.B. (1982). An algorithm for the equilibrium assignment problem with random link times. Networks, 12, 191-207.

Shimizu, H., Tanabe, H., & Yamamoto, M. (2008). The proposal system for Shinkansen using constraint programming. In: Proceedings of the 8th World Congress of Railway Research.

Shultze, K.. (1985). Modell Für Die Asynchrone Simulation Des Betriebes. In Teilen Des Eisenbahnnetzes Dissertation, Veröffentlichungen des Verkehrswissenschaftlichen Institutes der RWTH Aachen, Vol. 38.

Smith, A. S. (2012). The application of stochastic frontier panel models in economic regulation: Experience from the European rail sector. Transportation Research Part E: Logistics and Transportation Review, 48(2), 503-515.

Strahl (1913). Die Berechnung der FAhrzeiten und Geschwindigkeiten von Eisenbahnzugen aus den Belastungsgrenzen der Lokomotiven, Glasers Annalen fur Gewerbe- und Bauwesen 73, 869-871.

the 5th World Congress on Rail Research, Cologne, Germany.

Theeg, G. & Vlasenko, S. (2009). Railway Signalling & Interlocking - International Compendium. Eurailpress.

Toriumi, S., Nakamura, Y., & Taguchi, A. (2005). Delay calculation model of commuter trains (in Japanese). Communications of Operations Research of Japan 50 (6), 409-416.

Törnquist, J. (2012). Design of an effective algorithm for fast response to the rescheduling of railway traffic during disturbances. Transportation Research Part C, 20, 62-78.

Törnquist, J., & Persson, J.A. (2007). N-tracked railway traffic re-scheduling during disturbances. Transportation Research Part B 41 (3), 342-362.

Trozzi, V., Gentile G., Bell M.G.H., & Kaparias, I. (2013). Dynamic User Equilibrium in public transport networks with passenger congestion and hyperpaths. Transportation Research part B, 57, 266-285.

UIC. (2004). Code 406 – Capacity. 1st Edition.

Wardman, M. (2004). Public transport values of time. Transport Policy, 11(4), 363-377.

Wardman, M., & Whelan, G. (2011). Twenty years of rail crowding valuation studies: evidence and lessons from British experience. Transport reviews, 31(3), 379-398.

Wiklund, M. (2007). Serious breakdowns in the track infrastructure – calculation of the effects on rail traffic. In: Proceedings of the 2nd International Seminar on Railway Operations Modelling and Analysis – RailHannover 2007, Hannover, Germany.

Winter, P. (2009). Compendium on ERTMS - European Rail Traffic Management System. Hamburg: Eurailpress.

Yuan, J. (2002). Train dwell times in stations. P.H.L. Bovy, Trailblazing into the future! Proceedings of the 7th TRAIL congress, Delft University Press, Delft, The Netherlands.

Legislative references:

Cenelec (1999). Railway applications - Specification and demonstration of Reliability, Availability, Maintainability and Safety (RAMS). EN50126

Council Directive 95/18/EC of 19 June 1995 on the licensing of railway undertakings

Council directive of 29 July 1991 on the development of the Community's railways.

D.Lgs.422/97: Decreto Legislativo n.422 del 19 novembre 1997 "Conferimento alle regioni ed agli enti locali di funzioni e compiti in materia di trasporto pubblico locale, a norma dell'articolo 4, comma 4, della legge 15 marzo 1997, n. 59".

Decreto del presidente del consiglio dei ministri 30 dicembre 1998 "Schema generale di riferimento per la predisposizione della carta dei servizi pubblici del Settore Trasporti"

Decreto Legislativo 10 agosto 2007, n. 162 “Attuazione delle direttive 2004/49/CE e 2004/51/CE relative alla sicurezza e allo sviluppo delle ferrovie comunitarie”.

Decreto Legislativo 10 agosto 2007, n. 163 “Attuazione della direttiva 2004/50/CE che modifica le direttive 96/48/CE e 2001/16/CE relative all'interoperabilità del sistema ferroviario transeuropeo”.

Decreto legislativo 25 gennaio 2010, n. 15 “Attuazione della direttiva 2007/58/CE, che modifica le direttive 91/440/CEE, relativa allo sviluppo delle Ferrovie Comunitarie, e 2001/14/CE relativa alla ripartizione della capacità di infrastruttura ferroviaria e all'imposizione dei diritti per l'utilizzo dell'infrastruttura ferroviaria”.

Decreto legislativo 30 dicembre 2010, n. 247 “Attuazione della direttiva 2007/59/CE relativa alla certificazione dei macchinisti addetti alla guida di locomotori e treni sul sistema ferroviario della Comunità.”

Decreto Legislativo 8 luglio 2003, n. 188 “Attuazione delle direttive 2001/12/CE, 2001/13/CE e 2001/14/CE in materia ferroviaria”.

Decreto Legislativo 8 ottobre 2010, n.191 “Attuazione della direttiva 2008/57/CE e 2009/131/CE relativa all'interoperabilità del sistema ferroviario comunitario”.

Directive 2001/12/EC of the European Parliament and of the Council of 26 February 2001 amending Council

Directive 2001/12/EC of the European Parliament and of the Council of 26 February 2001 amending Council Directive 91/440/EEC on the development of the Community's railways.

Directive 2001/13/EC of the European Parliament and of the Council of 26 February 2001 amending Council Directive 95/18/EC on the licensing of railway undertakings

Directive 2001/14/EC of the European Parliament and of the Council of 26 February 2001 on the allocation of railway infrastructure capacity and the levying of charges for the use of railway infrastructure and safety certification

Directive 2001/16/EC of the European Parliament and of the Council of 19 march 2001 on the interoperability of the trans-European conventional rail system

Directive 2004/49/EC of the European Parliament and of the Council of 29 April 2004 on safety on the Community's railways

Directive 2004/49/EC of the European Parliament and of the Council of 29 April 2004 on safety on the community's railways and amending council directive 95/18/EC on the licensing of railway undertakings and directive 2001/14/EC on the allocation of railway infrastructure capacity and the levying of charges for the use of railway infrastructure and safety certification.

Directive 2004/50/EC of the European Parliament and of the Council of 29 April 2004 amending Council Directive 96/48/EC on the interoperability of the trans-European high-speed rail system and Directive 2001/16/EC of the European Parliament and of the Council on the interoperability of the trans-European conventional rail system.

Directive 2004/51/EC of the European Parliament and of the Council of 29 April 2004 amending Council Directive 91/440/EEC on the development of the Community's railways

Directive 2004/51/EC of the European Parliament and of the Council of 29 April 2004 amending Council Directive 91/440/EEC on the development of the Community's railways.

Directive 2007/58/EC of the European Parliament and of the Council of 23 October 2007 amending Council Directive 91/440/EEC on the development of the Community's railways and Directive 2001/14/EC on the allocation of railway infrastructure capacity and the levying of charges for the use of railway infrastructure.

Directive 2007/59/EC of the European Parliament and of the Council of 23 October 2007 on the certification of train drivers operating locomotives and trains on the railway system in the Community.

EN13816: European Standard EN13816 “Transportation – Logistic services – Public passenger transport – Service quality definition, targeting and measurement”.

IEC (2001). Functional safety of electrical/electronic/programmable electronic safety-related systems – Part 4: Definitions and abbreviations. EN 61508.

Regulation (EC) No 1371/2007 of the European Parliament and of the Council of 23 October 2007 on rail passengers’ rights and obligations.

Regulation (EC) No 881/2004 of the European Parliament and of the Council of 29 April 2004 establishing a European railway agency.

Regulation (EEC) no 1191/69 of the Council of 26 June 1969 on action by Member States concerning the obligations inherent in the concept of a public service in transport by rail, road and inland waterway.

Regulation (EEC) No 1893/91 of 20 June 1991 amending Regulation (EEC) No 1191/69 on action by Member States concerning the obligations inherent in the concept of a public service in transport by rail, road and inland waterway.