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 CIVILI – XXX CICLO

Modelling of interactions between rail service and travel demand: A passenger-oriented analysis

Marilisa Botte

Coordinatori di dottorato/Ph.D. course coordinators:

Prof. Andrea Papola

Prof. Elvira Petroncelli

Tutor/Supervisor:

Prof. Luca D'Acierno

CONTENTS

Chapter 1: Introduction ... 1

1.1 Research relevance and contributions ... 1

1.2 Thesis outline ... 5

Chapter 2: Literary review ... 7

2.1 Simulation models for transport systems ... 7

2.1.1 Delimitation of the study area ... 8

2.1.2 Zoning ... 8

2.1.3 Relevant infrastructures and services ... 10

2.1.4 The supply model ... 11

2.1.5 The demand model ... 18

2.1.6 The assignment models ... 23

2.2 Rail simulation models ... 29

2.3 Estimation techniques for travel demand flows ... 42

2.4 Simulation algorithms ... 51

2.5 Optimisation models for transport systems ... 61

2.6 Rail optimisation models ... 63

2.6.1. The timetabling phase ... 64

2.6.1.1 Dwell times estimation techniques ... 68

2.6.2. The rescheduling problem ... 73

2.7 Optimisation algorithms ... 83

2.8 Energy issues related to rail systems ... 95

Chapter 3: Proposed methodology for managing rail systems both in ordinary and disruption conditions ... 104

3.1 Optimisation framework ... 106

3.2 Basic simulation architecture 109

3.3 Extended simulation architecture 111

3.3.1 Stochastic simulation framework 112

3.3.2 Decision support system for implementing energy saving strategies 113

3.3.3 Modelling of the snowball effect 121

3.3.4 Travel demand estimation 131

3.3.4.1 Analytical methodology for extending passenger counts 133

3.3.4.2 A long-term evaluation of travel demand.............................. 140

3.4 Concluding remarks 152

Chapter 4: Applications to real network contexts of the proposed approach.. 155

4.1 Case studies 155

4.2 Rescheduling applications.. 162

4.2.1 Evaluation of unconventional rescue strategies for managing disruption conditions 178

4.3 Energy saving policies applications 183

4.4 Planning tasks: estimation of dwell times as flow-dependent factors.. 189

4.4.1 A comparison between FIFO and RIFO queuing rules 202

4.5 Travel demand estimation applications... 210

4.5.1 Calibration and validation of space-time relations representative of passenger flow data 210

4.5.2 A cost-benefit analysis relative to the implementation of an innovative signalling system in a regional context 216

Chapter 5: Conclusion 227

5.1 Resume of the main achievements 227

5.2 Research prospects 229

Dissemination of the main research achievements.................................... 232

References 234

CHAPTER 1: INTRODUCTION

1.1 Research relevance and contributions

The proposed research is situated in the field of design, management and optimisation in railway network operations. Rail transport has in its favour several specific features which make it a key factor in public transport management, above all in high-density contexts. Indeed, such a system is environmentally friendly (reduced pollutant emissions), high-performing (high travel speeds and low values of headways), competitive (low unitary costs per seat-km or carried passenger-km) and presents a high degree of adaptability to intermodality. However, it manifests high vulnerability in the case of breakdowns. This occurs because a faulty convoy cannot be easily overtaken and, sometimes, cannot be easily removed from the line, especially in the case of isolated systems (i.e. systems which are not integrated into an effective network) or when a breakdown occurs on open tracks. Thus, re-establishing ordinary operational conditions may require excessive amounts of time and, as a consequence, an inevitable increase in inconvenience (user generalised cost) for passengers, who might decide to abandon the system or, if already on board, to exclude the railway system from their choice set for the future. It follows that developing appropriate techniques and decision support tools for optimising rail system management, both in ordinary and disruption conditions, would consent a clear influence of the modal split in favour of public transport and, therefore, encourage an important reduction in the externalities caused by the use of private transport, such as air and noise pollution, traffic congestion and accidents, bringing clear benefits to the quality of life for both transport users and non-users (i.e. individuals who are not system users).

Managing to model such a complex context, based on numerous interactions among the various components (i.e. infrastructure, signalling system, rolling stock and timetables) is no mean feat. Moreover, in many cases, a fundamental element, which is the inclusion of the modelling of travel demand features in the simulation of railway operations, is neglected. Railway transport, just as any other transport system, is not finalised to itself, but its task is to move people or goods around, and, therefore, a realistic and accurate cost-benefit analysis cannot ignore involved flows features. In particular, considering travel demand into the analysis framework presents a two-sided effect.

Primarily, it leads to introduce elements such as convoy capacity constraints and the assessment of dwell times as flow-dependent factors which make the simulation as close as possible to the reality. Specifically, the former allows to take into account the eventuality that not all passengers can board the first arriving train, but only a part of them, due to overcrowded conditions, with a consequent increase in waiting times. Due consideration of this factor is fundamental because, if it were to be repeated, it would make a further contribution to passengers' discontent. While, as regards the estimate of dwell times on the basis of flows, it becomes fundamental in the planning phase. In fact, estimating dwell times as fixed values, ideally equal for all runs and all stations, can induce differences between actual and planned operations, with a subsequent deterioration in system performance. Thus, neglecting these aspects, above all in crowded contexts, would render the simulation distorted, both in terms of costs and benefits.

The second aspect, on the other hand, concerns the correct assessment of effects of the strategies put in place, both in planning phases (strategic decisions such as the realisation of a new infrastructure, the improvement of the current signalling system or the purchasing of new rolling stock) and in operational phases (operational decisions such as the definition of intervention strategies for addressing disruption conditions). In fact, in the management of failures, to date, there are operational procedures which are based on hypothetical times for re-establishing ordinary conditions, estimated by the train driver or by the staff of the operation centre, who, generally, tend to minimise the impact exclusively from the company's point of view (minimisation of operational costs), rather than from the standpoint of passengers. Additionally, in the definition of intervention strategies, passenger flow and its variation in time (different temporal intervals) and space (different points in the railway network) are rarely considered. It appears obvious, therefore, how the proposed re-examination of the dispatching and rescheduling tasks in a passenger-orientated perspective, should be accompanied by the development of estimation and forecasting techniques for travel demand, aimed at correctly taking into account the peculiarities of the railway system; as well as by the generation of *ad-hoc* tools designed to simulate the behaviour of passengers in the various phases of the trip (turnstile access, transfer from the turnstiles to the platform, waiting on platform, boarding and alighting process, etc.).

The latest workstream in this present study concerns the analysis of the energy problems associated to rail transport. This is closely linked to what has so far been described. Indeed, in order to implement proper energy saving policies, it is, above all, necessary to obtain a reliable estimate of the involved operational times (recovery times, inversion times, buffer times, etc.). Moreover, as the adoption of eco-driving strategies generates an increase in passenger travel times, with everything that this involves, it is important to investigate the trade-off between energy efficiency and increase in user generalised costs.

Within this framework, the present study aims at providing a DSS (Decision Support System) for all phases of planning and management of rail transport systems, from that of timetabling to dispatching and rescheduling, also considering space-time travel demand variability as well as the definition of suitable energy-saving policies, by adopting a passenger-orientated perspective.

Therefore, the provided contributions can be outlined as follows.

- Creating a dynamic database representing a decision-making tool for assisting dispatchers in handling both ordinary and disruption conditions. In particular, for each possible intervention strategy, related or not to a specific failure event, such database provides the identification and the quantification of relevant impacts on each part of the analysed system. In this way, dispatchers can be fully aware of the consequences of their own decisions and, thus, face the perturbed conditions in an appropriate manner, never opting again for the *non-intervention strategy*; moreover, response times can be made comparable with real-time rescheduling approaches, without, however, the computational effort they require.

- Developing an analytical framework which allows an accurate estimation of operational times within timetable as a support tool for the implementation of eco-driving strategies. Indeed, such policies imply an increase in travel times and, therefore, result feasible exclusively in the event of extra time rates available, which have to be suitably designed during the timetabling process.
- Defining a simulation-based methodology for computing dwell times as flow-dependent factors, rather than as fixed values. This task is fundamental in order to design a robust timetable, with a high degree of resilience to delays, and grows in importance in overcrowded contexts. Indeed, the dynamic interaction between rail service and passengers flows, which occurs on the interface platform-train, gives rise to the so called *snowball effect*: the number of passengers on the platform influences the dwell times of trains at stations, which may cause delays; these, in turn, produce an increase in headways which generates more passenger flows on the platform providing a further extension of dwell times and, therefore, additional delays. In particular, two different boarding behavioural patterns (i.e. FIFO and RIFO) are modelled and compared in terms of effects on rail service and passenger satisfaction.
- Customising travel demand estimation and forecasting techniques proposed in the literature to the specific features of rail transport, related to the discontinuous fruition in space and time which it offers. The relevance of this lies in the fact that, each planning task, both in the case of short and long term policies, requires an estimation of involved passenger flows as input information.

1.2 Thesis outline

This section provides a brief foreword to each chapter of the presented work.

Chapter 2 is focused on the literary review of research fields of concern. Specifically, the comprehensive nature of the proposed approach gives rise to the necessity of investigating a wide range of operational issues related to planning and management tasks in rail transport. Therefore, after a general analysis of simulation and optimisation models adopted for transportation systems, a focus on such techniques in the case of rail systems is provided. Additionally, both simulation and optimisation algorithms are described. Moreover, the estimation and forecasting techniques for travel demand are evaluated, with the aim of adapting them to the peculiarities of rail systems. Finally, an analysis of the main issues related to the application of energy savings policies in the rail field is given, with a focus on the existing deep relationship between eco-driving strategies and operational parameters within the planned timetable.

Chapter 3 describes the developed decision support tool which is based on suitable simulation models, properly integrated into an optimisation layout. In particular, it is possible to define a basic simulation structure which is improved and made more accurate by means of the development of methodological frameworks enabling the modelling of crucial operational factors, such as stochasticity of rail operations, the interaction between rail service and travel demand as well as energy saving issues. The adopted perspective is passenger-centric which means that the goal is to improve service quality so as to drive the modal split towards systems based on railway technology which is sustainable and high-performing.

Chapter 4 aims at pointing out the effectiveness of the proposed methodology, by applying it to real network contexts. In particular, most of the presented applications are focused on metro systems which, generally, operate in high-density conditions and, frequently, have to address overcrowded situations. Therefore, in such circumstances, the necessity of properly modelling the interaction between rail service and travel demand, as well as the need for ensuring a certain service quality, grow in importance. The second case-study is represented by a regional rail line, with the aim of showing the capacity of the proposed approach of dealing with different network contexts. Clearly, the differences between the two analysed systems have been duly taken into account. Indeed, a metro service is affected by urban user flows, while a regional network has to deal with extra-urban (i.e. rural) trips; moreover, the former are frequency-based, while the latter operates according to specific departure/arrival times at each station dictated by the planned timetable.

Finally, concluding remarks and research prospects are provided in chapter 5.

CHAPTER 2: LITERARY REVIEW

The proposed framework for managing railway systems is characterised by a simulation-optimisation integrated approach and, therefore, in this chapter both simulation and optimisation models presented in the literature, with related resolutions methods, are investigated. In particular, after an analysis concerning transportation systems in general, a deepening of railway contexts is carried out. Moreover, given the crucial role played in this work by the travel demand, related estimation techniques are assessed with the aim of customising them to the railway case. Finally, environmental issues relative to railway systems are described with particular attention to energy saving strategies involving the design of eco-driving profiles and the adjustment of operational times within the planned timetable.

2.1 Simulation models for transport systems

Transport systems are made up of physical and organisational elements which interact with each other to produce transport opportunities and satisfy travel demand which, in turn, is the result of the interactions among the various social and economic activities localised in a specific area.

Mathematical models concerning transport systems aim at simulating the interaction between demand flows and supply performance, both for existing contexts (operational phase) and hypothetical ones (planning phase). Therefore, such models, and the different techniques which they make use of, are fundamental tools for the assessment and/or design of interventions concerning physical (e.g. a new railway line) and/or functional elements (e.g. a new railway timetable) of a transport system. According to the analysed context, the elements considered relevant to the problem are identified and these, together with their reciprocal interactions, make up the *analysis system*; while, the remaining is known as the *external environment* and it is taken into account solely due to its relationship with the analysis system. Therefore, the transport system in a certain area can be seen as a subset of a wider territorial system with which it strongly interacts. Schematically, as shown by Cascetta (2009), its modelling can be summarised according to the following phases:

1. delimitation of the *study area*;
2. *zoning*, consisting in the subdivision of the study area into traffic zones;
3. definition of the *base network*, consisting in the identification of relevant transport infrastructures and services;
4. development of the transport supply model;
5. development of the travel demand model;
6. implementation of simulation models reproducing the interaction between supply and demand features.

In particular, the first three phases listed are preliminary to the development of the entire demand and supply models, since they define the spatial delimitation of the analysed system and the level of disaggregation to which the following assessments will be referred.

2.1.1 Delimitation of the study area

The study area is the geographic area including the transportation system which is involved in the planned measures and the elements which are mostly affected by the project. Indeed, the analysis cannot be focused in the sole area where the planned measures will be carried out, but must concern a much wider area including the other elements which will inevitably be influenced by the effects of the modifications on the analysed transport system. Therefore, it has to be properly designed. The limit of the study area is usually indicated as the *area cordon* or *boundary*.

2.1.2 Zoning

Modelling users' trips requires the definition of the departure and the destination locations. Obviously, trips which occur in a given area can start and end in a very high number of points on the territory; therefore, for allowing their simulation, it is necessary to subdivide the study area into a finite number of geographic units known as *traffic zones*.

In particular, a traffic zone represents a portion of the territory with homogenous features in relation to the activities, the accessibility, the infrastructures and the transport services. Trips between two different zones are named *inter-zonal trips*, while *intra-zonal* trips are those starting and ending within the same traffic zone.

To each zone is associated a fictitious node, named *centroid,* which represents the actual starting and terminal point of trips beginning or ending in each traffic zone. In such representation, intra-zonal trips start and end in the same point and, therefore, they are not simulated on the network. Moreover, external centroids are positioned on the intersections between cordon and the considered infrastructures and services. Through them, the trips from and towards the external zones, entering, leaving and crossing the study area, are modelled.

Zoning clearly depends on the scale of the problem under study and, in particular, on the features of trips to be simulated. Thus, according to the diverse level of detail, a traffic zone may include a building, a set of buildings, neighbourhoods, towns, regions or a whole country. Of course, the denser the zoning, the more exact the representation of the real system, but, at the same time, the greater the computational complexity. However, there are some guiding principles for the identification of traffic zones which are listed below.

Firstly, it is worth pointing out that traffic zones are generally obtained as aggregations of administrative territorial units (e.g. census sections, municipalities or provinces), in order to be able to carry out the socio-economic statistical data related to each zone, such as population or employment. Moreover, physical geographic separators (rivers, railway line sections, etc.) are usually used as zone borders because they imply different accessibility conditions to the infrastructures and transport services. Finally, in the definition of the limits of the traffic zones, it is desirable to aggregate areas which perform in a homogeneous manner in terms of land-use (e.g. residential or commercial zones in urban areas or rural municipalities in extra-urban areas) and accessibility to transportation facilities.

2.1.3 Relevant infrastructures and services

Regarding phase 3, the definition of the base network (which can be monomodal or multimodal) occurs by selecting the relevant supply elements.

The relevant infrastructures and transport services are identified on the basis of their role in connecting the traffic zones in the study area and the external zones. This implies a mutual dependence between the zoning phase and the definition of the base network (figure 2.1). Basically, since intra-zonal trips are neglected, the supply elements related to journeys between two places belonging to the same traffic zone are omitted; on the contrary, supply elements which are needed to connect two places belonging to different zones are taken into account, clearly according to the addressed scale problem.

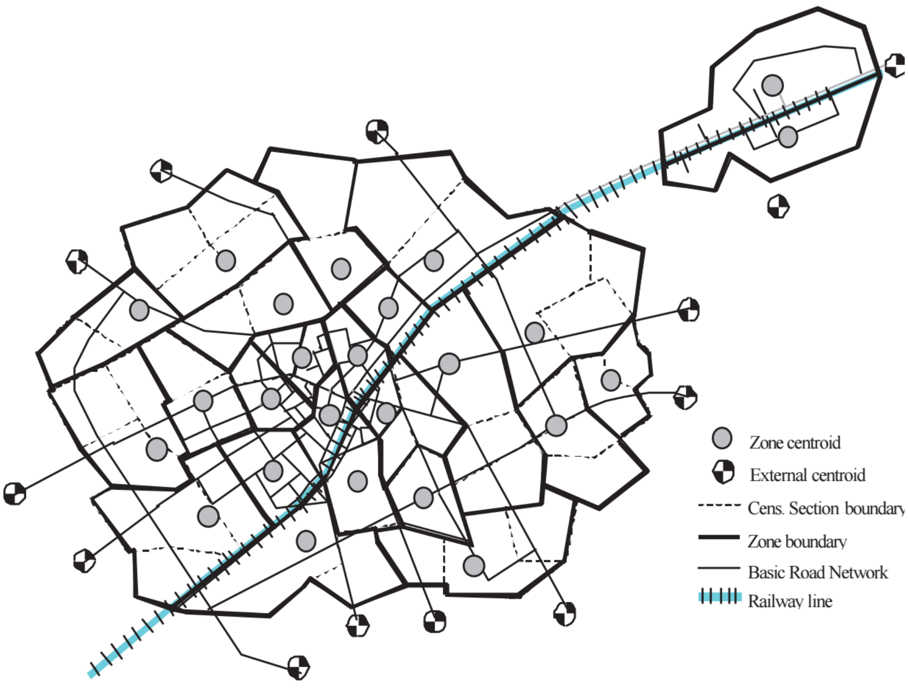


Figure 2.1 Zoning and base network (source: Cascetta, 2009)

After these preliminary phases, it is possible to build the supply and demand models, whose interaction is replicated by means of the assignment models.

2.1.4 The supply model

Regarding the supply representation, it is characterised by a topological and an analytical model.

The former is based on the use of the graph theory (Hauptmann, 2000; Newell, 1980; Potts and Oliver, 1972; Radtke and Watson, 2007). In particular, as shown by Hansen and Pachl (2008), a valued graph can be defined as follows:

$$G := (V, E, c) \tag{2.1}$$

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:

$$c(e) \geq 0 \quad \forall e \in E \tag{2.2}$$

Moreover, some definitions are provided: a graph is defined as '*directed*', if two adjacent nodes are linked by at least one connection and the direction is indicated by an arrow; '*simple*', if the graph does not contain parallel links or loops and '*connected*', if for any two nodes of the graph, links exist connecting the node.

Specifically, in a graph illustrating a transport network, the nodes correspond to noteworthy events and each link represents a phase of the trip; while, a succession of consecutive links connecting the origin and destination nodes identifies a path. Two fundamental quantities are associated to each link and path, which are flows and costs.

Obviously, the features of the graph and the quantities associated to each element vary according to the type of transportation system to be analysed. In particular, it is possible to distinguish:

- *continuous services*: available at every instant in time and accessible from any spatial point. Typical examples are individual modes which use road systems (cars, motorbikes, bicycles, pedestrian).
- *discontinuous or scheduled services*: available only in some instances in time and accessible only in given spatial points. Examples of such systems are scheduled services (bus, train, plane) which can be used only

between terminals (i.e. stops, stations, airports etc.) and are available only in certain instances, according to the planned timetable.

In the first case, nodes are localised at the intersections between road segments or in correspondence of particularly significant variations in the geometric and/or functional features of a single road segment, such as changes in road section or in terms of slope. Links, generally, correspond to the connections between nodes allowed by the road circulation rules. In this case, real links can represent, for instance, road segments, while fictitious links can simulate waiting phenomena at intersections or toll pay barriers.

Regarding cost items to be considered, it is worth pointing out that the term cost, generally used with the meaning of an exclusively monetary nature, in the economy of transport represents a linear combination of weights and attributes which defines the level of performance of the element to which it is associated or, in other words, the general impedance perceived by the users on a particular trip. It is known as *user generalised cost*.

Therefore, in the case of continuous services, typical rates of user generalised cost are running time, waiting time at intersections and monetary cost, generally corresponding to the fuel consumption and tolls, if any.

On the other hand, for discontinuous services, it is necessary to make a further distinction according to the system features and the assumptions about user choice behaviour.

In the case of services with a *high frequency* (e.g. one run every 5-15 minutes) and *low regularity*, it is usually assumed that user does not choose a specific run, but rather a service line (i.e. a set of runs characterized by the same terminals, same stations and same performance) or a group of lines.

Therefore, the implemented graph is the so-called *line graph*, depicted in figure 2.2, whose elements are:

- *access nodes,* representing the arrival of the user at the stop;
- stop *nodes* or *diversion nodes*, representing the boarding of a vehicle;

- *line nodes,* representing the arrival and departure of vehicles of a given line at a given stop;
- *access links* representing access trips between access nodes;
- *waiting links,* representing the waiting at the stop;
- *boarding and alighting links,* representing boarding and alighting process from the vehicles of a line;
- *on-board links,* representing the trip from one stop to another of the same line;
- *dwelling links,* representing vehicle dwelling at the stop.

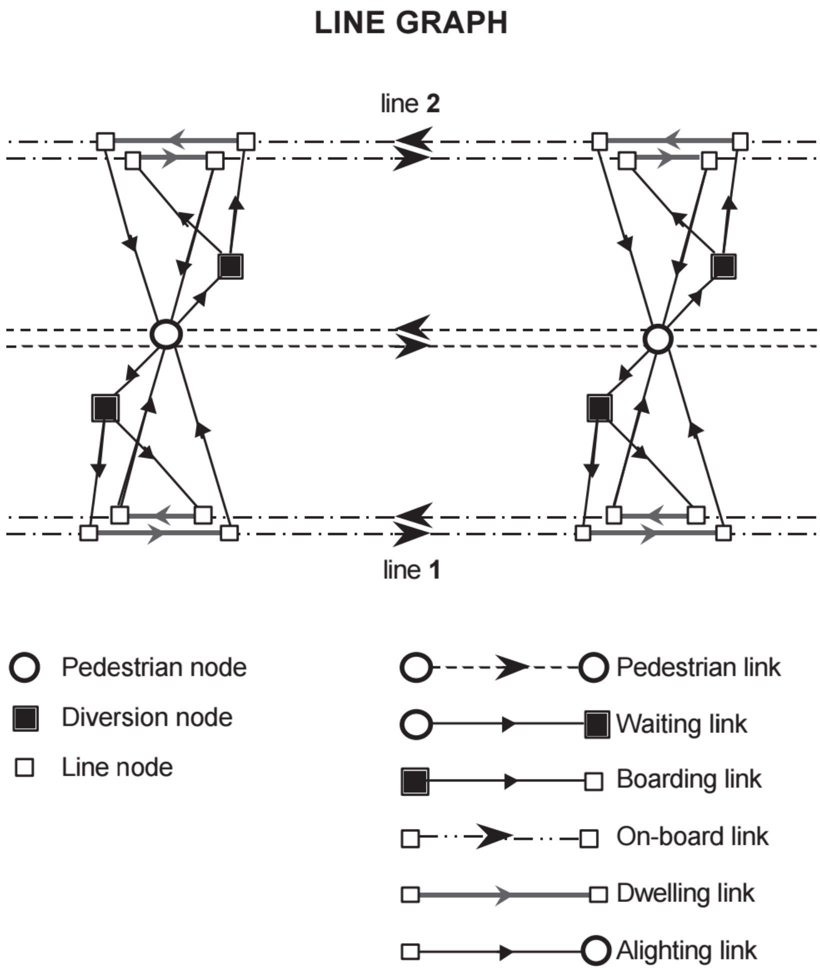


Figure 2.2 Line graph for urban transit systems (source: Cascetta, 2009)

In such a graph, in the case of railway transport, a real node can identify, for instance, a level crossing or a station; while, a fictitious node can represent a diversion point (i.e. where the user decides what train line to choose). Equally, a link can represent real elements, such as a rail section, or fictitious elements, such as the waiting time of a user on a platform.

Consequently, the aliquots of user cost to be taken into account are: on-board travel times, dwell times at stops, waiting times, boarding/alighting times and access/egress times which generally correspond to walking or driving time for reaching a point (both in spatial and temporal terms) where the service is available.

On the contrary, services with a *low frequency* and *high regularity* imply that the user chooses a specific run, with its own features. In this case, the assumption of within-day stationary does not hold, thus, it is necessary to deal with intra-period dynamic systems. Therefore, a more complex graph has to be adopted, known as *run graph* or *diachronic graph*, where, in order to simulate passenger behaviour in an appropriate manner, it is necessary to introduce additional information such as *desired departure and arrival times.*

In order to clarify the existing correlations between the involved variables, the following notations are introduced:

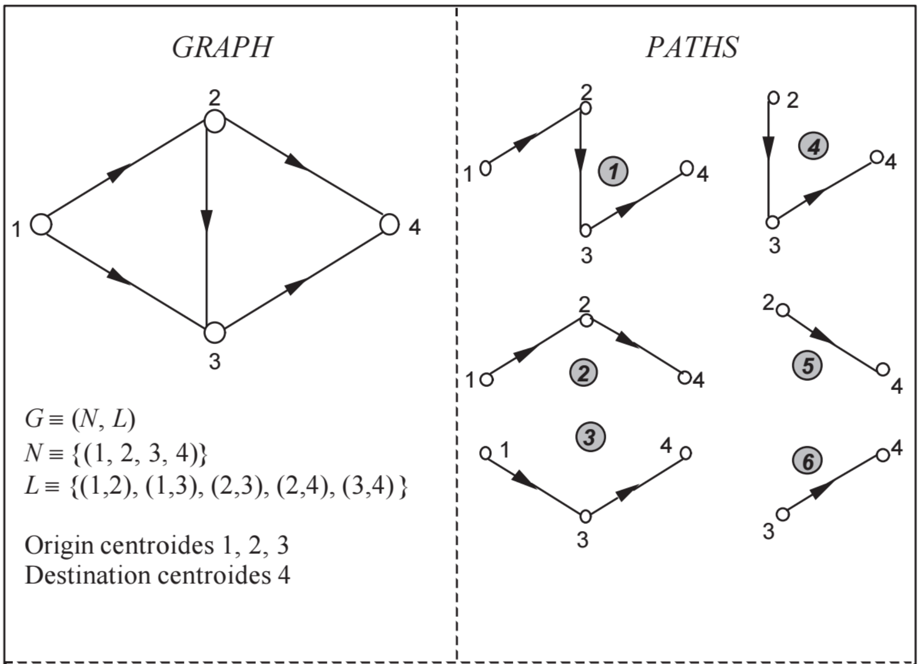
$\boldsymbol{f}$ is the link flow vector, whose dimension is $(n_L \times 1)$ with n_L number of links in the network and whose generic element is the flow f_l on link l.

$\boldsymbol{h}$ is the path flow vector, whose dimension is $(n_P \times 1)$ with n_P number of paths in the network and whose generic element is the flow h_k on path k.

$\boldsymbol{c}$ is the link cost vector, whose dimension is $(n_L \times 1)$ with n_L number of links in the network and whose generic element is the cost c_l on link l.

$\boldsymbol{g}$ is the path cost vector, whose dimension is $(n_P \times 1)$ with n_P number of paths in the network and whose generic element is the cost g_k on path k.

In particular, costs related to a path can be additive or non-additive: an additive path cost can be obtained by summing the costs related to the links included in the path (e.g. on-board time); on the contrary, a non-additive path cost is independent on cost values of each link (e.g. waiting time at stops for high frequency transit systems). For the sake of simplicity, the non-additive path costs will be neglected in the following discussion.



Link flow vector *Path flow vector* *Link-path incidence matrix*

$$\boldsymbol{f} = \begin{bmatrix} f_{12} \\ f_{13} \\ f_{23} \\ f_{24} \\ f_{34} \end{bmatrix} \qquad \boldsymbol{h} = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \end{bmatrix} \qquad \Delta = \begin{array}{c|cccccc} & 1 & 2 & 3 & 4 & 5 & 6 \\ \hline 1,2 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1,3 & 0 & 0 & 1 & 0 & 0 & 0 \\ 2,3 & 1 & 0 & 0 & 1 & 0 & 0 \\ 2,4 & 0 & 1 & 0 & 0 & 1 & 0 \\ 3,4 & 1 & 0 & 1 & 1 & 0 & 1 \end{array}$$

Link flow vector

$$\boldsymbol{f} = \begin{bmatrix} f_{12} \\ f_{13} \\ f_{23} \\ f_{24} \\ f_{34} \end{bmatrix} = \boldsymbol{\Delta h} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \\ h_6 \end{bmatrix} = \begin{bmatrix} h_1 + h_2 \\ h_3 \\ h_1 + h_4 \\ h_2 + h_5 \\ h_1 + h_3 + h_4 + h_6 \end{bmatrix}$$

Figure 2.3 Link and path flows (source: Cascetta, 2009)

The analytical modelling of the transportation supply system is based on the set of equations described below.

Firstly, the *network loading or (static) flow propagation model* defines the relationship among link and path flows (see figure 2.3):

$$\boldsymbol{f} = \boldsymbol{\Delta} \cdot \boldsymbol{h} \tag{2.3}$$

where $\boldsymbol{\Delta}$ represents the *link-path incidence matrix* whose dimension is $(n_L \times n_P)$. It is a binary matrix in which the generic element δ_{lk} is equal to 1, if the link l belongs to path k ($l \in k$), and 0 otherwise ($l \notin k$).

Then, we have the *link performance model*, expressed by means of the so called *cost functions*, which allow to take into account the phenomenon of congestion, according to which the performance of a link depends on the involved flows.

$$\boldsymbol{c} = \boldsymbol{c}(\boldsymbol{f}) \tag{2.4}$$

Examples of performance attributes affected by congestion are road travel times or alighting/boarding times on platform in the case of a transit system. On the contrary, a parameter unaffected by the phenomenon of congestion is on-board time. Indeed, the time required for a train to travel between two stations is independent of the number of passengers which are on-board.

Moreover, if a link cost is influenced exclusively by the flow on the link itself, the related function cost is defined as *separable*:

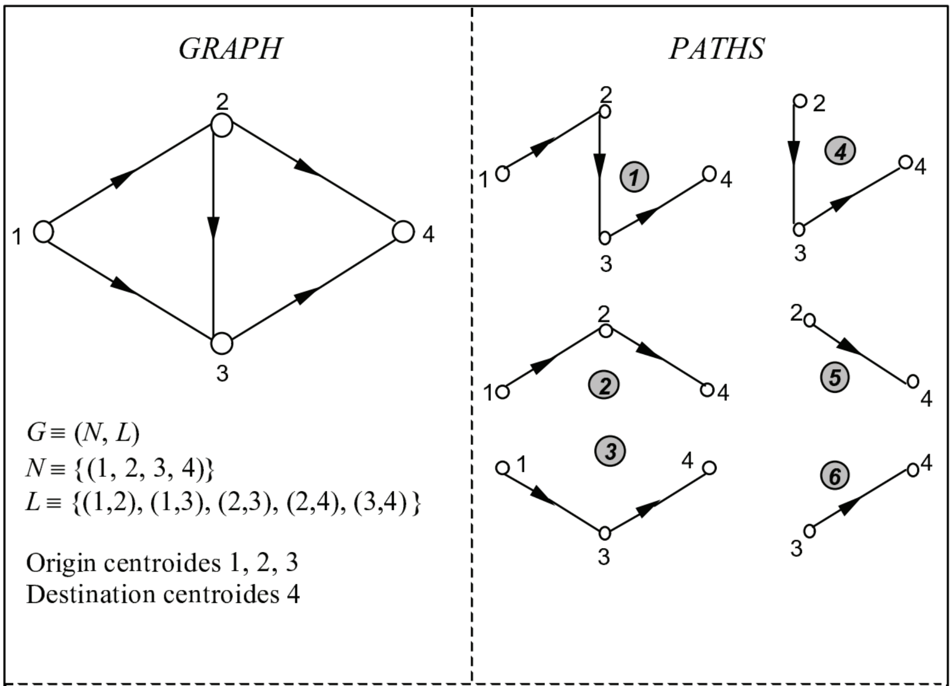
$$c_l(\boldsymbol{f}) = c_l(f_l) \tag{2.5}$$

On the other hand, if the link cost is influenced by the flow on the link itself and also by flows on other links, the related cost function is defined as *non-separable*.

A separable cost function is that representing, for instance, the waiting time at signalised intersections; while, an example of non-separable cost function concerns the dwell time of a bus or a train at stops, since it depends on the flows on three different links: boarding, alighting and on-board flows.

In the literature, several cost functions have been proposed: for urban road links (Festa and Nuzzolo, 1989), for extra-urban road links (HCM, 2000), for modelling waiting times at signal controlled intersections (Webster, 1958;

Webster and Cobbe, 1966), for toll-barrier links (Kleinrock, 1975; Newell, 1971) and parking links (Bifulco 1993). An example of a cost function concerning transit system can be found in Bouzaiene-Ayari et al. (1998).



Link cost vector

$$\boldsymbol{c} = \begin{bmatrix} c_{12} \\ c_{13} \\ c_{23} \\ c_{24} \\ c_{34} \end{bmatrix}$$

Link-path incidence matrix

$\boldsymbol{\Delta}=$

	1	2	3	4	5	6
1,2	1	1	0	0	0	0
1,3	0	0	1	0	0	0
2,3	1	0	0	1	0	0
2,4	0	1	0	0	1	0
3,4	1	0	1	1	0	1

Path cost vector

$$\boldsymbol{g}^{ADD} = \boldsymbol{\Delta}^T \boldsymbol{c} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_{12} \\ c_{13} \\ c_{23} \\ c_{24} \\ c_{34} \end{bmatrix} = \begin{bmatrix} c_{12} + c_{23} + c_{34} \\ c_{12} + c_{24} \\ c_{13} + c_{34} \\ c_{23} + c_{34} \\ c_{24} \\ c_{34} \end{bmatrix}$$

Figure 2.4 Link and path costs (source: Cascetta, 2009)

Finally, the *path performance model* connects the performance of single elements (links) with those of whole paths between any origin-destination pair (see figure 2.4):

$$\boldsymbol{g} = \boldsymbol{\Delta}^{T} \boldsymbol{c} \tag{2.6}$$

By combining equations (2.3), (2.4) and (2.6), the following relation is obtained:

$$\boldsymbol{g} = \boldsymbol{\Delta}^{T} \boldsymbol{c}(\boldsymbol{\Delta} \cdot \boldsymbol{h}) \tag{2.7}$$

which represents the mathematical formulation of the supply model.

2.1.5 The demand model

As stated by Cascetta (2009), the transportation demand model can be defined as a mathematical relationship associating the average values of demand flows with their relevant characteristics to a given activity and transportation supply system. It can be expressed by means of the following analytical formulation:

$$d_{od}[K_1, K_2, ..., K_n] = d(\boldsymbol{SE}, \boldsymbol{T}, \boldsymbol{\beta}) \tag{2.8}$$

where d_{od} is the average travel demand flow between zones *o* and *d*; $K_1, K_2, ..., K_n$ are the relevant characteristics of d_{od} flow; $\boldsymbol{SE}$ is the vector of socio-economic variables; $\boldsymbol{T}$ is the vector of level-of-service attributes of the transportation supply system; $\boldsymbol{\beta}$ is the vector of parameters to be properly calibrated.

This formulation expresses the fact that travel demand arises from the necessity to move in order to perform activities in different places. The estimation techniques for demand flows will be described in detail in paragraph 2.3; while, the following will be restricted to giving some fundamental notions, regarding the analytical representation of travel demand, which are preliminary to the description of assignment models.

Both spatial and temporal dimensions of travel demand play an important role. The spatial characterization is based on a matrix type representation, depicted in figure 2.5, which is known as the *origin-destination matrix* (*O-D matrix*).

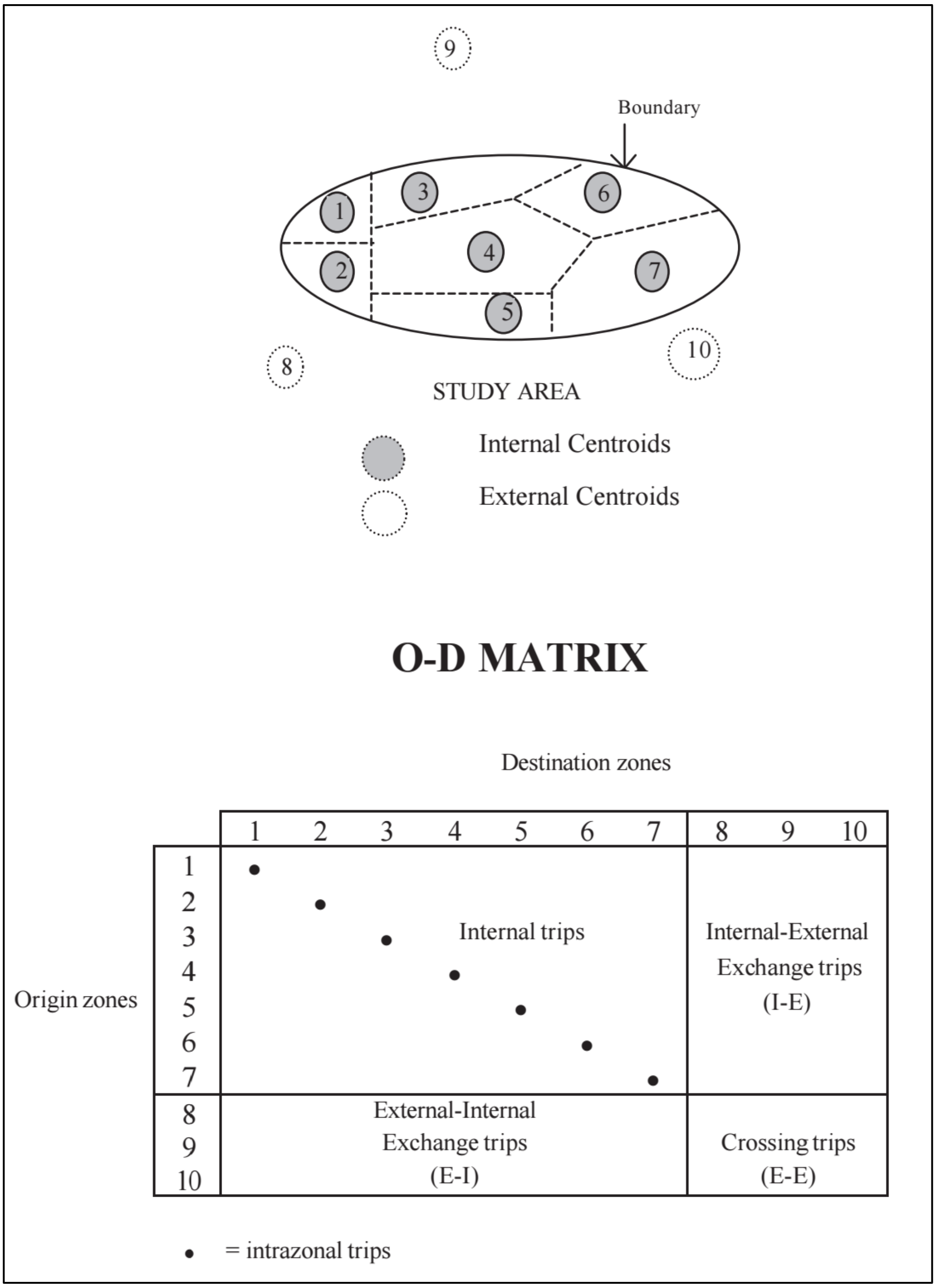


Figure 2.5 Trip types and their identification in the O-D matrix (source: Cascetta, 2009)

It is a square (*n* x *n*) matrix with *n* number of zones identified in the zoning phase. The generic element d_{od} provides the number of trips made in the reference period from origin zone *o* to destination zone *d* (i.e. *O-D flow*).

By summing the elements of the row *o*, it is possible to compute the flow emitted by zone *o*:

$$d_{o\cdot} = \sum_{d} d_{od} \tag{2.9}$$

By summing, instead, the elements of the column *d*, it is possible to compute the flow attracted by zone *d*:

$$d_{\cdot \mathrm{d}} = \sum_{o} d_{od} \tag{2.10}$$

Finally, by summing all the elements of the O-D matrix it is possible to obtain the total number of trips performed in the study area in the reference interval:

$$d_{\cdot\cdot} = \sum_{o} \sum_{d} \; d_{od} \tag{2.11}$$

Moreover, the O-D matrix can be subdivided into different sub-matrices, according to the type of involved trips:

1. *internal trips*: the origin and destination zones are both inside the study area; in particular, as already pointed out, the trips which start and end in the same zone are defined as intra-zonal and are neglected in the analysis.

2. *exchange trips*: the origin and destination zones are one within and the other without the study area or vice versa.

3. *crossing trips*: the origin and destination of trips are both external the study area.

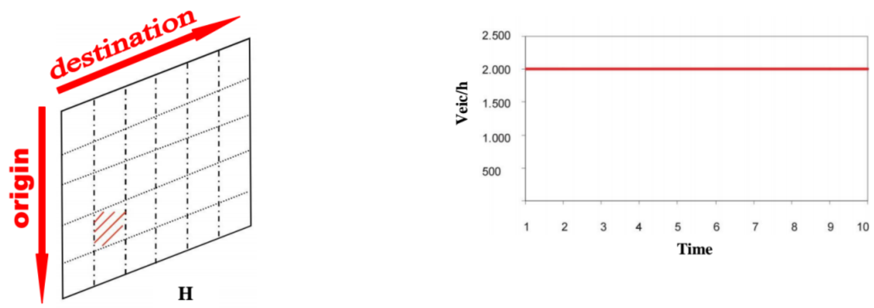
Furthermore, regarding the temporal dimension, two different approaches can be adopted, according to the assumption of *intra-period stationarity* or *intra-period dynamics* made.

Traditional planning models generally adopt the first hypothesis, according to which demand flows are constant for a sufficiently lengthy period of time to enable the analysed system to reach a steady-state condition (figure 2.6a). On the other hand, by adopting the intra-period dynamics approach, the time variability of travel demand is explicitly simulated by introducing as many

O-D matrices as the time intervals in which the analysed time period has been subdivided into (figure 2.6b).

Therefore, in this case, the generic entry $d^h{}_{od}$ represents the number of trips from origin *o* to destination *d* in the time period *h*.

a)



b)

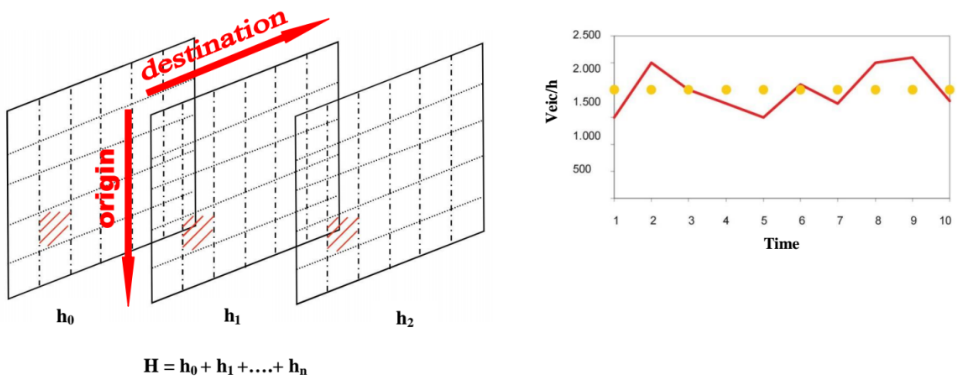


Figure 2.6 Static case (a) and dynamic case (b) in the O-D matrix representation (source: Nigro, 2009)

Cleary, this implies a major computational complexity, but allows to overcome several limits of the static simulation, which is inadequate to explicitly represent the capacity limits on the network and provide information on the propagation of congestion phenomena.

Moreover, as will be shown in paragraph 2.3, the intra-period dynamics assumptions have a key role in the adjustment process of the *O-D* matrix by means of aggregated data as traffic counts.

Besides the origin-destination pair (*od*) and the simulated period (*h*), other relevant features for modelling demand flows are: the purpose of the trip *s* (e.g. home-work trips, work-shopping trips), the mode adopted during the trip *m* (e.g. car, bus, bicycle) and the route used for the trip *k* (i.e. a series of links connecting the considered origin-destination pair). Moreover, if relevant for the analysis, also socio-economic features of users, such as income group or driving-license holding, can be taken into account. In this case, different homogeneous user groups (i.e. user categories) are identified and the generic user category is denoted by *i*.

Therefore, equation (2.8) becomes:

$$d^{i}{}_{od}\left[s,h,m,k\right]=d\left(\boldsymbol{SE,T}\right) \tag{2.12}$$

This formulation implies four different choice dimensions: if performing or not the trip (i.e. trip generation model), which destination to reach (i.e. trip distribution model), with which mode (i.e. modal split model) and through which route (i.e. route choice model). Generally, these choice dimensions are modelled by means of the *random utility theory* and the traditional approach for computing travel demand is based on the adoption of the so-called *partial share model*. For an exhaustive discussion concerning these models, the existing wide literature can be referred to (see, for instance, Ben-Akiva and Lerman, 1985; Cascetta, 2009; Domencich and McFadden,1975); while, below, the analytical formulation of demand model is described.

In particular, it can be outlined in aggregate form as:

$$\boldsymbol{h}=\boldsymbol{P}\left(-\boldsymbol{g}\right)\cdot\boldsymbol{d} \quad \in S_h \ \forall \ \boldsymbol{g} \tag{2.13}$$

where S_h is the set of feasible path flows and $\boldsymbol{P}$ is the *path choice probabilities matrix*, with a column for each *OD* pair and a row for each path *k*. The generic entry *k,od* is given by *p[k/od]* if path *k* connects the *OD* pair, otherwise it is null.

As will be described in the following, the characterisation of such a matrix, according to the path choice behaviour assumptions, plays a key role in the assignment models.

2.1.6 The assignment models

Assignment models reproduce supply-demand interactions and allow the evaluation of system performance by providing the basis for every technical assessment. These models can be classified according to the following assumptions:

1. dependence of link performance variables on flows: if link costs are independent of flows, the network is defined *uncongested*; otherwise, we have a *congested* network and an equilibrium problem where it is necessary to find configurations in which demand, path, and link flows are mutually consistent with the costs that they imply.
2. path choice behaviour: *deterministic* choice models assume that the perceived utility of a path is deterministic and all users choose a minimum cost alternative; while, *stochastic* choice models assume that the perceived utility of a path is a random variable.
3. dependence of O-D flows on path costs: if demand flows are independent of cost variations due to network congestion we have *rigid* demand models; otherwise, the demand models are defined as *elastic*.

By combining these assumptions it is possible to obtain different assignment models:

- Deterministic Uncongested Network (DUN)
- Stochastic Uncongested Network (SUN)
- Deterministic User Equilibrium (DUE)
- Stochastic User Equilibrium (SUE)

Their mathematical formulations are set out below.

In the case of *Deterministic Uncongested Network* assignment models, link costs do not depend on flows and the perceived utility of a path is assumed as deterministic. The mathematical formulation of this assignment model can be obtained by combining equations (2.3), (2.6) and (2.13), as follows:

$$\boldsymbol{f}_{DUN} = \boldsymbol{f}_{DUN}(\boldsymbol{c},\boldsymbol{d}) = \boldsymbol{\Delta} \cdot \boldsymbol{P}\left(-\boldsymbol{\Delta}^T \cdot \boldsymbol{c}\right) \cdot \boldsymbol{d} \qquad \in S_f(\boldsymbol{d}) \;\; \forall \;\; \boldsymbol{c} \tag{2.14}$$

where S_f is the set of feasible link flows, which can be obtained by the set S_h of the feasible path flows by means of equation (2.3).

Given the deterministic characterisation of the path choice model, the generic element of the link-path incidence matrix, $p(k/od)$, is expressed as follows:

$$p(k/od) = \begin{cases} \geq 0 & \text{if path } k \text{ is the minimum cost path} \\ = 0 & \text{otherwise} \end{cases}$$

with

$$\sum_{k \in K_{od}} p(k/od) = 1$$

On the other hand, in the case of *Stochastic Uncongested Network* assignment models, the assumption of the dependence of link costs on flows link is the same, but the path choice model is assumed as stochastic:

$$\boldsymbol{f}_{SUN} = \boldsymbol{f}_{SUN}(\boldsymbol{c},\boldsymbol{d}) = \boldsymbol{\Delta} \cdot \boldsymbol{P}\left(-\boldsymbol{\Delta}^T \cdot \boldsymbol{c}\right) \cdot \boldsymbol{d} \qquad \in S_f(\boldsymbol{d}) \;\; \forall \;\; \boldsymbol{c} \tag{2.15}$$

This implies that, nominally, the mathematical expression does not change as shown by equation (2.15); however, the link-path incidence matrix is no longer binary. Indeed, since the stochastic choice path model assumes that the perceived utility of a path is a random variable, the generic element has to be calculated by means of random utility models such as, for instance, the Multinomial Logit model (Ben-Akiva and Lerman, 1985; Domencich and McFadden, 1975) or the Probit model (Daganzo and Sheffi, 1977).

The *Multinomial Logit model* is one of the most widely used random utility models, since it has the great advantage of being able to be expressed in a closed form. However, its property known as *independence from irrelevant alternatives*, in some contexts, could provide an overestimation of choice probabilities, by leading to unrealistic results. This is due to the assumption on the independence of random residuals, according to which similar paths (i.e. paths made up of the same links in a relevant part) are perceived as distinct by the decision-maker.

A first extension to the Multinomial Logit model, which allows to partially overcome the assumption of independent random residuals, by preserving the closed form, is the *Nested Logit model* (Daly and Zachary, 1978; Williams, 1977). Specifically, it can consider different levels of interdependence among groups of alternatives (i.e. nests) in a choice set; the alternatives belonging to the same subset are intended as perceived in a similar way by the decision-maker and, therefore, they present a non-zero covariance among their random residuals. As a consequence, the variance-covariance matrix of random residuals has a block diagonal structure in which the covariance between each pair of alternatives belonging to the same group is constant, while the covariance between alternatives belonging to different groups is null. Another model developed for overcoming the drawback of the Multinomial Logit model, by maintaining the closed form, is the one proposed by Cascetta et al. (1996), known as *C-logit*. It is based on the introduction of a *commonality factor* which reduces the systematic utility of a path according to its degree of overlapping with other paths. Additionally, very recently, Papola et al. (2017) have proposed the so called *CoNL* route choice model, whose cdf is defined as a finite mixture of different Nested Logit cdfs. Thereby, such a model is characterised by closed-form expressions for choice probabilities, covariances and elasticities; moreover, it enables a very flexible correlation pattern.

By leaving the models belonging to the Logit class, it is worth mentioning the *Probit model* (Daganzo, 1979; Horowitz et al., 1982; Langdon, 1984) by means of which unbiased estimates of path choice probabilities, and of the corresponding path flows, can be obtained by using a Monte Carlo sampling technique of paths random residuals. Therefore, it does not present a closed-form formulation; however, in response to a more complex analytical tractability, offers a more realistic evaluation of choice probabilities and gets over most of the drawbacks of the Logit model and its generalizations.

By moving to the equilibrium models, it is necessary to take into account the circular dependence between flows and costs. In particular, the equilibrium configuration of the system is a condition whereby demand, path, and link flows are mutually consistent with the costs that they induce.

In the case of *Deterministic User Equilibrium* assignment models, link costs depend on flows and the perceived utility of a path is assumed as deterministic. The latter implies some mathematical complications due to the fact that the demand model is expressed by a one-to-many map and, therefore, path flows are not uniquely defined. For this reason, the properties of deterministic equilibrium are usually studied by means of indirect formulations called *variational inequality models* (Dafermos, 1980; Smith, 1979), expressed as follows:

$$\boldsymbol{c}(\boldsymbol{f}_{DUE})^{\mathrm{T}} \cdot (\boldsymbol{f} - \boldsymbol{f}_{DUE}) \geq 0 \quad \forall\, \boldsymbol{f} \in S_f(\boldsymbol{d}) \tag{2.16}$$

Sufficient conditions for the existence and uniqueness of deterministic flows $\boldsymbol{f}_{DUE}$ are assured respectively by the *continuity* and *monotonicity* of the cost functions.

In particular, the variational inequality problem (2.16) has at least one solution if the conditions of the Brouwer's theorem (Brouwer, 1912) are respected, i.e. cost functions are continuous and defined in the compact and convex non-empty set of link flows S_f.

The variational inequality problem (2.16) has at most one solution if cost functions are strictly monotone increasing functions in the set of feasible link flows S_f:

$$[\boldsymbol{c}(\boldsymbol{f}') - \boldsymbol{c}(\boldsymbol{f}'')]^{\mathrm{T}} (\boldsymbol{f}' - \boldsymbol{f}'') > 0 \quad \forall\, \boldsymbol{f}' \neq \boldsymbol{f}'' \in S_f(\boldsymbol{d}) \tag{2.17}$$

In fact, under these assumptions, it is possible to demonstrate, by means of a reductio ad absurdum, that two distinct equilibrium flow vectors $\boldsymbol{f}_{DUE}$ cannot co-exist. To be precise, these conditions are able to guarantee the uniqueness of link costs $\boldsymbol{c}_{DUE}$ and path costs $\boldsymbol{g}_{DUE}$ at the equilibrium, respectively by means of equations (2.4) and (2.6). However, the same cannot be stated for the equilibrium path flows, since different path flow vectors might exist associated with the same link flow vectors. Though, this is not matter of concern, since the final assignment objective is to carry out link attributes (e.g. link flows) so as to be able to derive network performance and, thus, path flows are considered only as a means to achieve that.

Finally, in the case of *Stochastic User Equilibrium* assignment models, link costs depend on flows and the perceived utility of a path is assumed as stochastic. The mathematical formulation of this assignment model can be obtained by combining equations (2.3), (2.4), (2.6) and (2.13), and results in a *fixed-point problem* in which it is necessary to find a flow vector that reproduces itself on the basis of the correspondence defined by the supply and demand models.

$$\boldsymbol{f}^{*}_{SUE} = \boldsymbol{f}_{SNL}\left(\boldsymbol{c}\left(\boldsymbol{f}^{*}_{SUE}\right), \boldsymbol{d}\right) = \boldsymbol{\Delta} \cdot \boldsymbol{P}\left(-\boldsymbol{\Delta}^{T} \cdot \boldsymbol{c}\left(\boldsymbol{f}^{*}_{SUE}\right)\right) \cdot \boldsymbol{d} \quad \in S_{f}\left(\boldsymbol{d}\right) \forall \boldsymbol{c} \tag{2.18}$$

Since in this case, as already mentioned, the perceived utility of a path is a random variable, for calculating the elements of the link-path incidence matrix it is necessary to rely on suitable random utility models.

At this point, it is worth clarifying the properties of fixed-point problems.

In general, the existence and uniqueness of the solution of a fixed point problem are assured by conditions of Banach's theorem (Banach, 1992) which also enables the specification of an asymptotically convergent algorithm. However, since only a limited class of functions meets these requirements, in the following we will refer to weaker conditions and, clearly, to the specific case under examination.

The fixed point problem (2.18) has, at least, one solution if stochastic uncongested network assignment function $\boldsymbol{f} = \boldsymbol{f}_{SUN}\left(\boldsymbol{c}, \boldsymbol{d}\right)$ and cost functions $\boldsymbol{c} = \boldsymbol{c}\left(\boldsymbol{f}\right)$ are continuous. In particular, the property of the continuity is assured by the Brouwer's theorem (Brouwer, 1912).

The fixed point problem (2.18) has, at most, one solution if stochastic uncongested network assignment function $\boldsymbol{f} = \boldsymbol{f}_{SUN}\left(\boldsymbol{c}, \boldsymbol{d}\right)$ is non-increasing monotone with respect to the link costs and cost functions $\boldsymbol{c} = \boldsymbol{c}\left(\boldsymbol{f}\right)$ are strictly increasing over the set of feasible link flows. Sufficient condition so that cost functions are strictly monotone is that $\boldsymbol{Jac}[\boldsymbol{c}(\boldsymbol{f})]$ is positive definite over the set of feasible link flows S_f.

The conditions of existence and uniqueness of link flows $\boldsymbol{f}_{SUE}$, guarantee, in turn, the uniqueness of link costs at the equilibrium $\boldsymbol{c}_{SUE}$, which can be obtained in correspondence with equilibrium link flows by means of equation (2.4). Moreover, as opposed to the *DUE* case, by means of equations (2.3) and (2.13), it can be stated that these properties of existence and uniqueness are valid also for path costs and flows.

A recap of the described models is set out in table 2.1.

	PATH CHOICE MODEL	
	Deterministic	Stochastic
Uncongested network	DUN	SUN
Congested network	DUE	SUE

Table 2.1 Assignment models classification (source: Cascetta, 2009)

For the sake of completeness, it is worth pointing out that in the case of the equilibrium models, it can be meaningful to implement an elastic demand assignment which assumes that demand flows depend on congestion costs (i.e. path costs resulting from congestion). Clearly, by adopting the assumption of the elasticity of demand, the degree of complexity arises and, in particular, two different methods can be adopted: an internal or an external approach (for further insights, see Cantarella et al., 2015; Cascetta, 2009).

Although the provided discussion concerning assignment models represents the bedrock for any kind of assessment related to transportation systems, it is worth specifying that it is not exhaustive. More complex assignment models are, for instance, assignment models with pre-trip/en-route path choice, relevant in public transport systems with high frequency and low reliability conditions. In particular, in this case, we are talking about trip strategies (i.e. hyperpath) characterised by the fact that en-route choices are made during the trip itself at each diversion node where different lines are available. For further insight, see Nguyen and Pallottino (1986), Nuzzolo et al. (2002), and Spiess and Florian (1989). Other aspects which have been addressed in the literature are multiuser class assignment and elastic demand (Cantarella, 1997), multi-modal assignment (D'Acierno et al., 2002; 2011), system optimal assignment models (Mahmassani and Peeta, 1995), intra-period (within day) dynamics models (Ben-Akiva et al., 1984; Cascetta, 2009), inter-period (day-to-day) dynamics assignment (Cantarella and Cascetta, 1995).

2.2 Rail simulation models

As shown by Montella et al. (2000), both in the design phase and in the management phase, it is necessary to rely on suitable simulation techniques, which allow to identify the effects of any intervention, before being put into practice, so as to give an adequate support to the decision making process. The railway simulation models proposed in the literature can be classified according to different criteria. A first criterion concerns the level of detail adopted for the representation of the network and enables the distinction among macroscopic, microscopic and mesoscopic models.

Macroscopic rail models describe the network by means of a graph whose nodes indicate the various stations and whose links usually define the frequency and the travel times of the various trains. The main advantages of macro approaches lie in the fact that a limited set of input data and a low computational effort are required. This makes them able to deal with large-size networks in a reasonable computation time and, therefore, they are usually implemented in planning phases to carry out strategic evaluations related to different infrastructure scenarios or to solve routing problems which consist in the definition of train paths without time restrictions. On the other hand, the low degree of detail adopted in the network representation affects the accuracy of results. Indeed, it usually contains a notably inferior number of nodes and links compared to the microscopic model and considers infrastructure in a more abstract manner. This implies the inability of these models to reproduce some aspects such as signalling equipment installed and layout of station tracks; therefore, they are incapable, for instance, to detect train conflicts or to provide a reliable estimation of running times.

Microscopic rail models, by contrast, portray the networks in great detail. They take into account information concerning tracks (e.g. the number, the length and the alignment of the block sections, speed, gradient), features of the signalling system (e.g. signal position, release points, permissive occupancy), layout of stations (e.g. number of tracks, length of platforms, shunting yards, points, vehicle depots), characteristics of the rolling stock (e.g. acceleration/deceleration features, tractive/effort diagram, total and adherence load), operational information (e.g. departure/arrival times, routes, alternative platforms, timing points, dwell times, connections between runs) as well as safety conditions. The variation of each attribute leads to the creation of a new node and, therefore, necessarily of a new link. Such a modelling of the infrastructure can be used for operational needs such as calculating travel times, performing timetable, detecting probable train conflicts, addressing disruptions conditions and testing rescheduling strategies. Therefore, they provide very accurate results against the necessity of collecting a large amount of data and the need of a high computational effort.

Mesoscopic rail models represent an intermediate approach between the macroscopic and microscopic models and, hence, they are described at this stage. They simulate the performance of the network at an aggregate level, by using aggregate variables such as capacity, flow and density. Traffic, therefore, is represented by convoy packets with identical characteristics (destination, routing behaviour, etc.) which propagate on the network. The main advantage of such models concerns the minimisation of the effort necessary for the representation of complex problems. Indeed, they allow to focus only on the effectively relevant elements and neglect factors which, on the contrary, are not pertinent to the true aim of the study. This permits a simplified simulation of articulated contexts in order to respond to both strategic and tactical needs.

Hereafter, various software packages and tools implementing these models are described.

A first example of macro-simulation model is *NEMO* (*Network Evaluation MOdel*), which has been developed by Sewcyk and Kettner (2001) for supporting the planning phase. Indeed, it can compute arising costs and earnings for different scenarios thus allowing a comparison between them on the basis of an economic evaluation.

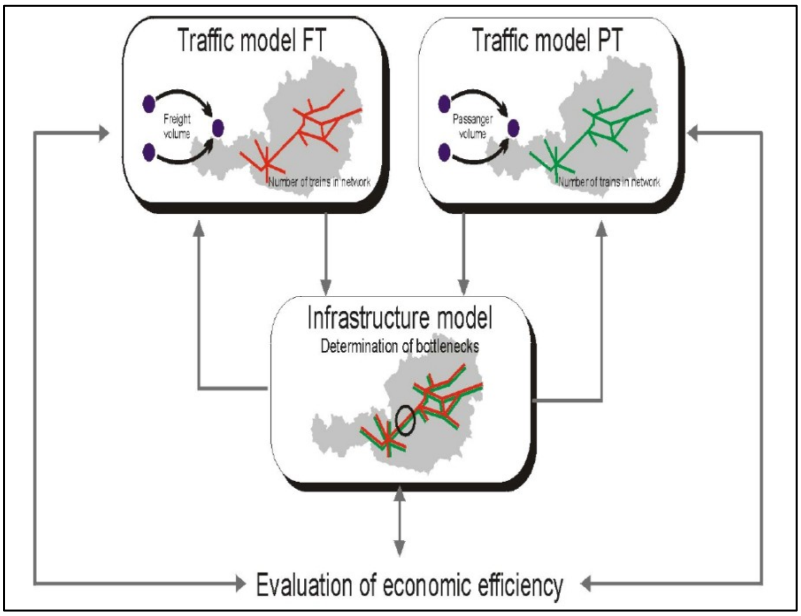


Figure 2.7 *NEMO* model (source: Sewcyk and Kettner, 2001)

NEMO is composed of four different modules (figure 2.7): the *infrastructure module*, in which the railway network is stored as a link-oriented graph, the two *traffic modules*, which treat separately the case of passengers and freight, and the *evaluation module*. In particular, the two traffic modules have a very similar layout, since both of them compute the traffic volume and perform the route search as well as the evaluation of train composition. However, in the case of passengers, the model analyses the traffic relations by dividing traffic according to different destination regions and assigning it to the network; on the other hand, in the case of freight, the traffic volume is assigned to production systems by taking into account their available capacity. Therefore, traffic volume and train composition are combined for obtaining the total infrastructure load for each considered time slice. Finally, the evaluation module derives costs and earnings on the basis of computed trains, demand and load of infrastructure, as well as of previously fixed quantities, such as earnings for a given transport service and costs of infrastructure and rolling stock.

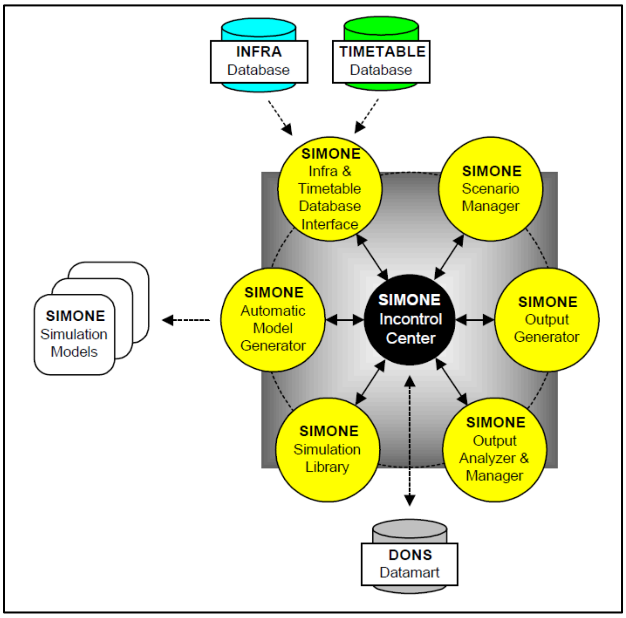


Figure 2.8 *SIMONE* framework (source: Middelkoop and Bouwman, 2001)

Another macro-simulation model is *SIMONE* (*SImulation of MOdel NEtwork*), developed by Middelkoop and Bouwman (2001), whose possible applications regard strategic planning decisions (e.g. the possibility to build a new railway infrastructure or the allocation of network capacity to train operating companies) and the assessment of the stability and robustness of timetables. As shown in figure 2.8, SIMONE presents a quite complex architecture. The core of the model is the *Incontrol Centre* which represents the simulation environment where all information is collected, processed and reworked in a comprehensive view. Other crucial elements are represented by the *Automatic Model Generator* and the *Infra and Timetable Database Interface*. In particular, the former can generate a simulation model without user intervention, on the basis of the set of models stored in the *Simulation Library*; the latter allows the model to interface with the DONS database (i.e. the database of the Dutch railway network). Therefore, this module can, by combining data from DONS database related to infrastructure, timetable and travel demand, generate automatically cyclic timetables. Finally, it is worth quoting the presence of two specific modules for managing the output which are the *Output Generator*, aimed at producing the outcome, and the *Output Analyser and Manager* used for analysing the results related to the different examined scenarios and comparing them in terms of impacts on the system.

Among macro-simulation software packages, it is worth citing also *TransCAD* which is a commercial GIS (Geographic Information System) software specifically developed for transportation analysis. Indeed, besides graphic elements and related databases, it holds transportation network modelling skills by means of which it is able to simulate the supply and demand features of a certain transport system and their interactions. In particular, it can perform different traffic assignment procedures such as 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 as well as dynamic equilibrium traffic assignment.

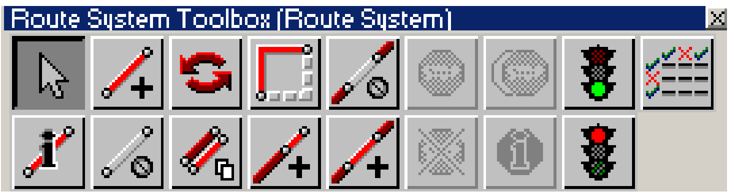


Figure 2.9 Toolbox for transit networks

Specifically, railway lines are treated by the software as transit lines and their representation is carried out by means of the *Route System Toolbox* (figure 2.9). In this way, it is possible to model the line with its route and stops and, additionally, associate to it relevant features, such as headway and capacity, which are stored in the related database. Two different kinds of stop can be set up, namely:

- *physical stop* indicating the physical presence of a station where different lines can stop;
- *route stop* associating to a particular rail service line.

Then, network representation can be completed with additional links (e.g. pedestrian links, connectors) or nodes (e.g. centroids), by characterising them with the related attributes which are saved in the associated dataview.

After the implementation of the transit network, a matrix of passenger flows between origin and destination locations can be uploaded and the transit assignment can be performed. The output is a specific database (i.e. *Transit flows*) which provides link levels and aggregate ridership statistics at every stop along each route such as, for instance, boarding and alighting counts, stop-to-stop flows and route-to-route transfers.

An example of transit network representation by means of TransCAD is shown in figure 2.10.

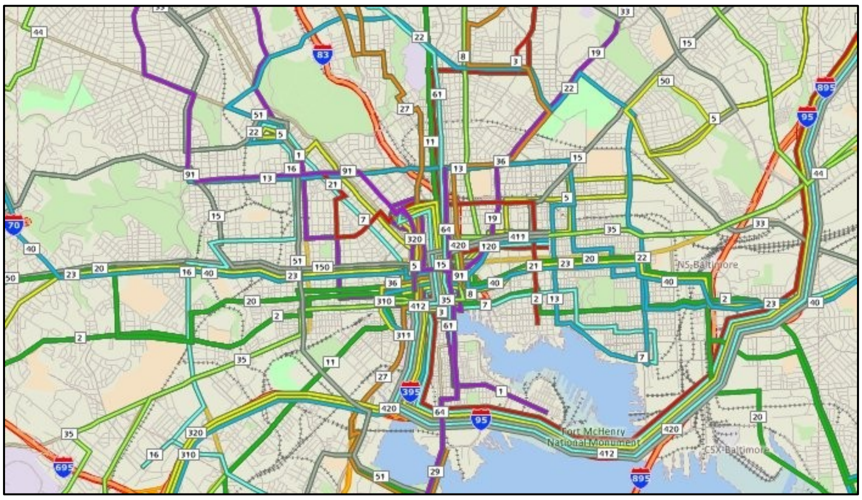


Figure 2.10 Transit network representation in TransCAD (source: www. calliper.com)

On the other hand, among the micro-simulation models, it is worth mentioning the software *RailySis*, developed by Radke and Bendfeldt (2001). It is essentially aimed at simulating different operational scenarios and comparing them in terms of timetable. For this purpose, a very detailed modelling of delays is performed by properly taking into account their stochastic nature. The structure of the model is shown in figure 2.11.

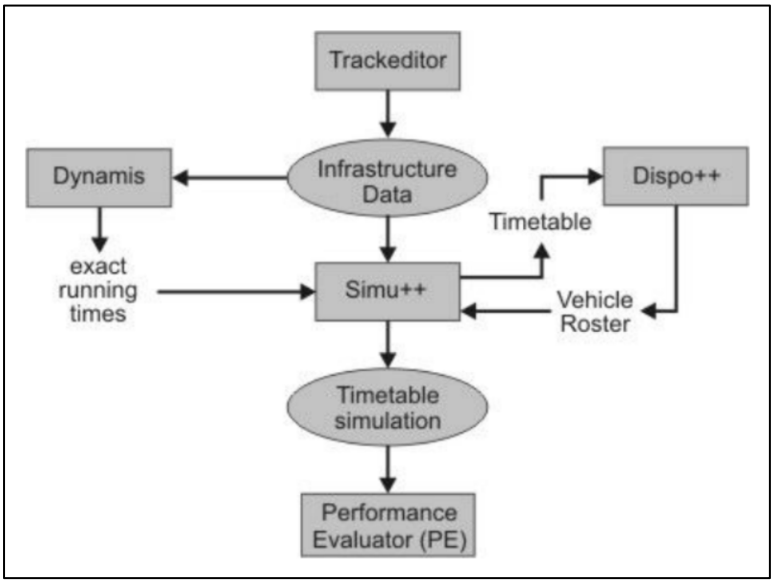


Figure 2.11 *RailSys* architecture (source: Radke and Bendfeldt, 2001)

First of all, the infrastructure data have to be implemented by means of a dedicated editor. Then, these data are used by the modules *Dynamis* and *Simu++* which interact with each other, as well as with *Dispo++*, for carrying out a feasible timetable. In particular, the interaction among the three above mentioned modules consists in the following steps: first of all *Dynamis*, on the basis of infrastructure data and vehicle features, computes running times which are transferred to the *Simu++*; then the latter derives a first attempt timetable, on the basis of which *Dispo++* provides the vehicle roster. At this point, *Simu++* detects eventual inconsistencies between timetable and vehicle allocations (e.g. unfeasible connection times and different transit tracks in stations for some trains), if any, and re-transfers the new timetable to *Dispo++*, until a feasible condition is reached. Finally, all feasible timetable configurations obtained are evaluated and compared with each other, in terms of stability, by the *Performance Evaluator*.

In the following, other two micro-simulation software packages are described: *OpenTrack* (Huerlimann, 2001; Nash and Huerlimann, 2004) and *EGTRAIN – Environment for the design and simulaTion of RAIlway Networks* (Quaglietta, 2011; Quaglietta and Punzo, 2013), whose structures are depicted respectively in figures 2.12 and 2.13.

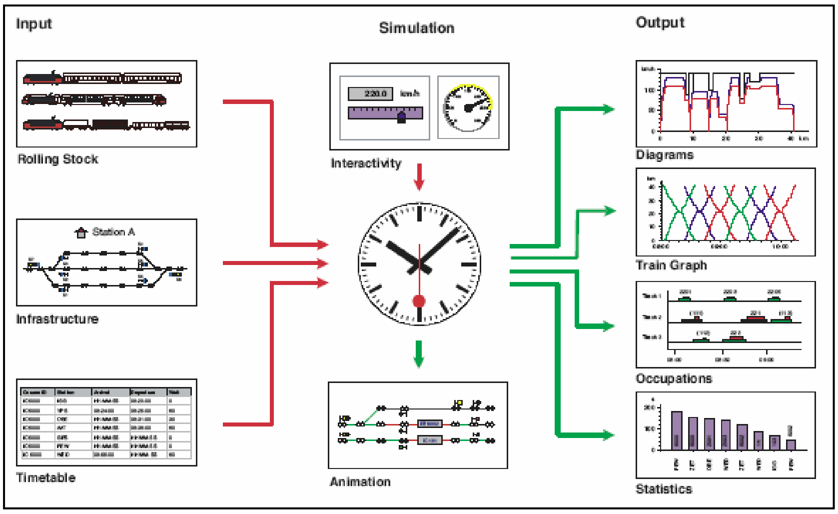


Figure 2.12 *OpenTrack* structure (source: Nash and Huerlimann, 2004)

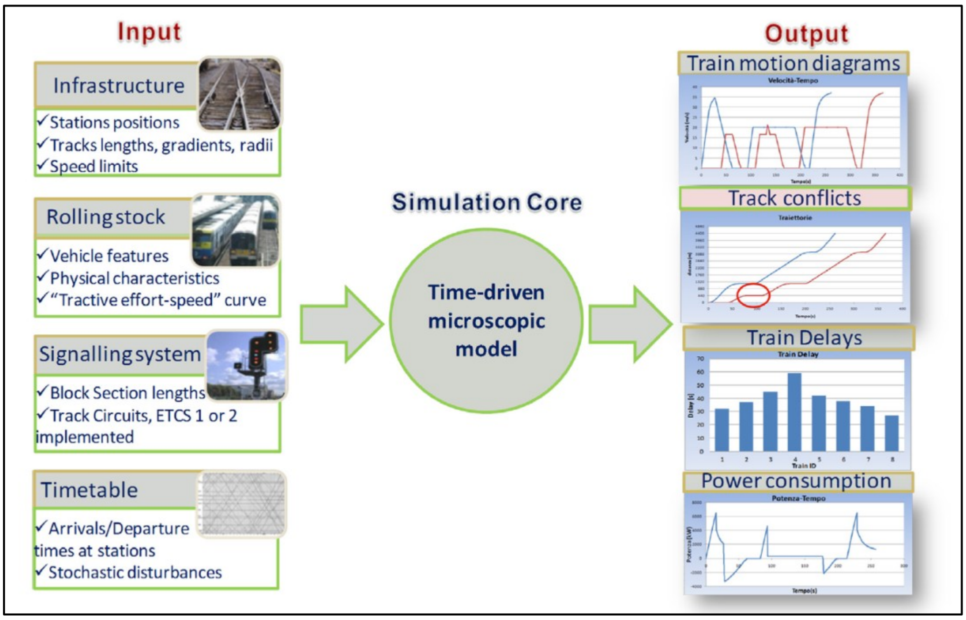


Figure 2.13 *EGTRAIN* framework (source: Quaglietta, 2011)

As can be seen, they are built on a very similar architecture which is based on: different modules providing input data, a simulation core and several possible outputs to be carried out. The input modules provide data concerning infrastructure, signalling systems, stations, rolling stock and planned timetable which are modelled with a high degree of detail.

In particular, in order to represent the motion of rail convoys in the most realistic way, OpenTrack adopts the so-called *colon-graph or double vertexes graph*. In such a graph, each node can be crossed if and only if both vertexes of the node are crossed and this allows to obtain a valid representation of the network, especially in the case of points. Moreover, it follows a hierarchical structure which dictates to carry out specific elements in a given order, that is block sections, routes by connecting contiguous sections, itineraries by connecting contiguous routes, runs by combining an itinerary with a specific kind of train and a specific departure/arrival time and, finally, the planned timetable made up of all runs on duty. Regarding the simulation core, this software performs a mixed discrete/continuous simulation process which calculates both the continuous numerical solution of the differential motion equations for the trains, by means of the Euler's method (Butcher, 1987), and the discrete processes of signal box states and delay distributions. Furthermore, it presents a very user-friendly GUI (Graphical User Interface) which displays the infrastructure as a double-vertex graph, together with the animation of trains along their route, and offers the possibility of visualising interactive messages and measurement tools during the simulation. EGTRAIN, by contrast, does not provide any GUI (the interface with the user is constituted by a simple Win-32 Console window) and performs a time-discrete simulation (i.e. the clock goes ahead with discrete time where each time instant t is obtained as the sum of the previous time instant $t-1$ and the defined time step $\Delta t : t = t - 1 + \Delta t$).

Finally, both simulation tools provide similar outputs: train motion diagrams (speed-distance, speed-time, distance-time trajectories); occupation times of rail sections (in both numerical and graphical format); track conflicts; statistics, such as the percentage of delayed trains at a certain station and the overall train punctuality (fixing a certain delay threshold); energy consumption diagrams (electrical or mechanical power-time diagrams, electrical or mechanical energy-space diagrams). However, these software packages are aimed at simulating exclusively train service, without considering its interactions with travel demand, whose influence has an impact on the estimation of dwell times to be implemented in the simulated timetable. Moreover, it is worth noting that, while OpenTrack is a commercial software whose code is clearly unknown, since it works as a black-box, EGTRAIN is a software developed for research purposes in *C++* language and, therefore, offers the possibility of developing new functions and performing interactions with other models in a very simple way.

By moving to the class of mesoscopic models, it is worth citing the contribution of Quaglietta et al. (2011), concerning the development of an event-driven multi-train simulation model which has been implemented by means of the *Stochastic Activity Networks* (*SANs*) formalism (Meyer et al., 1985; Movaghar and Meyer, 1984). Specifically, as shown by Sanders and Meyer (2001), *SANs* can be considered a stochastic variant of Petri nets developed for dealing with non functional properties of a system such as its performability. Indeed, this mesoscopic model aims to perform a *RAMS* analysis (CENELEC, 1999), so as to assess global effects of breakdowns on rail service and simulate strategic operations for re-establishing ordinary service conditions (e.g. moving a broken train to the depot and substituting it with a spare). The computational efficiency of this model is due to the fact that only main events, such as modifications in signal aspects or train arrivals/departures from sections, joints and stations, are taken into account; while, events relative to train acceleration/deceleration phases are neglected.

Moreover, Marinov and Viegas (2011) developed a mesoscopic model for simulating freight train operations in a rail network, by means of the event-based simulation computer package *Simul8* which adopts a decomposition approach. In particular, it consists in separating the whole system into its components (i.e. rail lines, rail yards, rail stations, rail terminals and junctions) and capturing the interactions existing among them by modelling the network as an interconnected queuing system, so as to preserve the global perspective in the estimation of operating performance. More in detail, rail freight terminals are modelled by means of *Work Centres* and *Storage Areas*, which are interconnected by means of *Work Flow Arrows* representing the routing of trains. Specifically, Work Centres (i.e. where a freight train is served by a component of the rail system) simulate the operating procedures in the case of freight trains and the related attributes are inbound traffic, service pattern and outbound traffic. The service times in Work Centres are set up equal to the dwell terminal times of freight trains. On the other hand, Storage Areas replicate the waiting phase of a freight train which holds on to be processed by a component of the network. Their control parameter is represented by the capacity. Finally, in order to simulate the departure and arrival patterns of each work item, i.e. a freight train, *Work Entry Points* and *Work Exit Points* are introduced in the modelling framework. Network performance is measured by means of the following indexes: total number of freight trains processed by a given Work Centre, 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.

Finally, De Fabris et al. (2013) proposed a mesoscopic network model for addressing the timetabling design problem. In particular, this model is implemented in the tool *TTPSW* which is able to iteratively generate different timetables, in a reasonable time, so as to perform the cyclic optimisation procedure depicted in figure 2.14.

Approaches aimed at transforming a micro model into a macro one and vice versa are possible. In particular, as shown by Hansen and Pachl (2008), the derivation of a macroscopic model from a microscopic framework is known as *bottom-up* approach; while, the *top-down* approach can be used for generating artificial microscopic infrastructure whose level of detail depends on the addressed problem and the analysed perspective.

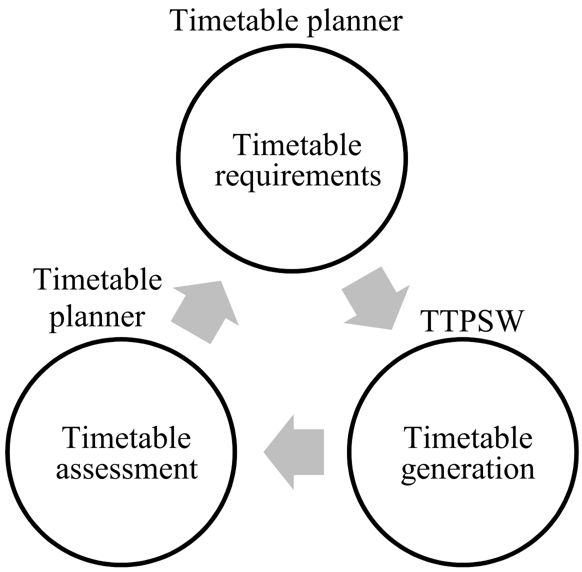


Figure 2.14 Timetable generation cycle (source: De Fabris et al., 2013)

Clearly, the bottom-up approach is the most used, since it is straightforward to be implemented inasmuch the final model requires less information than the starting one. In this context, Eickmann et al. (2003) developed a particular interface enabling data migration between Railsys and NEMO models with the aim of supporting the generation of a conflict-free timetable. Moreover, Schlecthe et al. (2011) derived a macroscopic framework by starting from a microscopic model, implemented in OpenTrack, for determining conflict-free track allocations. In particular, the transformation occurs by means of the aggregation of block sections and station areas, together with the introduction of '*pseudo-nodes*' which replicate the interactions among different convoys. In addition, after having derived the macroscopic model from the microscopic one, the proposed procedure combines them with each other in order to validate the solutions carried out.

Indeed, several contributions in the literature arranged frameworks based on the combination of two different approaches, so as to be able to exploit the advantages of both of them and overtake their drawbacks.

Generally, macro e micro approaches are merged together or combined in an iterative manner. In this context, Bešinović et al. (2015), developed a micro-macro framework for timetable design which consists in performing an iterative adjustment of train running and minimum headway times, with the aim of determining a feasible timetable and, in addition, analysing its stability and robustness features. Moreover, Middelkoop (2010) illustrated the tool *ROBERTO* based on a microscopic representation of the network, whose outputs (i.e. running and headway times) are used as inputs for the macroscopic timetabling model DONS (Kroon et al., 2009).

In the meanwhile, a first attempt of combining mesoscopic and microscopic models was made by Quaglietta et al. (2011) with the aim of carrying out a stochastic analysis in a rescheduling framework. More in detail, the idea consists in exploiting the major computational efficiency of a mesoscopic model for performing millions of ordinary service simulations and, only when a failure occurs, bringing into play the microscopic simulator (i.e. EGTRAIN) so as to derive a more accurate and focused analysis.

Another classification criterion is the implemented processing technique, according to which it is possible to distinguish between *synchronous* and *asynchronous* simulation models. In particular, synchronous approaches simulate the events as they occur in reality; therefore, a chronological progression is followed, with no chance of returning to previous states. Hence, this kind of simulation follows an event-driven approach and is generally applied to evaluate network performance, by taking into account interactions among trains. In asynchronous models, on the other hand, the convoys are simulated according to their class of priority, which means that the simulation is divided into more steps on the basis of a particular principle which is related to the category the trains belong to: the trains belonging to the higher categories are firstly simulated and, therefore, they are not affected by the interactions with other convoys; then, progressively, trains belonging to other categories are processed, until the service is completely simulated. A typical application of such a hierarchical procedure is the construction of a timetable in the planning phase. Examples of synchronous commercial simulation tools, besides the above mentioned OpenTrack and Railsys, 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. On the other hand, examples of asynchronous models are BABSI (Gröger, 2002) and STRESI (Shultze, 1985), both developed at the RWTH Aachen University in Germany.

Finally, according to the assumptions on the distribution of the involved parameters, it is possible to distinguish between *deterministic* and *stochastic* simulation models. The deterministic case deals with parameters characterised by a steady value equal to their average, which means that departure/arrival times, dwell times, travel times etc. are constant. On the other hand, in the case of stochastic simulations, involved parameters are considered as random variables and, therefore, they are modelled by means of their probability density function (*pdf*), as well as the mean and the standard deviation of the pdf itself. Generally, deterministic models are implemented in design phases, while stochastic ones are most suitable for evaluating network performance (e.g. robustness of timetables, stability against disturbance, impacts of operational strategies), since they better reflect the actual conditions. Many types of software, such as the above-mentioned OpenTrack and EGTRAIN, can perform both deterministic and stochastic simulations. Specifically, EGTRAIN allows to simulate stochastic delays and failures; while, OpenTrack is able to take into account stochasticity of train performance, dwell time and delays by performing a set of simulations by randomly changing input parameters.

2.3 Estimation techniques for travel demand flows

In order to carry out accurate results by means of simulation frameworks, the explicit modelling of travel demand has a fundamental role. Indeed, the reconstruction, estimation and prediction of travel demand (Hazelton, 2001) represent a key factor to be addressed in any kind of assessment regarding transportation systems, so as to optimise both planning and management phases. Therefore, the issue of estimating travel demand, in terms of current and potential or expected passenger flows with related characteristics (i.e. departure and arrival stations, adopted time slot, trip duration, etc.), has always been subject of attention in the literature.

In particular, three different methods can be adopted in order to perform this task: *direct estimation* (see, for instance, Brog and Ampt, 1982; Ortuzar and Willumsen, 2011; Smith 1979), *disaggregated estimation* (see, for instance, Ben-Akiva and Lerman, 1985; Domencich and McFadden, 1975; Horowitz, 1981; Manski and McFadden, 1981; Novačko et al., 2014), and *aggregated estimation* (see, for instance, Bera and Rao, 2011; Barcelo and Montero, 2015; Cascetta and Nguyen, 1988; Cipriani et al. 2014).

Direct estimation enables to reconstruct exclusively the present demand, without any capacity for future prediction. Strictly speaking, the O-D matrix is not directly observable in its entirety; in fact, given the huge quantity of data to be collected, carrying out a census would not be economically doable even if, in certain instances, technically feasible. Thus, actually, direct estimation consists in making use of sampling techniques together with inferential statistics methods for extending the information content of a sample to the whole analysed system. Different kinds of surveys may be carried out such as *on-board surveys* (also named as *cordon surveys* if aimed at estimating the crossing demand), *households surveys, destination surveys* and (*e*)*mail surveys.* Additionally, in the last years, given the esponential growth in the field of information and communication technology, further survey methods have been developed, namely CAPI (Computer Assisted Personal Interviewing), CATI (Computer Assisted Telephone Interviewing) and CAWI (Computer Assisted Web Interviewing). However, whatever the adopted approach, as shown by Cascetta (2009), a preparatory design phase of the survey is required, which consists in the definition, first of all, of the sampling unit and the sampling strategy, which could generally be a *simple random sampling,* a *stratified random sampling* or a *cluster sampling*. Then, according to the adopted sampling method, it is necessary to set up the sample size and the estimator to be applied.

In the case, however, that the request arises for a certain predictive capacity, it is necessary to make use of a *disaggregate estimation* of the O-D matrix which consists in specifying (i.e. providing the functional form and related variables), calibrating (i.e. determining numerical values of model parameters) and validating (i.e. verifying the ability of the model to reproduce original data), by means of proper information, a model which manages to reproduce the variations in travel demand as a result of modifications to transport system performance or socio-economic changes. In this case, two different survey approaches can be implemented: *the revealed preference* (*RP*) approach (Cascetta, 2009) which is based on the use of data related to real traveller behaviour; and the *stated preference* (*SP*) approach (see, for instance, Ben-Akiva and Morikawa, 1990; Ortuzar, 1992), which is based on the statements of travellers related to their potential choices in the case of a hypothetical scenario, which has to be appropriately described and illustrated in order to make user declarations as reliable as possible. With the use of this second approach, the predictive capabilities of the calibrated demand models can be improved. Hence, once the functional formulation of the model, together with the types of attributes to be considered, are specified, it is necessary to carry out the calibration phase by means of which a numerical value is associated to each involved parameter. Generally, in order to calibrate a disaggregate demand model, the *Maximum Likelihood* (*ML*) method is performed. This approach consists in calculating numerical values of the unknown parameters by maximising the probability of observing the choices made by a sample of users. The formulation of the likelihood function $L(\cdot)$, under the assumption of a simple random sampling, is the following:

$$L(\boldsymbol{\beta},\boldsymbol{\theta}) = \prod_{i=1,\ldots,n} p^i[j(i)] \cdot (\boldsymbol{X}_i, \boldsymbol{\beta}, \boldsymbol{\theta}) \tag{2.19}$$

where $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ are the vectors of the model parameters; $p^i[j(i)]$ is the probability that each user i chooses the alternative $j(i)$; $j(i)$ is the alternative actually chosen by the individual; $\boldsymbol{X}_i$ is the vector of the explanatory variables for the user i.

Therefore, by maximising equation (2.19), or its natural logarithm, it is possible to carry out the maximum likelihood estimate of the vectors of the parameters $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$. Lastly, a validation phase with proper statistical tests has to be performed.

Finally, the *aggregate estimation* of travel demand indicates a correction procedure of the O-D matrix which consists in updating a previously known matrix, through aggregate type data, such as traffic counts, in order to improve its reliability and guarantee an accurate assessment of the system status in the forecasting phase. In this context, it is worth noting that, in contrast to the sample surveys which are complex and expensive, counts require inferior cost and can be obtained automatically.

This approach is expressed by Cascetta and Nguyen (1988) in terms of an optimisation problem:

$$\boldsymbol{d}^* = argmin_{x\geq 0}\left[\ z_1(\boldsymbol{x},\hat{\boldsymbol{d}}) + z_2(\boldsymbol{v}(\boldsymbol{x}),\hat{\boldsymbol{f}})\ \right] \tag{2.20}$$

where $\boldsymbol{x}$ is the unknown demand vector; $\hat{\boldsymbol{d}}$ is a prior estimate demand vector which is considered the target demand vector; $\boldsymbol{v}(\boldsymbol{x})$ is the vector of link flows obtained by assigning the demand vector $\boldsymbol{x}$ to the network; $\hat{\boldsymbol{f}}$ is the vector of detected link flows.

The aim is to obtain a matrix $\boldsymbol{d}^*$ which is as close as possible to the prior estimate, and that, once assigned to the network, is able to re-produce link flows as close as possible to those detected. Therefore, this procedure can be considered as the inverse assignment problem (figure 2.15). Indeed, in the assignment process, starting from the knowledge of supply, demand and the model which regulates path choice, link flows on the network are defined; on the contrary, in the estimation of the *O-D* matrix, starting from the detected link flows, together with the knowledge of supply and path choice models, the computation of demand is performed.

The importance of the presence of a prior estimate demand vector $\hat{\boldsymbol{d}}$ lies in the fact that, since the number of OD pairs is generally much higher than the number of detected link flows, without such a vector, the problem would result intrinsically undetermined.

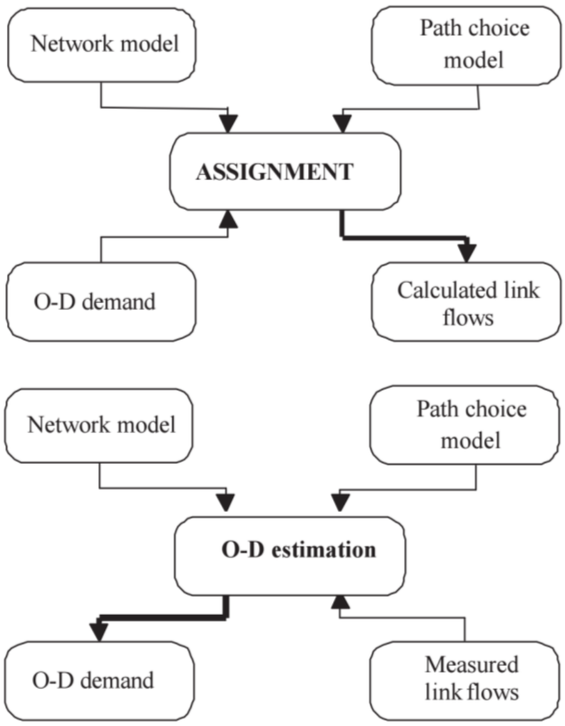


Figure 2.15 Relationship between estimation of O-D flows with traffic counts and traffic assignment (source: Cascetta, 2009)

In the case of congested networks, the estimation problem of *O-D* matrix can be formulated as a fixed-point problem or, alternatively, by means of a bi-level optimisation framework. In this second approach, the upper level represents the estimation problem, while the lower level addresses the network assignment problem. In particular, Cascetta and Postorino (2001) proposed different fixed-point algorithms for congested networks. On the other hand, contributions related to the bi-level optimisation methodology are, for instance, those of Florian and Chen (1995), Yang and Yagar (1995), and, more recently, Lu et al. (2013a), and Walpena et al. (2015) which addressed a *DUE* assignment; while Lo and Chan (2003), Wang et al. (2016), and Yang et al. (2001) dealt with a *SUE* approach.

At this point, it is worth addressing the definition of functions $z_1(\cdot)$ and $z_2(\cdot)$ in equation (2.20), which represent goodness of fit measures and can be expressed by means of different estimators. In particular, for the static approach, we can have:

- Maximum Likelihood (Bell, 1983; Maher, 1983; Cascetta and Nguyen, 1988);
- Generalized Least Squares (Cascetta, 1984);
- Bayesian (Maher, 1983).

A complete overview of the features and statistical principles of such estimators can be found in Cascetta (2009).

By moving to *within-day dynamic* contexts, where travel demand varies within the reference time period, the matrix representation consists in a certain number of matrices: they are as many as the temporal intervals into which the reference period has been subdivided. The introduction of a time dimension leads to two different estimation approaches, namely sequential and simultaneous, as shown by Cascetta et al. (1993). In order to describe such approaches, let the total study period H be divided into n_h intervals h (with $h = 1 .... n_h$) of equal lenght T, so that $H = n_h \cdot T$.

In particular, the *simultaneous* estimation can be specified as follows:

$$\left(\boldsymbol{d}_1^* ... \boldsymbol{d}_{n_h}^*\right) = argmin_{\left(x_1 ... x_{n_h}\right) \geq 0} \left[z_1\left(\boldsymbol{x}_1 ... \boldsymbol{x}_{n_h}, \hat{\boldsymbol{d}}_1 ... \hat{\boldsymbol{d}}_{n_h}\right) + z_2\left(\boldsymbol{v}_1 ... \boldsymbol{v}_{n_h}, \hat{\boldsymbol{f}}_1 ... \hat{\boldsymbol{f}}_{n_h}\right)\right] \tag{2.21}$$

The aim is to identify matrices $\boldsymbol{d}_1^* ... \boldsymbol{d}_{n_h}^*$, for each interval h into which the reference time period H is split, which minimise the 'distance', on one side, between the unknown demand vectors $\boldsymbol{x}_1 \ldots \boldsymbol{x}_{n_h}$ and the above-mentioned prior estimate demand vectors $\hat{\boldsymbol{d}}_1 \ldots \hat{\boldsymbol{d}}_{n_h}$; and, on the other side, between the flow vectors $\boldsymbol{v}_1 \ldots \boldsymbol{v}_{n_h}$ (obtained by assigning demand vectors $\boldsymbol{x}_1 \ldots \boldsymbol{x}_{n_h}$ to the network) and the traffic counts vectors $\hat{\boldsymbol{f}}_1 \ldots \hat{\boldsymbol{f}}_{n_h}$.

While, in the case of *sequential* estimation, occurs:

$$\boldsymbol{d}^*_{n_h} = argmin_{x_{n_h} \geq 0} \left[z_1\left(\boldsymbol{x}_{n_h}, \hat{\boldsymbol{d}}_{n_h}\right) + z_2\left(\boldsymbol{v}_{n_h}\left(\boldsymbol{x}_{n_h} \middle| \boldsymbol{d}^*_1 \ldots \boldsymbol{d}^*_{n_{h-1}}\right), \hat{\boldsymbol{f}}_{n_h}\right)\right] \tag{2.22}$$

In this context, one matrix at a time is identified, starting from the first temporal interval and proceeding until the ultimate, maintaining the previously calculated matrices fixed time after time.

Generally, conversely to a simultaneous approach which is usually adopted for off-line estimations, the sequential approach can be used for on-line applications. Indeed, it provides a lower computational complexity, by splitting the addressed problem into different sub-problems which are easier to analyse and, thus, the matrix estimated for each time slice can be used as estimation input for the following time period. On the other hand, however, it presents the drawback of considering, for each demand vector $\boldsymbol{d}^*_h$, limited information consisting in traffic counts associated exclusively with the same interval to which is referred (i.e. $\hat{\boldsymbol{f}}_h$). Therefore, in order to rectify this aspect, different contributions based on the adoption of Kalman filtering methodologies have been proposed (see, for instance, Ashok and Ben-Akiva 1993; Okutani and Stephanades, 1984).

While, regarding the simultaneous approach with the adoption of an assignment matrix, as shown by Bierlaire and Crittin (2004), Cascetta and Russo (1997), and Toledo et al. (2003), it turns out to have a prohibitive computational complexity, even in the case of medium-size networks. Therefore, in order to deal with more feasible computational times and, at the same time, adopt a simultaneous approach which is the most suitable from a conceptual point of view, several *non-assignment based methods* for dynamic O–D matrix estimation have been developed. Within this framework, after pioneering works by Cremer and Keller (1981), Cremer and Keller (1984, 1987), and Nihan and Davis (1987, 1989), more recent contributions proposed the adoption of evolutionary methods (Appiah and Rilett, 2010; Kattan and Abdulhai, 2012, 2011; Kim et al., 2001; Park and Zhu, 2007; Tsekeris et al., 2007), simulating annealing techniques (Stathopoulos and Tsekeris 2004), Bee Colony Optimisation (Caggiani et al., 2012), probe vehicles data (D'Acierno et al., 2009) and artificial intelligence approaches (De Luca and Gallo, 2017; Huang et al., 2013; Kattan and Abdulhai, 2006). Other assignment matrix-free methods are provided by Balakrishna et al. (2008) and Cipriani et al. (2011) which address the simultaneous adjustment of a dynamic traffic demand matrix by means of a gradient-approximation approach representing a variant of the Simultaneous Perturbation Stochastic Approximation (SPSA) path search optimisation method proposed by Spall (1992; 1998). Further variants of the SPSA approach are W-SPSA proposed by Lu et al. (2015) and Antoniou et al. (2015), and c-SPSA presented by Tympakianaki et al. (2015). Similarly, Toledo and Kolechkina (2013) proposed a method based on the use of linear approximations of the assignment matrix in the optimisation iterations and tested several specific solution algorithms which differ in the search direction.

On completion of the above mentioned contributions, regarding both static and dynamic frameworks, it is worth citing the work by Cascetta et al. (2013) proposing a GLS-based within-day dynamic estimator which provides room for significant improvements of the unknowns/equations ratio, thanks to a quasi-dynamic approach based on the intuitive assumption of considering distribution rates as constant within a wider time interval compared to the within-day variation of the generation rates.

Another point, strictly related to these aggregated estimation techniques, is the *Network Sensor Location Problem* (*NSLP*) which addresses the relationship between survey costs and estimation accuracy (Bao et al., 2016; Chung, 2001;

Ehlert et al., 2006; Fei et al., 2007; Shao et al., 2016; Simonelli et al., 2012; Yang and Zhou, 1998). Basically, it consists in identifying the optimal location (i.e. the location which maximises the information content) under a budget constraint (i.e. a given number of count sections). With the advance of *Information and Communications Technologies (ICT)*, in addition to traffic counts, also other kinds of data sources have been implemented to carry out a reliable estimation of O-D matrix, such as GPS data (Moreira-Matias et al., 2016), video recordings (Savrasovs and Pticina, 2017) and mobile-phone data (Tolouei et al., 2016).

However, in the case of railway systems, it is fundamental to take into account some specific issues due to the intrinsic features of the addressed context. First of all, the flows of concern are related to the number of passengers, rather than the amount of vehicles. That gives rise to a first issue to be faced, concerning the kind of passenger flows to be considered, such as flows at turnstiles, boarding or alighting flows, waiting flows and on-board flows. This results in a *spatial* problem related to 'where' to detect passengers. In the case that counting them at the turnstiles is selected, the measurement would be affected by an uncertainty about the direction of the trip. Alternatively, would be possible to acquire data from a single gate, but, in such circumstances, it might not be possible to know how many passengers are not able to board the train because of overcrowding. While, such information could be obtained by carrying out the counts on platforms. In addition to this, a *temporal* problem should be taken into account which lies in the difficulty of identifying a proper reference time interval, given the fact that rail service is a scheduled service. It is this very discontinuity which makes counting at the turnstiles susceptible to a certain degree of uncertainty due to the gap between the time of registration of the users' passage and the moment they reach the platform. Therefore, it appears clear that it is necessary to opportunely design the data acquisition phase according to the target.

As already stated, differently from the sample surveys which are complex and expensive, counts are cheaper to carry out and can be obtained automatically.

The use of automatic devices makes the detection task as easier and more efficient; however, it is not immune to incidents. First of all, it might happen that, because of a device failure, there could be effects on the entire measurement. A typical situation in which this could happen is if the target is to reconstruct the distribution of the passengers on the platform by carrying out counts at each gate. Indeed, in this case, if a detector of a single gate was damaged, this would make also the counts of the other gates as useless, invalidating the measurement for the whole platform. Other issues to be taken into account are the presence of exchange points between two lines and, in some contexts, also the possibility of evasion.

However, in the literature, a large part of contributions related to the detection of passengers flows in railway contexts are essentially based on the use of e-ticketing and automatic fare collection systems and neglect these matters (see, for instance, Gavriilidou et al., 2017; Nagy, 2016; Tavassoli et al., 2017; Zhao et al., 2007). Therefore, starting from the above mentioned aggregate estimation techniques of the O-D matrix, which are considered acquired by the literature from a conceptual point of view, this work aims at customising them to the railway case, by developing resolution methods suitably designed for satisfying specific requirements related to the intrinsic nature of such a transport system.

2.4 Simulation algorithms

After the analysis of assignment models and their properties of existence and uniqueness, in this section, related resolution methods are described. For this purpose, it is necessary to provide some basic concepts about the structure of *shortest path algorithms*, since they are used as sub-routine within assignment procedures. Therefore, their computational efficiency is an important factor to be evaluated.

Since applications relative to transportation network assignment require the determination of the shortest path only among centroids, rather than among all the possible pairs of nodes, the case of concern is the correspondence one-to-many: from an origin centroid o to all the network nodes (*forward* *shortest paths*) or from all the network nodes to a destination centroid *d* (*backward shortest paths*).

Such methods are based on the iterative updating of a tentative shortest path tree, by means of the application of the *Bellman principle* (Bellman, 1958), which states that *a shortest path is made up of shortest paths*.

Let,

$c_l = c_{ij} \geq 0$ be the cost on link $l = (i, j)$;

$Z_{ij} \geq 0$ be the cost of the shortest path between any pair of nodes *i* and *j*.

Hence, if the link (i, j) belongs to the shortest path between *o* and *d*, then

$Z_{o,i} + c_{ij} + Z_{j,d} = Z_{o,d}$;

otherwise:

$Z_{o,i} + c_{ij} + Z_{j,d} \geq Z_{o,d}$.

They stop when no further updates can be performed.

Furthermore, according to the adopted node-list management strategy, these algorithms can be classified in *label-setting* and *label-correcting* methods. In the first case, the node-list adopts an increasing order and nodes are made definitive in order of increasing shortest path cost. On the other hand, label correcting algorithms do not require an ordered node-list and cannot guarantee that each node will be examined once. Moreover, in this case, all tentative values become definitive only at the end of the algorithm. Examples of label setting and label correcting algorithms are respectively *Dijkstra* (Dijkstra, 1959) and *L-deque* (Pape, 1974). Clearly, the best ever algorithm does not exist and, therefore, it is necessary to select which is better to use on a case-by-case basis, according to the specific context under examination.

Within the assignment models, we can have different criteria for analysing the available paths. First of all, it is possible to take into account all feasible paths (*exhaustive approach*) or only those which meet specific conditions (*selective* *approach*). Moreover, they can adopt an *explicit* or *implicit* path enumeration. In many cases, because of the large number of paths involved, an explicit approach is not feasible and the implicit enumeration is preferred. The convenience of adopting this method, which overrides the definition of path features, lies in the fact that, actually, the ultimate aim of assignment is obtaining link attributes so as to calculate the network performance and, thus, paths are considered only a means of reaching this target.

After these general considerations, in what follows, we will proceed to analyse the most frequently used algorithms in the case of assignments models described in paragraph 2.1.

Deterministic Uncongested Network models are solved by means of a very simple procedure known as *All or Nothing* which, according to the result of the implemented shortest path algorithm, assigns, for each *OD* pair, the total demand flow to the links belonging to the shortest path and no flow to all other links. This algorithm can be based on a simultaneous or a sequential approach. Obviously, the former is more efficient but, in this case, an algorithm with ordering is required for shortest path trees to each destination.

Regarding the *Stochastic Uncongested Network,* mainly two different algorithms have been proposed: the *Dial algorithm* (Dial, 1971) which is based on a Multinomial Logit path choice model and the *MonteCarlo* algorithm (Sheffi and Powel, 1982) which is based on a Probit path choice model. Therefore, they differ in the specification of path choice model; however, both are based on an implicit path enumeration.

In particular, the *Dial algorithm* adopts a selective approach which considers as belonging to the set of relevant paths only the *efficient paths with respect to the origins,* which are made up of links $l = (i, j)$, termed *efficient links*, such that the cost of the shortest path to reach the initial node *i,* from the origin *o,* is inferior to the cost of the shortest path to reach the final node *j*, say $Z_{o,i} < Z_{o,j}$ (Cascetta, 2009). As an efficiency criterion, it is possible also to define *efficient paths with* *respect to the destinations* or *efficient paths with respect to both the origin and the destination*. For the sake of simplicity, only the case of *efficient paths with respect to the origins* will be considered. The definition of an efficient path can be carried out by associating to the links the cost or any other positive attribute, such as the length or the zero flow cost. The adoption of quantities which are independent of congestion is relevant when a stochastic network loading function is used for the definition of a stochastic user equilibrium model as it guarantees the sufficient condition for the uniqueness of the equilibrium state as well as for the convergence of the stochastic equilibrium algorithms.

By using the well-known Multinomial Logit model, the probability $p_{od,k}$ of choosing a generic path *k*, belonging to the set K_{od} of paths which connect the origin-destination pair *od*, is given by:

$$p_{od,k} = \frac{exp(-g_k / \theta)}{\sum_{k' \in K_{od}} exp(-g_{k'} / \theta)} \tag{2.23}$$

where g_k and $g_{k'}$ are the costs associated respectively to the paths *k* and *k'*; θ is a parameter of the model to be calibrated.

By expressing path costs as a function of link costs, according to the relation (2.6), the previous equation can be re-written as follows:

$$p_{od,k} = \frac{\prod_{(i,j) \in k} exp(-c_{ij} / \theta)}{\sum_{k' \in K_{od}} exp(-g_{k'} / \theta)} \tag{2.24}$$

If each path is considered as a sequence of nodes j and links (i,j), the probability $p_{od,k}$ can be expressed as the product of the probabilities, $Pr[(i,j)/ j]$, to choose each link (i,j) belonging to the path *k*, conditional upon the crossing of the final node *j*, i.e.:

$$p_{od,k} = \prod_{(i,j) \in k} Pr[(i,j)/ j] \tag{2.25}$$

The formulation (2.25) is the equivalent to (2.24) if the probability $Pr[(i,j)/j]$ of choosing the link (i,j), conditional on the final node *j*, is defined with a Multinomial Logit model of parameter θ in which: the alternatives are the efficient links (i,j) incident to node *j* (i.e. all the efficient links entering node *j*) and the systematic utility of each alternative $V_{ij/j}$ is the sum of the opposite of the link cost c_{ij} and of a logsum variable Y_i, synthetically taking into account the utilities of all the efficient paths, from the origin *o* to the initial node *i* of the link.

Therefore, under these assumptions, the probability $Pr[(i,j)/j]$ can be expressed by the relation between the weight of link (i,j), indicated by w_{ij}, and the weight of node *j*, indicated by W_j:

$$Pr[(i,j)/j] = \frac{w_{ij}}{W_j} \tag{2.26}$$

The weight of a generic link (i,j) can be calculated on the basis of the cost of the link and the weight of the initial node *i*, as follows:

$$w_{ij} = \begin{cases} exp(-c_{ij}/\theta)\cdot W_i & \text{if } Z_{o,j} < Z_{o,j} \\ 0 & \text{if } Z_{o,j} \geq Z_{o,j} \end{cases} \tag{2.27}$$

Moreover, the weight of a generic node *j* can be calculated noting the weights of all the links belonging to the backward star of node *j*, as follows:

$$W_j = \sum_{(m,j)\in BS(j)} w_{mj} \tag{2.28}$$

where $BS(j)$ is the backward star of node *j*.

Equations (2.27) and (2.28) allow the computation of the weights of the links and nodes by starting from a certain origin and, progressively, moving away from it; while equation (2.26) enables the allocation of flows by starting from the most distant nodes from the examined origin and, gradually, closing in on it. Dial's algorithm is based on this principle.

Clearly, the use of the Multinomial Logit model necessarily implies the fact that very similar paths (that are made up of a relevant part of the same links) are perceived as being independent and, therefore, their probabilities are overestimated. However, as already pointed out, it is possible to remedy to this drawback by adopting some variants of the Logit model, such as the C-logit model with an appropriate specification of the commonality factor.

On the other hand, the *MonteCarlo* algorithm adopts a Probit path choice model (Daganzo and Sheffi, 1977). This method is able to take into account overlapping paths by introducing a positive covariance between the perceived utilities of two paths sharing some links. In particular, the adopted assumption is that the random residuals are distributed according to a Multivariate Normal $MVN(0,\Sigma)$ with null mean and variance-covariance matrix Σ which can be any symmetric positive definite matrix. Moreover, it is assumed that, for each *OD* pair, only *elementary paths* (i.e. those without loops) are relevant.

However, this approach does not enable an explicit calculation of paths choice probabilities and, therefore, it is necessary to rely on a Monte Carlo technique in order to obtain unbiased estimates of path choice probabilities and related path and link flows. In particular, a sample of perceived link cost vectors has to be generated; then, for each sampled vector, demand flows are assigned to the shortest paths with an All-or-Nothing assignment algorithm and the mean of the link flows obtained for the different link cost vectors of the sample is an unbiased estimate of Probit SUN link flows. Therefore, this algorithm does not provide link flow values, but only a sequence of unbiased estimates whose precision increases with the number of iterations. Hence, in practice, it proceeds until a halt condition occurs; for example, a pre-assigned maximum number of iterations i_{max}. In this case, the time necessary for performing the algorithm is

i_{max} times bigger than that necessary to carry out a deterministic loading of the network. For more details about this algorithm see Cascetta (2009).

By moving to the case of *Deterministic User Equilibrium* models, it can be proved that, if the Jacobian related to cost functions $\boldsymbol{Jac}[\boldsymbol{c}(\boldsymbol{f})]$ is symmetrical, the *DUE* problem can be formalised as an optimisation problem (Beckman et al., 1956):

$$\boldsymbol{f}^* = \underset{\boldsymbol{f} \in S_f}{argmin}\ z(\boldsymbol{f}) = \int_0^f \boldsymbol{c}(\boldsymbol{x})^T d\boldsymbol{x} \qquad \text{with } \boldsymbol{f} \in S_f(\boldsymbol{d}) \tag{2.29}$$

where, according to Green's theorem, the integral is independent of the integration path.

It is worth noting that, if cost function $\boldsymbol{c}(\boldsymbol{f})$ is continuous and differentiable with symmetrical and semi-definite positive Jacobian $\boldsymbol{Jac}[\boldsymbol{c}(\boldsymbol{f})]$, the optimisation framework and the equilibrium problem become equivalent; consequently, the solution of problem (2.29), obtained by means of a proper optimisation algorithm, becomes feasible for the deterministic equilibrium assignment and vice versa. In this case, the adopted resolution method is the *Frank-Wolfe* algorithm (LeBlanc et al., 1975).

It starts from a feasible initial solution $\boldsymbol{f}^0 \in S_f(\boldsymbol{d})$, which can be easily obtained, for example, with a *DUN* algorithm using zero flow costs $\boldsymbol{f}^0 = \boldsymbol{f}_{DUN}(\boldsymbol{c}(\boldsymbol{f} = 0))$. Then, the algorithm proceeds by generating a succession of feasible link flow vectors $\boldsymbol{f} \in S_f(\boldsymbol{d})$ by solving a succession of linear problems which approximates the problem expressed by means of equation (2.29). Indeed, at a point $\boldsymbol{f} \in S_f(\boldsymbol{d})$, objective function $z(\boldsymbol{f})$ of problem (2.29) can be approximated with a linear function $\bar{z}(\boldsymbol{f})$ using Taylor's formula up to the first term:

$$z(\boldsymbol{f}) \cong z(\overline{\boldsymbol{f}}) + \nabla z(\overline{\boldsymbol{f}})^T \cdot \boldsymbol{f} = \bar{z}(\boldsymbol{f}) \tag{2.30}$$

By neglecting constant terms and considering that the integration and the derivation on the same variables are operations which cancel each other out, it is possible to obtain the following equation:

$$\bar{z}(\boldsymbol{f}) \cong \nabla z(\bar{\boldsymbol{f}})^T \cdot \boldsymbol{f} = c(\bar{\boldsymbol{f}}) \cdot \boldsymbol{f} \tag{2.31}$$

Hence, problem (2.29) can be expressed as follows:

$$\boldsymbol{f} = \underset{\boldsymbol{f} \in S_f}{argmin}\ c(\bar{\boldsymbol{f}})^T \cdot \boldsymbol{f} = \boldsymbol{f}_{DUN} \tag{2.32}$$

and, therefore, its solution is exactly a *DUN* flow vector, indicated as $\boldsymbol{f}_{DUN}^k$ at the generic iteration k, which coincides with a vertex of the polyhedron $S_f(\boldsymbol{d})$. In fact, such a method does nothing but resolve a problem of optimisation of a convex non-linear function on the $S_f(\boldsymbol{d})$ set defined by linear constraints. Hence, starting from the current solution, by means of the resolution of optimisation problem (2.32), that is none other than the calculation of a *DUN* link flow vectors as stated above, the algorithm is able to identify a direction along which the objective function is minimised so as to determine the new solution.

More in detail, by setting $k = 0$ at the beginning, the Frank-Wolfe algorithm for the calculation of *DUE* link flows with rigid demand and cost functions with symmetric Jacobian, can be described by the following system of recursive equations:

$$\boldsymbol{c}^k = \boldsymbol{c}(\boldsymbol{f}^{k-1}) \tag{2.33}$$

$$\boldsymbol{f}_{DUN} = \boldsymbol{f}_{DUN}(\boldsymbol{c}^k) \tag{2.34}$$

$$\mu^k = \underset{\mu \in [0,1]}{argmin}\ \psi(\mu) = z(\boldsymbol{f}^{k-1} + \mu(\boldsymbol{f}_{DUN}^k - \boldsymbol{f}^{k-1})) \tag{2.35}$$

In particular, the derivative $\frac{d\psi(\mu)}{d\mu}$ can be easily obtained from link costs:

$$\frac{d\boldsymbol{\psi}(\mu)}{d\mu} = \nabla z\left(\boldsymbol{f}^{k-1} + \mu \cdot \left(\boldsymbol{f}_{DUN}^{k} - \boldsymbol{f}^{k-1}\right)\right)^{T} \left(\boldsymbol{f}_{DUN}^{k} - \boldsymbol{f}^{k-1}\right) = \\ = c\left(\boldsymbol{f}^{k-1} + \mu \cdot \left(\boldsymbol{f}_{DUN}^{k} - \boldsymbol{f}^{k-1}\right)\right)^{T} \left(\boldsymbol{f}_{DUN}^{k} - \boldsymbol{f}^{k-1}\right) \tag{2.36}$$

It is worth pointing out that, since the calculation of the derivatives is replaced by the calculation of cost functions, it is possible to bypass the computation of function with a great simplification in the resolution procedure.

The termination test foresees that scalar product between the gradient and the direction of movement, made dimensionless, is lower than a pre-established threshold:

$$\frac{\left|\left(\boldsymbol{c}^{k}\right)^{T} \cdot \left(\boldsymbol{f}_{DUN}^{k} - \boldsymbol{f}^{k-1}\right)\right|}{\left|\left(\boldsymbol{c}^{k}\right)^{T} \cdot \boldsymbol{f}^{k-1}\right|} < \varepsilon \tag{2.37}$$

In the case of a non-symmetric Jacobian, it is necessary to rely on a *diagonalisation algorithm* (Florian and Spiess, 1982) whose convergence, however, cannot be strictly guaranteed on a mathematical basis.

Finally, *Stochastic User Equilibrium* assignment (2.18) is generally addressed as a fixed point problem for solving which it is necessary to adopt the *Method of Successive Averages -MSA* (Daganzo, 1983; Powell and Sheffi, 1982; Sheffi and Powell, 1981), whose convergence is assured by Blum's theorem (Blum, 1954).

There are two possible variations of such an algorithm which differ in the parameter which is mediated at each iteration, namely *MSA-FA* (*Flow Averaging*) and *MSA-CA* (*Cost Averaging*).

In particular, the *MSA-FA algorithm*, given $\boldsymbol{f}^{0} \in S_f$ and starting by $k = 0$, can be described by the following system of recursive equations:

$$k = k + 1 \tag{2.38}$$

$$\boldsymbol{c}^{k} = \boldsymbol{c}\left(\boldsymbol{f}^{k-1}\right) \tag{2.39}$$

$$\boldsymbol{f}_{SUN} = \boldsymbol{f}_{SUN}\left(\boldsymbol{c}^{k}\right) \tag{2.40}$$

$$\boldsymbol{f}^k = \boldsymbol{f}^{k-1} + \frac{1}{k}\left(\boldsymbol{f}_{SUN}^k - \boldsymbol{f}^{k-1}\right) \tag{2.41}$$

Therefore, at each iteration, a stochastic uncongested network assignment is carried out with costs corresponding to the current solution, and the solution, which is the average of the first *k SUN* assignments, is carried out.

On the other hand, *the MSA-CA algorithm*, given $\boldsymbol{f}^0 \in S_f$, $k = 0$, $\boldsymbol{c}^0 = \boldsymbol{c}\left(\boldsymbol{f}^0\right)$, can be described by the following system of recursive equations:

$$k = k + 1 \tag{2.42}$$

$$\boldsymbol{f}^k = \boldsymbol{f}_{SUN}\left(\boldsymbol{c}^{k-1}\right) \tag{2.43}$$

$$\boldsymbol{y}^k = \boldsymbol{c}\left(\boldsymbol{f}^k\right) \tag{2.44}$$

$$\boldsymbol{c}^k = \boldsymbol{c}^{k-1} + \frac{1}{k}\left(\boldsymbol{y}^k - \boldsymbol{c}^{k-1}\right) \tag{2.45}$$

In this case, it is worth noting that the link flow vector $\boldsymbol{f}^k$ at each iteration k is feasible.

The algorithm terminates if the *SUN* flows calculated with costs $\boldsymbol{y}^k$ are equal to the flows vector $\boldsymbol{f}^k : \boldsymbol{f}_{SUN}\left(\boldsymbol{c}\left(\boldsymbol{f}^k\right)\right) - \boldsymbol{f}^k$. This, in practice, corresponds to stop the procedure when the quantity $\dfrac{\left|\boldsymbol{f}_{SUN}\left(\boldsymbol{c}\left(\boldsymbol{f}^k\right)\right) - \boldsymbol{f}^k\right|}{\boldsymbol{f}^k}$ is below a pre-assigned threshold ε.

It is worth noting that the termination test is itself computationally demanding, since it requires an additional *SUN* assignment at each iteration.

In general, the convergence of the *MSA-CA algorithm* is slower with respect to the *MSA-FA algorithm*; however, the Cost Averaging method has the advantage that its convergence properties are guaranteed also in correspondence with non-separable cost functions.

Moreover, recently, D'Acierno et al. (2006) proposed a *MSA* algorithm based on an Ant Colony Optimisation (ACO) approach, where the average was applied to the pheromone trail.

A recap of the described algorithms is set out in table 2.2.

<table>
<tr><td colspan="2"></td><td colspan="2">PATH CHOICE MODEL</td></tr>
<tr><td colspan="2"></td><td>Deterministic</td><td>Stochastic</td></tr>
<tr><td colspan="2">Uncongested network</td><td>All-or-Nothing</td><td>Dial/ MonteCarlo</td></tr>
<tr><td rowspan="2">Congested network</td><td>Symmetric User Equilibrium</td><td>Frank-Wolfe</td><td>MSA-FA</td></tr>
<tr><td>Asymmetric User Equilibrium</td><td>Diagonalization</td><td>MSA-CA</td></tr>
</table>

Table 2.2 Assignment algorithms classification (source: Cascetta, 2009)

2.5 Optimisation models for transport systems

As shown by Gallo et al. (2011b), generally, the optimisation of transportation systems, under the assumption of rigid demand, can be formalised as follows:

$$\hat{\boldsymbol{y}} = \underset{\boldsymbol{y} \in S_y}{argmin}\ w(\boldsymbol{y}, \boldsymbol{f}^*) \tag{2.46}$$

subject to

$$\boldsymbol{y} \in S_y \tag{2.47}$$

$$\boldsymbol{f}^* = \psi(\boldsymbol{y}, \boldsymbol{f}^*, \boldsymbol{d}) \tag{2.48}$$

where $\boldsymbol{y}$ is the vector of decision variables to be optimised (e.g. transit line path, frequency, fares, travel times); $\hat{\boldsymbol{y}}$ is the optimal value of $\boldsymbol{y}$; S_y is the feasibility set of vector $\boldsymbol{y}$; $w(\cdot)$ is the objective function to be minimised; $\boldsymbol{f}^*$ is the vector of equilibrium flows; $\psi(\cdot)$ is the rigid demand assignment function; $\boldsymbol{d}$ is the travel demand vector.

Equation (2.47) summarises both technical and budget constraints on decision variables. For example, technical constraints can be related to the network features or to the planned timetable in the case of scheduled service; whereas, instances of budget constraints can be grants for an improving infrastructural intervention or the number of vehicle-km to be operated.

In addition, the assignment constraint (i.e. demand-supply consistency constraint) is represented by equation (2.48) which indicates the fixed-point problem arising in the case of a *SUE* assignment, as described in paragraph 2.1.6.

Under the assumption of elastic demand, the above-mentioned optimisation model becomes:

$$\hat{\boldsymbol{y}} = \underset{\boldsymbol{y} \in S_y}{argmin} \; w\left(\boldsymbol{y}, \boldsymbol{f}_m^*\right) \tag{2.49}$$

subject to

$$\boldsymbol{y} \in S_y \tag{2.50}$$

$$\boldsymbol{f}_m^* = \psi'\left(\boldsymbol{y}, \boldsymbol{f}_m^*, \boldsymbol{d}_m\left(\boldsymbol{y}, \boldsymbol{f}_m^*\right)\right) \tag{2.51}$$

where $\boldsymbol{f}_m^*$ is the multimodal equilibrium flow vector; ψ' represents the elastic demand assignment function; $\boldsymbol{d}_m(\cdot)$ is the multimodal demand vector.

In particular, multimodal equilibrium flow vector $\boldsymbol{f}_m^*$ is obtained by a joint estimation of equilibrium on the various transportation systems (car, rail, bus, etc.); while, the multimodal demand vector $\boldsymbol{d}_m(\cdot)$ arranges the transportation demand vectors for each transportation system. It depends on decisional variables $\boldsymbol{y}$ and on the multimodal equilibrium flows $\boldsymbol{f}_m^*$. The function that relates $\boldsymbol{y}$ and $\boldsymbol{f}_m^*$ to the multimodal demand vector is a mode choice model.

Besides constraints, other key elements to be defined in the above described optimisation frameworks, according to the specific problem to be addressed, are decision variables (e.g. waiting and travel times, frequency) and the objective function to be minimised (e.g. car and transit user costs, external costs).

2.6 Rail optimisation models

The issue of managing and optimising railway operations is addressed in the literature as *dispatching* and *rescheduling* problems, which consist in tasks of monitoring and controlling aimed at ensuring a smooth running of rail service, as well as re-establishing ordinary conditions, in response to any kind of system failure, by adjusting the planned service to the actual situations.

FACTORS OF COMPARISON	OFF-LINE TIMETABLING	REAL-TIME TRAFFIC MANAGEMENT
Main objective	Design optimal schedule	Implement optimal control
Schedule validity	Up to some years	Up to few perturbed hours
Degree of flexibility	Any change applicable	Minor timetable modifications
Traffic conditions	Usually ideal conditions	Perturbations or disruptions
Time span of prediction	Long time horizon	Up to some hours
Space span of prediction	Large traffic network	Rail junction or small network
Computation time	Up to several months	Up to few minutes

Table 2.3 Differences between off-line timetabling and real-time traffic management (source: D'Ariano, 2008)

In particular, there are two dimensions of interest: an off-line design phase and a real-time operational phase (D'Ariano, 2008). The first stage concerns the design of the railway timetable and the analysis of its stability; while, the second stage is related to the management of the service in real-time so as to properly react to system failure and provide an effective solution as rapidly as possible. Table 2.3 shows the main differences between the two phases.

	STATIC	DYNAMIC	
ON-LINE	Online static traffic rescheduling *Open Loop Control*	Reactive dynamic *Closed Loop Control*	Proactive dynamic *Closed Loop Control*
OFF-LINE	Train timetabling	—	

Table 2.4 On-line vs. off-line and static vs. dynamic approaches (source: Corman and Meng, 2015)

A further distinction, shown in table 2.4, has been proposed by Corman and Meng (2005), according to which rescheduling tasks can be performed statically or dynamically on the basis of the input information implemented, as opposed to the timetabling phase which is intrinsically an off-line and static process. In particular, in static methods, input data are processed only once with a fixed value; while, in dynamic approaches, the values of input parameters change over time. Moreover, dynamic rescheduling approaches can be distinguished into *reactive*, if they neglect a view of the future, and *proactive*, if they consider future conditions in a probabilistic and time-dependent way.

The close relationship between the two above-mentioned management dimensions is evident: a well-designed timetable, with a high degree of stability and robustness, makes the rescheduling process easier and smoother.

2.6.1. The timetabling phase

The *timetabling* process of a railway line consists in establishing the departure and arrival times of each convoy at each station being served, respecting the limits imposed by safety, law, infrastructure, signalling system and the necessity to guarantee a certain number of transfers. Such a phase is crucial for the entire railway operation as it influences, directly or indirectly, system performance, the degree of use of the infrastructure capacity, service quality, the management of rolling stock and the crew scheduling.

Generally, timetables are characterised by the adoption of cyclic structures, which can include particular properties namely periodicity, synchronisation and symmetry. The *periodicity* consists in setting regular intervals among trains during the whole service; the *synchronisation* regards the coordination of departure and arrival times of the planned runs in order to provide feasible transfers for passengers; the *symmetry* concerns the adoption of the same timetable features in both directions and, as shown by Liebchen (2004), it makes sense only if travel times and dwell times are the same in both ways and travel demand is symmetrical as well.

Periodic timetables are usually modelled by means of a *Periodic Event Scheduling Problem* (PESP), introduced by Serafini and Ukovich (1989). Contributions which adopt this approach are, for instance, Liebchen (2008), Liebchen and Möring (2002), Nachtigall (1996) and Peeters (2003). Moreover, Kroon et al. (2007) described a stochastic variant of the PESP which takes into account random disturbances to rail service; Wong and Leung (2004), and Wong et al. (2008) proposed a synchronisation model which minimises waiting times for passengers. Similarly, Guo et al. (2017) developed a timetable optimisation framework implementing the Particle Swarm Optimisation and Simulated Annealing for enhancing the performance of transfer synchronisation between different rail lines. Finally, an example of optimisation framework for symmetric timetables can be found in Bruglieri et al. (2017) whose approach duly takes into account modal split and travel demand.

Timetable performance measures are *reliability*, *punctuality* and *robustness*. The *reliability* is the ability of a system or a component to perform its required functions under stated conditions for a specified period of time (Rausand and Høyland, 2004); *punctuality* is usually defined as the probability that a train arrives less than *x* minutes late (Ceder and Hassold, 2015); *robustness* refers to the capability of avoiding delay propagation as much as possible (Cacchiani and Toth, 2012).

Generally, a robust timetable is carried out by properly introducing buffer times for absorbing potential delays. However, it is necessary to strike the right balance between the use of railway capacity and the robustness of the timetable (Barter, 2004; Carey and Kwiecinski, 1995; Landex et al., 2006; Wendler, 2001). In fact, with an increase in buffer times, the timetable presents a greater flexibility and, thus, an increased chance of absorbing delays, avoiding their spread; however, this could lead to an under-usage of system capacity.

In Goverde (2005) we can find an interesting design methodology for railway timetables, featured in figure 2.16, where two feedback cycles are proposed: one on the stability of the timetable (ex-ante analysis), and the other one on the punctuality of the system (ex-post analysis). As regards the stability analysis, the contribution extends to the railway case the methodology proposed by Baker (1993) and Subiono (2000), based on the Max-Plus Algebra, by introducing constraints dictated by the infrastructure and the signalling system. On the other hand, regarding the ex-post analysis, which requires the acquisition of measurements relative to the actual performed service, the author proposed a tool called TNV-Prepare.

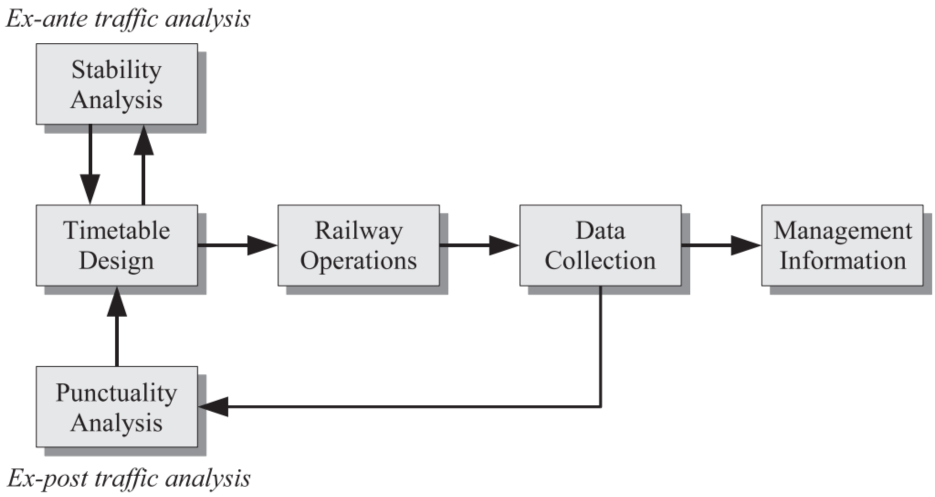


Figure 2.16 Feedback cycles in railway timetabling process (source: Goverde, 2005)

Moreover, Bešinović and Goverde (2016) proposed a two-stage model for carrying out robust timetables in which, after obtaining a stable timetable structure (i.e. a structure which minimises the trade-off between capacity utilization and travel times), the optimal allocation of time supplements and buffer times is derived. In addition, a delay propagation model is implemented for validating the obtained timetable. Similarly, Fischetti et al. (2009) developed a three-stage framework aimed at identifying robust timetable structures by means of a combination of linear programming with stochastic programming and robust optimisation techniques. In particular, firstly the Train Timetabling Problem (TTP) is modelled neglecting robustness; in the second step different training methods, which essentially test the impact on the system of the occurrence of delays, are implemented and, finally, a validation phase is performed. Furthermore, Yan and Goverde (2017) proposed an optimisation methodology for maximising timetable robustness in which the variability of dwell and travel times as well as the possibility of overtaking are considered. In addition, Sparing and Goverde (2017) improved the approach developed by Sparing and Goverde (2013) by proposing a method for generating periodic timetables aimed at maximising timetable stability indirectly, that is by optimising the cycle time. Indeed, as shown by Goverde (2007), a timetable can be stable only if the nominal timetable period is higher than the minimum cycle time; moreover, the degree of stability increases with the increase of the gap between these two quantities. Khadilkar (2017) developed a stochastic delay propagation model which evaluates timetable robustness by means of individual and collective measures, related respectively to primary and knock delays, and tested it on a portion of the Indian railway network. As to timetable performance addressed in the literature, it is worth citing the following contributions: Liebchen et al. (2009) which proposed an integrated timetabling/delay management framework by introducing a new concept of robustness, known as *recoverable robustness*, and Ciuffini (2014) which derived a method for comparing different timetable structures in terms of attractiveness for passengers, by computing the so called *time displacement* between what travellers desire and the scheduled service, whose formulation takes into account the frequency and the travellers' time adaptability.

Simulation-based approaches for performing the timetabling phase can be found in Bešinović et al. (2016) and Goverde et al. (2016). In particular, the former proposed an integrated framework which combines a micro and a macro network representation. More in deep, the timetable structure is carried out at the microscopic level, thanks to a very precise adjustment of running times and minimum headways; while, at the macroscopic level the trade-off between travel times and degree of robustness is performed. On the other hand, Goverde et al. (2016) proposed a design approach aimed at generating a robust and energy-efficient timetable by means of a three-stage process which combines different levels of analysis: microscopic, macroscopic, and a corridor fine-tuning level. The basic idea is to optimise each performance indicator at an appropriate level, so as to obtain a more reliable evaluation.

Clearly, bearing in mind the importance of the above-mentioned issues related to the stability and robustness of a timetable, different objective functions can be considered according to the examined contexts. For example, Canca et al. (2011) proposed a methodology to optimise the timetabling process so as to find the right balance between the quality of service and operational costs; while, Brännlund et al. (1998) introduced an optimisation problem in which the objective function to be maximised is the degree of use of the railway infrastructure. Moreover, Oliveira and Smith (2000), and Oliveira (2001) modelled the timetabling phase as a constrained job-scheduling problem, in which the objective function to be minimised is the total delay. In particular, the introduced restrictions are relative to travel demand and to the connections between runs, in order to guarantee a minimum number of transfers. Furthermore, the optimisation of the timetabling phase in an energy-efficient perspective can be found in Canca (2017) and Su et al. (2013). The close relationship among timetable, eco-driving profiles and energy saving strategies will be analysed in depth in paragraph 2.8.

The definition of a timetable involves different time rates such as train running times, blocking times and minimum headway between two successive convoys, dwell times at stations, buffer times and layover times. In particular, the following section is focused on the estimation techniques of dwell times, which have a key role in the timetable planning phase, especially in the case of congested lines, given their nature of flow-dependent factors.

2.6.1.1 Dwell times estimation techniques

One of the main consequences of the interaction between rail service and travel demand is the fact that dwell times cannot be derived as fixed values, but their estimation has to be carried out as function of the passenger flows involved in the boarding/alighting process. This requirement becomes increasingly felt in crowded situations where, as shown by Kanai et al. (2011), the phenomenon known as *snowball effect* takes place. In fact, according to the dynamic interaction between rail operations and traveller flows, the number of passengers on the platform influences the dwell times of trains at stations, which may cause increasing delays. Consequently, there occurs an increase in headways which could generate more passenger flow on the platform producing longer dwell times. However, the snowball effect does not evolve indefinitely, but tends to converge towards an equilibrium state according to proper theoretical conditions, as shown later in this work. Therefore, the importance of a suitable estimation of dwell times in order to ensure a high degree of timetable robustness, thus making the service more reliable and attractive in the eyes of passengers, appears clear (Carey and Carville, 2000; Dewilde et al., 2014; Hadas and Ceder, 2010; Heimburger et al., 1999).

Hence, estimating the number and characteristics of passengers (i.e. gender, age, mobility, luggage) is a key task for calculating the amount of time required for the boarding/alighting process (Daamen et al., 2008) and, therefore, for obtaining a reliable value of dwell time. Clearly, the matter is quite complex, since it involves some uncertain factors such as the interaction among different groups of passengers on the same platform and between passengers on the platform and those on-board (Puong, 2000). In addition, it is necessary to carry out also a forecast of queues due to lack of residual capacity on the convoy: this, in turn, may cause delays since it influences the subsequent alighting and boarding process. In this context, Xu et al (2013) proposed a tool for supporting pedestrian flow management which, by means of probabilistic theory and discrete time Markov chain theory, gives a theoretically quantitative prediction for the queue length of stranded passengers.

Other factors which may influence dwell time are rolling stock, station layout and rail operations. The relevance of rolling stock features lies in several factors: number and width of doors (Weston, 1989), kind of service performed (Jong and Chang, 2011), horizontal and vertical gaps between the train and the platform (Buchmuller et al., 2008; Wiggenraad, 2001), interior layout of the convoy, seen as number and position of seats (Harris, 2006) or as passenger distance to exit doors and potential free space which user is inclined to occupy (Baee et al., 2012). Strictly related to the internal layout of convoys, there is the fare collection method which, as shown by Fletcher and El-Geneidy (2013), may influence the time required for completing boarding and alighting process as well. Indeed, in the presence of a manual fare method, additional time is required due to the necessity of an interaction between the passenger and the driver, who has to select the proper ticket and eventually give the change. On the contrary, an automatic fare collection method could speed up the operation; however, in this case, the factors which may influence the boarding time are the number and position of stamping machines or fare boxes on-board, if any. Station features in terms of position of access/egress facilities (Kunimatsu et al.,2012) and train stop type, e.g. short stop or large stations (Li et al., 2014a), may affect dwell times as well. Regarding rail operations, in addition to the above mentioned reciprocal influence existing between rail service and travel demand, it is worth noting that dwell times have a large role also in containing the propagation of delays in order to avoid the arising of the so called secondary delays, which lead to a further deterioration of service quality (Büker and Seybold, 2012; Burdett and Kozan, 2014; Ceder and Hassold, 2015; Cui et al., 2016).

Basically, two kinds of estimation approach have been proposed in the literature for computing dwell times.

The first proposed contributions adopted statistical techniques, i.e. regression models (Lam et al., 1999; Wirasinghe and Szplett, 1984), which were borrowed from the bus service field (Guenthner and Hamat, 1998; Levine and Torng; 1994; Levinson, 1983). Specifically, regression methodologies are based on detected data and aim to express dwell times as a sum of constant and variable predictors. In particular, fixed values are related to the unlocking, opening and closing times of doors together with planned buffer times; in fact, they are invariant once rolling stock features and train dispatching times are set up. By contrast, variable parameters are function of the user alighting and boarding times which, in turn, depend on passenger flows. More recent contributions in this field are provided by Hansen et al. (2010), Harris and Anderson (2007), Kecman and Goverde (2015), and Vuchic (2005). However, broadly speaking, these models are too bound to the specific conditions in which they were developed for being applied to other contexts. Indeed, basically, they are descriptive models, since no details about passenger behavioural rules when a train arrives are considered and, thus, they can be useful only in the case of already existing rail services. Therefore, they have no predicted power and, hence, are not appropriate to be used in the planning phase. Finally, they are mostly deterministic, which means that their results can be viewed as the expected values of dwell times required for completing the boarding and alighting process, without any information about their statistical distribution.

On the other hand, the second approach proposed in the literature relies on micro-simulation tasks which are able to explicitly model pedestrian behaviour on platforms, especially in crowded conditions (Lam et al., 1998; Tirachini et al., 2013), and relate it to delays and to other aspects of rail service performance. These methods overtake the inconveniences of regression models, since they can be used also in a planned stage for modelling hypothetical contexts and, generally, are able to take into account the stochasticity of the phenomenon under observation, which may be due to several factors such as temporal and spatial distribution of travel demand, train delays, passenger and train driver behaviours. Stochastic variations in dwell time are modelled in Larsen et al. (2014); while, Longo and Medeossi (2013) computed dwell times by splitting the estimation procedure into both deterministic and stochastic sub-processes. Other micro-simulation approaches are proposed by Jiang et al. (2015) which extended the simulation methodology suggested by Jiang (2012) introducing the evaluation of mutual interactions between dwell times and train delays. By contrast, Zhang et al. (2008) simulated the cooperation and negotiation process between boarding and alighting passengers by means of a cellular automata-based model and Yamamura et al. (2013) proposed a multi-agent simulation method which is able to take into account passenger congestion both on platform and inside trains. Moreover, in order to overcome the drawbacks of discrete approaches, Bandini et al. (2014) proposed an innovative floor-field cellular automata pedestrian model which is specifically developed for simulating high-density contexts. Furthermore, Seriani and Fernandez (2015) addressed the problem of defining passenger service time and other related factors (i.e. user density on trains and platforms, pedestrian level of service and passenger dissatisfaction) by combining micro-simulation tasks with laboratory experiments.

Other methodologies proposed in the literature for estimating dwell times are based on the use of artificial neural networks: Berbey et al. (2012) modelled human behaviour and interactions among different groups of passengers by combining artificial intelligence-based techniques and a fuzzy logic approach; while, Chu et al. (2015) addressed the problem by using the so-called *Extreme Learning Machine* (ELM), a very fast training speed algorithm described by Huang et al. (2006).

Given the close relationship among dwell times, timetable and reliability of rail service (stated, for instance, by Pouryousef and Lautala, 2015; Sato et al., 2013), together with the need to evaluate boarding and alighting times as a function of passenger flows, the majority of contributions in the literature are focused on dwell time estimation in the planning phase: Wong at al. (2008) addressed the definition of running times and station dwell times in order to minimise transfer waiting times; while, Landex and Jensen (2013) analysed the possibility of adjusting dwell times so as to increase station capacity. Nevertheless, some have also proposed models for managing disruptions (Kepaptsoglou and Karlaftis, 2010) and real-time rescheduling tasks (Li et al., 2016), or even for putting in place effective energy saving measures (Xiaoming et al., 2016).

Clearly, in the case of existing services, mass transit agencies may adopt a statistical approach for determining dwell times as a function of a certain confidence level (expressed, for instance, in terms of percentiles) and perform some control strategies for increasing the compliance between the real and the scheduled timetables. However, this work is focused on the estimation of dwell times in planning stages (i.e. when the service is not yet in operation), during which it is necessary to rely on suitable modelling approaches for simulating passenger behaviour on platform when a train arrives in a very accurate and realistic manner.

2.6.2. The rescheduling problem

The *rescheduling* problem covers a large part of the railway operations research, since the advantages offered by rail transport, in terms of high travel speed and low values of headway (due to exclusive lanes, constrained drive and signalling system), are counterbalanced by an intrinsic fragility to failure phenomena. However, very frequently, dispatchers can count only on their experience (e.g. by presuming the amount of recovery times or the most successful intervention strategy) and, therefore, developing suitable decision support systems for helping them to deal with disruption conditions turns out to be fundamental.

Generally, recovery strategies are implemented according to three consecutive phases: timetable rescheduling, rolling stock rescheduling and crew rescheduling; however, what follows is essentially focused on timetable rescheduling.

As shown by Hansen e Pachl (2008), the rescheduling process consists in two successive steps. The initial phase concerns the identification of potential conflicts on the basis of the current state of the infrastructure, the characteristics of operational times, the availability of rolling stock, the position and travel speed of each convoy. This is followed by a problem solving phase which, according to the results of the previous step and the delays actually occurred, identifies the most appropriate strategies for re-establishing normal operating conditions.

In the following, a classification of rescheduling methodologies proposed in the literature, according to different criteria, will be provided.

Frequently, rescheduling problems are addressed by means of simulation-based methods and, therefore, railway optimisation models, similarly to simulation ones, can be classified into macroscopic and microscopic, according to the degree of detail implemented. Moreover, two main modelling approaches are generally adopted which are based on the implementation respectively of the so called *Alternative Graph (AG)* and *Mixed Integer Linear Programming (MILP)* formulations.

The *Alternative Graph* model was proposed by Mascis and Pacciarelli (2002), as a generalisation of the disjunctive graph formulation of Roy and Sussman (1964). Essentially, it allows to simulate railway operations as a job-shop scheduling problem, i.e. the problem of allocating machines to competing jobs over time, subject to the constraint that each machine can handle at most one job at a time. Therefore, each operation denotes the traversal of a resource (block/track section or station platform) by a job (train route). In particular, Mascis and Pacciarelli (2002) introduced additional constraints, known as blocking and no-wait constraints, modelling respectively the absence of storage capacity among machines and the condition in which two consecutive operations in a job must be processed without any interruption. Several rescheduling methodologies based on the implementation of the Alternative Graph, together with the blocking time theory (Hansen and Pachl, 2008), have been proposed in the literature for dealing with different dispatching problems. A first disruption management method based on this approach can be found in D'Ariano et al. (2008) which developed a decision support system for real-time traffic management, named ROMA (Railway traffic Optimisation by Means of Alternative graphs). In particular, such a tool solves the real-time train dispatching problem by subdividing it into four sub-problems:

- data loading and exchange of information with the field;
- assigning a passable route to each train in order to avoid blocked tracks;

- defining optimal train routes, ordering and specifying the exact arrival and departure times at stations and at a set of relevant points in the network;
- ensuring a minimum distance headway between trains while maintaining acceptable speed profiles.

In Quaglietta et al. (2013) ROMA was integrated with the microscopic traffic simulator EGTRAIN, so as to incorporate the dynamic evolution of traffic conditions into the dispatching procedure.

Other rescheduling approaches based on the adoption of the Alternative Graph concern: delay management problems (D'Ariano and Pranzo 2009); rerouting recovery actions (Corman et al. 2011b, Flamini and Pacciarelli 2008, Samà et al. 2017) and conflict resolution tasks (Corman et al. 2012). Also large network contexts and very severe disruption conditions can be addressed by means of such an approach (Corman et al 2011a). Moreover, formulations of the Alternative Graph targeting for dealing with disruptions conditions in rail lines with moving-block signalling systems (i.e. the headway is computed as a minimum time lag on each section for two consecutive trains) have been developed by Mazzarello and Ottaviani (2007), and Xu et al. (2017a; 2017b).

While rescheduling methods based on the Alternative Graph generally adopt a microscopic approach, works implementing Mixed Integer Linear Programming (MILP) formulations proposed in the literature deal with both microscopic (Boccia et al., 2013; Hirai et al., 2009; Pellegrini et al., 2012) and macroscopic (Acuna-Agost et al., 2011; Dundar and Şahin, 2013; Louwerse and Huisman, 2014; Min et al., 2011; Narayanaswami and Rangaraj, 2013; Shöbel, 2007; Törnquist and Persson, 2007) frameworks.

Moreover, Shakibayifar et al. (2017) proposed a real time recovery management model, for dealing with multiple disruptions, which adopts heuristic dispatching rules and integrates different intervention strategies such as reordering, retiming, speed control and dwell time adjustment. Meng and Zhou (2014) developed an integer programming model characterised by an innovative formulation with network-based cumulative flow variables for addressing a simultaneous train rerouting and rescheduling problem. Zhan et al. (2015) formulated a Mixed Integer Programming model, for handling a complete blockage disruption on high speed lines, whose aim is to minimise the total weighted train delay and the number of cancelled trains, in accordance with headway and station capacity constraints. Finally, Huo et al. (2016) addressed the timetable rescheduling problem by developing a binary mixed-integer programming model aimed at minimising the time difference between the planned timetable and the rescheduling one which is expressed in terms of train order entropy.

The main advantage offered by macro approaches lies in the lower computational effort which, for example, allows to deal with complex objective functions, like in Binder et al. (2017) where a macroscopic multi-objective framework, taking into account passenger satisfaction, operational costs and deviations from the undisrupted timetable, is proposed. On the other hand, micro-simulation approaches, as already pointed out, allow to explicitly model the interactions among system components (i.e. infrastructure, signalling system, rolling stock, timetable and travel demand) and compute involved quantities in an accurate manner (e.g. running times, dwell times, headways). Therefore, in order to benefit from advantages of both approaches, also integrated frameworks which combines these two simulation techniques have been proposed in the literature. In this context, Placido et al. (2014a) proposed a rescheduling method including both a macroscopic and a microscopic model of the network. In particular, the macroscopic representation is implemented in an optimisation framework, based on the model developed by Cadarso et al. (2013), whose aim is to derive timetable and rolling stock schedule in the case of failure. On the other hand, the microscopic representation is used for the simulation model, which is based on the proposal of D'Acierno et al. (2013a), whose structure includes the Service Simulation Model (SSM) and the On-Platform Model (OPM) for assigning travel demand to the rail network. Dollovoet et al. (2014) developed an iterative optimisation framework in which a delay management problem is solved macroscopically and, then, validated microscopically by means of a train scheduling model taking into account the limited capacity of stations. Specifically, the original timetable and travel demand flows are given as initial input data, together with a set of delays computed on arrival events. With this information, the algorithm solves the delay management problem by identifying the connections to be maintained and carrying out an expected macroscopic timetable. Then, the output of the delay management problem becomes the input of the train scheduling problem, whose resolution consists in analysing potential conflicts around stations and estimating delay propagation. The fact that these delays are computed by means of a microscopic approach ensures an accurate degree of estimation; hence, they are, in turn, implemented in the delay management problem which is run again and, at the end of the iterative process, a timetable minimising passenger delays is carried out.

The above-mentioned works adopt a synchronous approach, since the aim of the analysis is to address events like deviations from the planned service, propagation of delays and system failures. On the other hand, the asynchronous approach is generally implemented for solving conflicts between trains belonging to different categories, by always giving priority to trains in a higher category (see, for instance, Jacobs 2004). However, clearly, asynchronous solving conflict algorithms cannot guarantee a global optimum as solution.

Finally, albeit in the literature it is possible to find some deterministic rescheduling methodologies (see, for instance, D'Acierno et al., 2013b; Ho and Yeung, 2001; Schöbel, 2007), the stochastic approach is the most accurate, given the random nature of the involved factors. In particular, the importance of taking into account the stochasticity of events lies in the fact that the stability of rail service is very sensitive to the presence of even small variations in the performance of convoys or dwell times, above all for the risk of a knock-on effect of propagation of delays which would negatively affect the entire system. To this end, Hansen (2006) described the influence on system performance of the stochasticity of design variables within the railway timetable. In this context, Yuan (2006) proposed a probabilistic analytical model which makes a realistic estimate of delay propagation and provides an assessment of delay impact on the punctuality of the service. Conte and Shöbel (2007) developed a stochastic simplified graphical modelling approach, for identifying dependencies among delays, which is based on the so called Tri-graph (proposed by Wille and Bühlmann, 2004; 2006) allowing a compact representation of different kinds of delay: primary delays, secondary delays (due to the propagation of primary delays) and delays due to the restricted capacity of the railway infrastructure. The relevance of considering delays as time-dependent random variables is stated also by Kecman et al. (2015a) and Kecman et al. (2015b) which modelled the uncertainty of train delays respectively by means of a Markov stochastic process and Bayesian networks. While, stochasticity of arrival and recovery times is taken into account in the rescheduling models proposed by Davydov et al. (2017), and Li et al. (2014b). Moreover, Larsen et al. (2014) analysed the impact of considering uncertainty in the rescheduling framework by comparing results of different algorithms, both in deterministic and stochastic scenarios. In particular, train delays are modelled by means of a statistical distribution, while running and dwell times are perturbed with stochastic variations. Similarly, stochasticity of train performance and dwell times are modelled in D'Acierno et al. (2016a). In addition, the uncertainty of the disruption information are addressed by Meng and Zhou (2011) which developed a stochastic and dynamic rescheduling model aimed at minimising the total train delay in the case of a single-track rail line. More in deep, the proposed approach is implemented in a rolling horizon framework: the robustness of rescheduling strategies is evaluated considering random segment running times and a segment capacity breakdown with an uncertain duration. Finally, Yin et al. (2016) developed a metro rescheduling model which takes into account the stochasticity of travel demand: the arriving ratio of passengers at each station is modelled as a non-homogeneous Poisson distribution in which the intensity function is treated by means of time-varying origin-destination matrices.

In rescheduling problems, two fundamental issues have to be taken into account, which are strictly related to each other: on one side, the interaction between rail operations and travel demand and, on the other side, capacity constraints of rail service. In particular, the interface between rail operations and passenger flows is represented by the boarding and alighting process which is obviously affected by the available capacity. For the sake of clarity, it is worth noting that the fact of addressing the problem by taking into account the influence of travel demand on the service is aimed at making the simulation as more realistic as possible, disregarding the final purpose of the analysis (i.e. whether or not the final aim is to satisfy passenger needs). However, issues related to the impact of travel demand on rail service and the minimisation of passengers discomfort are generally addressed together, due to their strict relationship. Indeed, boarding, alighting and on-board flows affect the performance of rail service and, therefore, their attractiveness, which in turn affect passenger satisfaction. Hence, a realistic modelling of boarding and alighting process allows a more accurate estimation of passenger inconvenience, for example in terms of waiting times for users on platform or in terms of total travel times for users on-board. Rescheduling methodologies which fulfil these requirements can be found in Kepaptsoglou and Karlaftis (2010), which dealt with post-disruption operations at station-platform level and D'Acierno et al. (2012) which introduced capacity constraints for taking into account the fact that, especially in crowded contexts, not all passengers waiting on the platform are actually able to board the first arriving train. Furthermore, Xu et al. (2017c) developed a rescheduling framework for minimising delay time of alighting passengers and penalty time of stranded passengers. Zhu and Goverde (2017) developed a dynamic passenger assignment model which implements an event-based simulation technique for modelling alighting and boarding process. In particular, passengers' en-route travel decisions are considered and all phases occurring during a disruption event (i.e. the first transition phase from the planned timetable to the disruption timetable, the second phase where the disruption timetable is performed and the third recovery phase from the disruption timetable to the planned timetable) are modelled. This is a very relevant point, since passengers who start their trip in different phases, generally, are affected by the disruption in a different manner.

Moreover, time-variability of travel demand, disruption-induced service changes and capacity constraints of convoys are explicitly taken into account.

More in general, the dynamic interaction between rail service and travel demand is considered in the following contributions. Gao et al. (2016) proposed a disruption management approach, in the case of a metro system, based on a skip-stop pattern, which involves the analysis of time-dependent passenger flows under conditions of limited train capacity. Canca et al. (2012) developed a model for analysing short-turning and deadheading rescheduling solutions which takes into account the dynamic behaviour of travel demand along the considered planning horizon and aims at minimising passenger overload and improving service quality. Finally, Veelenturf et al. (2017) proposed a macroscopic rescheduling approach which combines rolling stock and timetable recovery strategies by considering adjustments of stopping patterns in a passenger-oriented perspective. In particular, the adopted resolution method is a greedy technique based on the passenger flow simulation algorithm proposed by Kroon et al. (2015).

Moreover, disruption management problems may concern metro (Bizhan and Mohammad, 2015; Gao et al., 2017) regional (Adenso-Diaz et al., 1999; Botte et al., 2017) or high-speed services (Wang et al., 2012; Zhan et al., 2016). Furthermore, different degrees of network complexity can be addressed. In particular, the level of complexity increases moving from a single-track case (Meng and Zhou, 2011) to a N-track context (Meng and Zhou, 2014) and from a single line (Xu et al. 2016a) to a large network (Corman et al., 2010a; D'Ariano et al., 2016; Kecman et al., 2013). Also networks characterised by a mixed traffic can be analysed, with a further increase in the degree of complexity tackled. For example, Corman et al. (2011c) developed an on-line rescheduling model for dealing with different types of train categories (both for passengers and freight) having different priority rules.

Another classification criterion for rescheduling approach is the analysed failure severity. Indeed, as shown by Cacchiani et al. (2014), it is possible to distinguish between disturbance and disruption: disturbances are generally considered as small perturbations influencing the system; while, disruptions indicate large external incidents which can lead to the cancellation of runs within the timetable or even to the interruption of the whole service. Clearly, the greater the severity of the failure, the greater the impact of the corrective measures to be adopted. For example, Dollevoet et al. (2012b) dealt with the problem of connection and re-routing in the case of a delay occurrence; similarly, Bauer and Schöbel (2014) developed a learning-strategy for the on-line delay management problem. On the other hand, more severe perturbations are addressed by Corman et al. (2010b) and Veelenturf et al. (2016). In particular, Corman et al. (2010b) analysed a serious disruption where some block sections have a reduced maximum speed, together with others which are totally unavailable for traffic, by implementing the alternative graph; while, Veelenturf et al. (2016) developed a macroscopic rescheduling approach for handling cyclic timetables, in presence of large scale disruptions, which is based on an Integer Linear Programming (ILP) formulation taking into account infrastructure and rolling stock capacity constraints. Moreover, Ghaemi et al. (2016) presented a macroscopic rescheduling model to compute the disruption timetable for a complete blockage with a focus on short-turning trains. Partial and complete blockages are also addressed in Louwerse and Huisman (2014), which developed integer programming formulations for maximising service quality and tested them on case-studies from Netherlands Railways.

Finally, different perspectives can be introduced in rescheduling models. Firstly, as already mentioned, several works proposed passenger-oriented methodologies. In addition to the already cited contributions, other passenger-centric approaches can be found in Binder et al. (2015), Kanai et al. (2011), Kumazawa et al. (2010), Placido et al. (2014b), Sato et al. (2013), Tanaka et al. (2009), Toletti and Weidman (2016). Typical measures of service quality used for determining passengers satisfaction resulting from rescheduling strategies are: cumulative delays, waiting times, user generalised costs, removed connections, penalty time of stranded passengers. Obviously, passengers are not the only players in the rescheduling process. Indeed, the other parties involved are infrastructure managers and train operating companies. On one hand, train operating companies are interested in minimising both passenger discomfort and operational costs associated to the implemented rescheduling strategies. On the other hand, infrastructure managers aim to reduce train delays, even if this implies cancelling runs or suppressing connection services. Works which, in addition to passenger needs, considered operational costs of train companies are those proposed by Binder et al. (2017) and D'Acierno et al. (2016b). In this context, it is worth citing also the contribution of Cadarso et al. (2015) which computed different measures of cost resulting from the disruption management process, such as total operational cost for passengers services, total operational cost for empty movements and total number of schedule changes (i.e. services, compositions and inventory train changes), as indicators of the effort made by rail companies for putting in place recovery strategies. Moreover, the trade-off between the targets pursued by the two above-mentioned stakeholders (i.e. infrastructure managers and train operating companies) is addressed in Corman et al. (2012; 2015) and D'Ariano et al. (2017).

In addition, since the reduction in energy consumption is one of the main goals of railway companies, optimisation methods which adopt an energy saving view have been also proposed. However, to be precise, energy issues are generally taken into account in the case of scheduling frameworks, such as timetabling optimisation methods and real-time control strategies (see, for instance, Albrecht and Oettich, 2002; Canca and Zarzo, 2017; Corman et al., 2009; Feng et al. 2017). Furthermore, Chevrier et al. (2013), and Yin et al. (2017) analysed the trade-off between passenger needs and energy-efficiency in the case of scheduling approaches. On the other hand, rescheduling approaches involving passenger services proposed in the literature, usually neglect energy saving perspectives. By contrast, this target is very felt in disruption management approaches in the case of freight trains (see, for instance, Toletti et al., 2016; Umiliacchi et al., 2016).

In the light of the above, it is understandable that the rescheduling problem is strongly NP-HARD and, therefore, for its resolution, it is necessary to rely on proper heuristic and metaheuristic methods which are able to find sub-optimal solutions within suitable computation times. An overview of such optimisation techniques will be provided in the following section.

2.7 Optimisation algorithms

The conceptually simplest technique for identifying the optimal solution in a combinatorial optimisation problem is based on the enumeration methods which evaluate all candidate solutions (*exhaustive approach* or *brute force search*), or select a set of efficient solutions (*implicit enumeration approach*), and choose the one which optimises specific criteria expressed by an objective function, to be minimised or maximised according to the specific addressed issue. Their computational cost depends on the number of candidate solutions and, therefore, they are typically used in problems of limited dimensions (small-size problems). On the other hand, in the case of real-scale networks where, generally, the number of feasible solutions to be analysed is very high and the objective functions are not convex, it is necessary to rely on suitable metaheuristic techniques which afford the possibility of finding near-to-optimal solutions within reasonable computation times. What follows, far from any claim of being exhaustive, provides some basic principles of the most frequently used metaheuristic algorithms in the field of rail transport, ranging from design problems to those of scheduling and routing.

Let us begin with the analysis of a series of algorithms belonging to the class of *Local Search methods,* whose common framework consists in starting from an initial feasible solution, trying iteratively to improve the current solution by means of more or less complex modifications (e.g. the exchange of elements belonging or not to the solution) and drawing to a close when no further improvements can be made.

Specifically, the following techniques will be described:

- Neighbourhood Search
- Heuristic Local Search
- Tabu Search
- Simulating Annealing

The *Neighbourhood Search Algorithm (NSA)* is a heuristic algorithm for solving discrete optimisation problems. Each vector $\boldsymbol{y}$ has an associated set of vectors $N(\boldsymbol{y}) \subseteq S_y$, called neighbourhood of $\boldsymbol{y}$, where the generic element $\boldsymbol{y} \in N(\boldsymbol{y})$ is obtained from solution $\boldsymbol{y}$ by an operation consisting in modifying only one component of vector $\boldsymbol{y}$. This algorithm can be implemented according to two different approaches: *Steepest Descent Method (SDM),* consisting in examining all elements of the neighbourhood and identifying the best solution (i.e. the solution with the best objective function value), and *Random Descent Method (RDM)* consisting in randomly extracting a solution from the neighbourhood and comparing it with the current one. In particular, if the new solution is better than the current one, it then becomes the current solution; otherwise, another solution is randomly extracted until the neighbourhood runs out, since all solutions inside have been explored. This algorithm is relatively simple, but its importance lies in the fact that, in many cases, it is implemented as a sub-routine in more complex techniques, such as the Heuristic Local Search approach, set out below.

The *Heuristic Local Search (HLS)* is made up of five phases which combine unconstrained optimisation steps with constrained ones.

More in detail, as shown by Gallo et al. (2011b), it can be outlined in the following steps:

1. Unconstrained Mono-Dimensional Optimisation (UMDO);
2. Unconstrained Starting Solution definition (USS);
3. Unconstrained Neighbourhood Search Optimisation (UNSO);
4. Constrained Starting Solution definition (CSS);
5. Constrained Neighbourhood Search Optimisation (CNSO).

In the first phase, each component of vector $\boldsymbol{y}$ is optimised, assuming the values of other components as constant. This phase may be addressed according to an exhaustive or a mono-dimensional NSA approach which is carried out by neglecting involved constraints. The second phase entails determining the first starting solution by setting each component of vector $\boldsymbol{y}$ at the optimal value obtained by the previous phase. In the third phase, involved constraints are neglected as well and an *NSA* approach is performed. In this phase, it is possible to rely on both *SDM* and *RDM* techniques. The fourth phase analyses all the solutions generated in the previous phases, selecting the one which optimises the objective function and, jointly, satisfies constraints. Finally, the last phase performs the *NSA* by considering involved constraints. In this case, the *NSA* technique is implemented by means of an *SDM* approach.

Similarly to the *NSA*, this algorithm, in many cases, is performed as sub-routine of more articulate metaheuristic procedures, as we will shortly see.

Within this framework, Gallo et al. (2010; 2012) developed metaheuristic procedures for solving the network design problem, respectively, in urban and regional contexts. Moreover, Gallo et al. (2011a) proposed a multimodal approach for bus frequency design, then improved in the case of rail frequencies in Gallo et. al. (2011b). In addition, Hassannayebi et al. (2016) proposed a *Variable Neighbourhood Search (VNS)* algorithm for minimising the average passenger waiting time in the case of a partial line blockage and Samà et al. (2017a) implemented the same optimisation technique for addressing the problem of train scheduling and routing under disruption conditions. Moreover, Canca et al. (2017) developed an *Adaptive Large Neighbourhood Search (ALNS)* algorithm as resolution method for a complex problem which involves both network design and line planning issues. Finally, De Los Santos et al. (2017) addressed a frequency optimisation problem, in a cost-oriented perspective, by comparing a heuristic local search algorithm with three different optimisation techniques: a Mixed Integer Linear Programming (MILP) model, a MIP-based iterative algorithm and a shortest-path based algorithm. In particular, travel demand and competition among modes are taken into account and numerical results, both on test networks and over a real context, show that heuristic local search provides the best compromise between computational effort and solution quality.

Tabu Search (TS) algorithm is a deterministic method proposed by Glover (1986) and formalised by Glover (1989; 1990). Basically, it is a search approach whose peculiar feature, as the name itself implies, consists in making prohibited, namely tabu, the opposite of the ultimate move carried out, in order to avoid going back to previously-visited solutions. In particular, this method is based on the use of a memory structure, known as *tabu list*, which can adopt a short, intermediate or long-term memory criterion. However, in order to avoid that the search gets trapped at a local minimum, an aspiration criterion, generally based on the objective function values, is set up. It states that the solution accessible by means of a forbidden move can be accepted if no improving moves are available outside the tabu list. Clearly, at each iteration, it is necessary to update the tabu list, generally by means of a *FIFO* approach: the move entering is the opposite of the ultimate action carried out and the move exiting is the one which has remained on the list for the longest time. Obviously, there are many variations which enrich this basic version, for instance by considering the frequency with which certain types of solution have been analysed or by introducing random elements.

In this context, Ho and Yeung (2001) addressed the problem of train conflict detection and resolution in real time, by performing a Tabu Search optimisation and comparing its performance with those of other heuristic methods with different neighbourhood definitions. Corman et al. (2010b) dealt with the same problem, by implementing a Tabu Search technique in the real-time traffic management system ROMA which is based on the alternative graph model (Mascis and Pacciarelli, 2002). Moreover, in Corman et al. (2010b), similarly to the previously cited contribution, different neighbourhood structures are assessed and the results are compared with those obtained by D'Ariano et al.

(2007; 2008) which implemented, respectively, a branch and bound algorithm and a local search method. Silvestrin and Ritt (2017) proposed a methodology for solving a particular vehicle routing problem which deals with vehicles with multiple compartments. The suggested procedure can be considered as an iterative local search method where the implemented local search technique is a Tabu Search algorithm. More in detail, the starting point is a local minimum obtained by applying any local search; then, at each iteration, the current local minimum is randomly perturbed and the Tabu Search is implemented in order to move on to another local minimum. The stopping criterion is based on the number of consecutive iterations which provide an improvement on the incumbent solution. Dewildea et al. (2014) described an optimisation procedure for increasing the robustness around large railway stations, which may represent a bottleneck for the whole system, based on the investigation of the interaction between the routing and scheduling of trains in the vicinity of the analysed area. Therefore, a route choice module and a timetabling module are implemented and the timetabling problem is addressed by means of a Tabu Search algorithm, whose aim is to increase the smallest minimum time span between two trains so as to improve the reliability of railway operations. Moreover, a simulation module is introduced to evaluate the achieved performance in the examined region.

Simulating Annealing (SA) is a stochastic metaheuristic method proposed as optimisation technique for the first time by Kirkpatrick et al. (1983). It is inspired by the process of annealing in metallurgy, i.e. a process by which a solid is firstly brought to the fluid state, by means of heating to high temperatures, and then brought back to a solid or crystalline form by gradually reducing the temperature. At a high temperature, the atoms are in highly disordered state and so there is a high level of energy in the system. In order to bring such atoms to a highly ordered (statistically) crystalline state, the temperature must be lowered. However, a fast reduction can cause flaws in the crystalline grid with consequent fissuring and fracturing of the grid itself (thermal stress). Annealing proceeds to a gradual cooling of the system, precisely in order to avoid this phenomenon. Although, in general, the solid is inclined to turn out into states with a lower level of energy, there is a slight chance that it increases its energy. This probability depends on the temperature and the variations of energy level associated with the transformation between the two states. In particular, it is regulated by the *Metropolis criterion* (Metropolis et. al., 1953) according to which the probability of transformation increases with the increasing in the temperature and the decreasing in the energy gap. It is this very criterion which determines if the solution being studied can become the new current solution or not. More precisely, the analogy between the physical system and the optimisation method is based on the following correspondences: the states of the physical system correspond to the solutions of the problem; the position of the particles corresponds to the value of decisional variables; the energy level related to a certain state corresponds to the value of the objective function which is associated with a certain solution. While, the temperature has no a direct analogy, but it represents a control parameter which implicitly defines the region of the state space being explored by the algorithm in a particular phase. At high temperatures, since bad solutions are easily accepted, the *SA* algorithm can cross almost all the state space. Following on, by lowering the value of the control parameter, the algorithm is confined to increasingly restricted regions of the state space. Therefore, it can be stated that, at high temperatures, the algorithm behaves more or less as a random search; while, at low temperatures, the *SA* is similar to the steepest descent methods. The algorithm stops when the temperature value needed to terminate the annealing process is reached and, hence, there are no further possibilities for improvement in terms of objective function.

This method has been implemented for solving many different transportation problems such as: minimising timetable cycle time (Burkolter, 2005), finding the optimum stop-skipping patterns in urban railway systems under uncertainty (Jamili and Pourseyed Aghaee, 2015), solving conflicts in railway traffic under disruption conditions (Törnquist and Persson, 2005), optimising energy consumption in train operations (Kim and Chien, 2011). Moreover, it has been adopted in the case of rail-car fleet sizing problem (Sayarshad and Ghoseiri, 2009), railway crew scheduling problem (Hanafi and Kozan, 2014), track allocation problem at railway stations (Wu et al., 2013), train platform problem (Kang et al., 2012), train transfer problem (Kang and Zhu, 2016), transit network optimisation problem (Zhao and Zeng, 2006), bottleneck routing problem at railway stations (Wu et al., 2013) and location routing problem with simultaneous pickup and delivery (Yu and Lin, 2014).

Regarding the evolutionary techniques, *Scatter Search (SS)* methods (Goldberg, 1989; Holland, 1975; Srinivas and Patnaik, 1994) and the *Genetic Algorithm (GA)* (Glover et al., 2003; Laguna, 2002) are addressed in the following. In particular, differently from the case of *Local Search* which is characterised by a neighbourhood-based approach, evolutionary procedures are population-based problem solvers and are inspired by principles of biological evolution.

As shown by Martì et al. (2003), the basic framework of the *Scatter Search* can be described as the sum of the following five methods.

1. *Diversification Generation Method* aimed at generating a collection of diverse trial solutions starting off from a seed solution. It is important that the generated trial set is characterised by a high variety of different solutions, so as to cover different parts of the solution space.
2. *An Improvement Method* represented by an algorithmic subroutine (e.g. *NSA* or *HLSA*) aimed at transforming a trial solution into one or more enhanced trial solutions. However, there is no guarantee of improvements and, hence, if no enhancing is possible, the improved solutions are considered to be the same which have been generated in the previous phase.
3. *A Reference Set Update Method* aimed at carrying out a reference set by selecting all the enhanced solutions or only a part of them, taking into account their quality, according to objective function values (good solutions), and their diversity in terms of distance from the best solution (scattered-solutions). By including scattered solutions in the reference

set, the algorithm is empowered to explore regions which, otherwise, would remain unexplored.

4. *A Subset Generation Method* aimed at manipulating the reference set, in order to produce a subset of its solutions as a basis for creating combined solutions.
5. *A Solution Combination Method* aimed at obtaining one or more combined solution vectors from the subset of solutions generated in the previous phase.

The output of the fifth phase is then improved, as described in the second step, so as to create a new reference set and so on. The procedure stops when the reference sets in two successive iterations are equal or when a pre-fixed number of iterations is reached.

By moving to the *Genetic Algorithms*, the evolutionary rationale is more evident, starting from the terminology adopted: each solution is indicated as a chromosome and each solution component as a gene. This method can be summarised in the following phases: *initialisation, selection, reproduction and termination*. The first phase is aimed at generating a set of initial solutions which represents the initial population. In the second phase, two fundamental tasks are carried out. First of all, for each chromosome, the objective function and the related fitness function are calculated and, then, on the basis of these values, the parent selection is performed. It consists in extracting the best solutions from a population so as to enable them to successfully pass on their genes to the next generation. This step can make use of many different techniques such as roulette wheel selection, fitness proportionate selection, rank selection, random selection, tournament selection and stochastic universal sampling. Therefore, once two elements have been selected as parents, the reproduction phase is performed by means of two processes: crossover and mutation. The former produces an offspring by combining two different solutions (i.e. parents); while, the latter by producing random variations to a single parent. Then, the best solution in the previous population is enriched by the generated offspring and the procedure carries on from the selection phase, until the maximum number of iterations is achieved or the optimal values of objective function are the same in two successive iterations.

Contributions related to the implementation of such evolutionary techniques in transportation optimisation problems concern several issues: network design, routing and scheduling problems, timetabling and rescheduling tasks and energy consumption optimisation, as set out below. D'Acierno et al. (2014) addressed different network design problems, related to urban and extra-urban (i.e. rural) contexts as well as road and transit transport modes, by implementing Scatter Search and comparing the results with those obtained by means of other metaheuristic techniques such as Genetic Algorithm and Local Search methods. Moreover, Khooban et al. (2015) proposed a mixed network design problem with the aim of maximising the reserve capacity of the whole system and solved it by means of a hybrid Scatter Search method which incorporates the golden section search; while, Zhang at al. (2012) implemented both *SS* and *GA* for facing a stochastic travel-time vehicle routing problem with simultaneous pick-ups and deliveries. Similarly, Sun et al. (2014) proposed a genetic technique for solving a train routing problem combined with train scheduling, by taking into account average travel time, energy consumption and passenger satisfaction. In addition, Albercth (2009) proposed an automated timetable design method, characterised by a demand-oriented perspective, in which the computation of optimal departure times is performed by means of a Genetic Algorithm. Likewise, Niu and Zhou (2013) implemented a Genetic Algorithm, based on a binary coding approach, for solving a timetable optimisation problem in an urban rail line, by considering the time variability of travel demand and a multiple origin-to-destination demand pattern. Furthermore, Dundar and Sahin (2013) addressed a train rescheduling problem by implementing a Genetic Algorithm for minimising delays in conflict resolutions, together with an artificial neural network approach for simulating decision-making process of dispatchers during the failure management phase. Yang (2012) combined a

Genetic Algorithm with simulation techniques in order to identify the optimal energy-saving strategies to be implemented.

The *Ant Colony Optimisation (ACO)* method, instead, belongs to the family of swarm intelligence methodologies, based on the modelling of the collective behaviours of social insects, such as colonies of ants and termites or flocks of birds, which adopt decentralized control and self-organisation. The *ACO* was introduced by Colorni et al. (1992a; 1992b) and Dorigo (1992), and its basic principles are described in the following.

The idea was inspired by the exploitation of food resources by ants. These insects, although within the limits of cognitive capacities of the single ant, are able to collectively find the shortest path between a source of food and their nest. This is because they leave a trail of pheromones which attracts other ants.

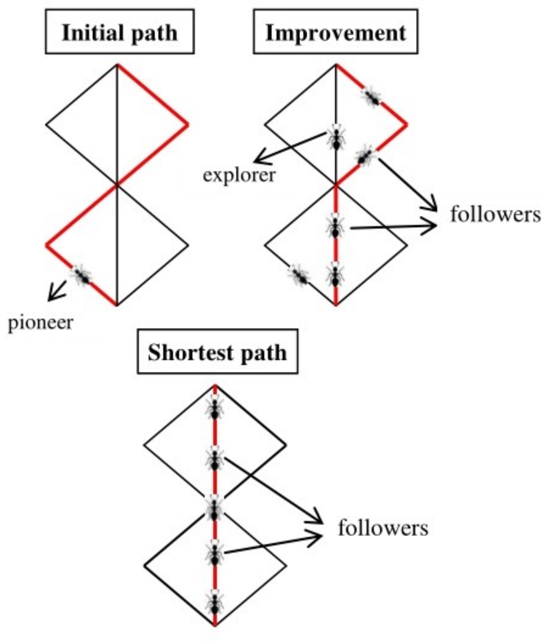


Figure 2.17 Ants behaviour

Specifically, when an ant is exploring an area in search of food, it leaves a trace. If it finds food, it returns and thus reinforces the trace. Hence, since pheromones are subject to evaporation, the shortest path will continuously be reinforced and will become the most attractive; while, the longest path will end up by disappearing and so, finally, all the ants will take the shortest path (figure 2.17).

The mathematical formulation adopted by the algorithm to model this phenomenon is set out below. In particular, the probability $P_{i,j}^{k}(t)$ with which the *k*-th ant, at the instant *t,* moves from state *i* to state *j* belonging to the set N_i^k, is expressed as follows:

$$P_{i,j}^{k}(t)=\frac{\left(\tau_{i,j}(t)\right)^{\alpha}\cdot\left(\eta_{i,j}(t)\right)^{\beta}}{\sum_{j\in N_i^k}\left(\tau_{i,j}(t)\right)^{\alpha}\cdot\left(\eta_{i,j}(t)\right)^{\beta}} \tag{2.52}$$

with

$$\eta_{i,j}=\begin{cases}\dfrac{1}{d_{i,j}} & \text{if } i\neq j\\ 0 & \text{otherwise}\end{cases} \tag{2.53}$$

where $\tau_{i,j}$ is the trail level of pheromone on the link *(i,j)*, i.e. a posteriori desiderability of the move; $\eta_{i,j}$ is the attractiveness of the move, i.e. a priori desirability of the move; α is the control parameter for the trail level ($\alpha\geq 0$); β is the control parameter for the attractiveness ($\beta\geq 1$); $d_{i,j}$ is the distance between nodes *i* and *j.*

The trial level of pheromone on the link *(i,j)* is updated as follows:

$$\tau_{i,j}(t)=\rho\cdot\tau_{i,j}(t-1)+\sum_{k=1}^{n}\Delta_{i,j}^{k}(t) \tag{2.54}$$

with

$$\Delta_{i,j}^{k}(t)=\begin{cases}\dfrac{Q}{L_k(t)} & \text{if link } (i,j) \text{ is chosen by the} k\text{-th ant}\\ 0 & \text{otherwise}\end{cases} \tag{2.55}$$

where ρ is the pheromone evaporation coefficient (0< ρ <1), n is the number of ants, $L_k(t)$ is the cost, generally in terms of length path, of k-th ant, at instance t; Q is a constant.

Many different variants of such a method are presented in the literature: ant system, elitist ant system, rank-based ant system, MAX–MIN ant system, ant colony system. An extended overview of ant-based algorithms can be found in Dorigo and Stützle (2004). Several transportation issues are addressed by means of *ACO* techniques such as assignment problems, optimal control theory and energy-saving tasks, vehicle routing problems and re-scheduling approaches. For instance, D'Acierno at al. (2006) integrated *ACO* into a *MSA* framework, in order to solve a Stochastic User Equilibrium assignment, and demonstrated the convergence of the proposed approach from a theoretical point of view by means of Blum's theorem; while, Ke et al. (2011) applied the so called MIN-MAX ant system, in order to optimise speed profiles of convoys between two stations thus providing a support tool for implementing strategies aimed at reducing energy consumption. In particular, in the proposed approach a cab-signalling system is considered and a fuzzy-PID gain scheduling mechanism is implemented for train acceleration. Moreover, thanks to its efficiency in terms of calculation time, the *ACO* is often implemented for real-time management approaches. In this context, Yan et al. (2016) proposed an *ACO* technique for implementing real-time energy saving policies in the case of high speed trains. In particular, the heuristic information parameter is designed according to the system status, in terms of delays, in order to adjust the trajectory planning procedure and allow the convoy to reduce the energy consumption by exploiting trip time redundancy. Likewise, Samà et al. (2016) implemented *ACO* in order to deal with the real-time problem of routing trains in a railway, which consists in re-optimising the routing of convoys under disruption conditions by identifying the potential best routing alternatives for each train and deciding which to implement with the purpose of re-establishing ordinary conditions as soon as possible. In addition, Samà et al. (2017b) implemented *ACO* for the same problem (i.e. train routing selection problem) by comparing its application, and the relative issues, in the case of two different dimensions, namely the tactical level and the operational stage. Furthermore, *ACO* techniques were implemented to address a railway junction rescheduling problem when a delay occurs, both in dynamic and static environments, respectively by Eaton and Yang (2016), and Fan et al. (2011). The latter also provides an interesting comparison between *ACO* and other seven optimisation approaches, among which Genetic Algorithm, Tabu Search and Simulating Annealing.

Obviously, the above mentioned contributions cannot in any way be considered exhaustive with regard to the copious number of applications of these techniques in the field of rail service management; however, they may make the reader aware of the numerous potentialities of such metaheuristic approaches.

2.8 Energy issues related to rail systems

In recent years, besides improving performance of rail systems so as to drive the modal split towards such a sustainable transport mode, thus reducing pollution and congestion effects due to private car use, considerable attention has been focused on energy issues for reducing energy consumption of systems based on a rail technology.

For this purpose, different approaches have been proposed in the literature, such as the adoption of eco-driving profiles, the regenerative braking, the introduction of timetable adjustments, the exploitation of on-board and way-side storage systems, the use of reversible substations. Clearly, they are strictly related to each other. In particular, the design of energy-efficient speed profiles consists in identifying the pattern which minimises the tractive energy consumption, given a running time to be respected (see, for instance, Albrecht et al., 2013; Miyatake and Ko, 2010); while, strategies based on the exploitation of regenerative braking aim to re-use the amount of kinetic energy produced during the braking phase by converting it back to the electrical one. In this case, the traction motor acts also as a generator and the recovered energy can be used at the exact time or stored for later use by means of energy storage devices. For instance, an on-board storage device allows to temporarily accumulate the excess regenerated energy and release it for the next acceleration phase of the same train (see, for instance, Miyatake and Matsuda, 2009; Steiner et al., 2007); while, the aim of a wayside storage device is to release it when required for other convoys' acceleration (see, for instance, Romo et al., 2005; Teymourfar et al., 2012). In this context, a timetable optimisation, aimed at synchronising acceleration and deceleration phases of convoys operating in the network, represents a key task for maximising the receptivity of the line (see, for instance, Kim et al., 2011; Nasri et al., 2010; Ramos et al., 2007; Yang et al., 2013). Additionally, the role of an energy-efficient timetabling phase lies in a suitable design of all operational times involved, such as running times, buffer times, dwell times and reserve times (Canca and Zarzo, 2017; D'Acierno et al., 2017; Wong and Ho, 2007). Moreover, by means of reversible or active substations, the regenerated energy can also be traced back to the medium voltage distribution network (Cornic, 2011; Ibaiondo and Romo, 2010).

An extensive overview of regenerative braking issues and energy storage systems, together with the above-mentioned related concerns, can be found respectively in Ghavihaa et al. (2017), and Gonzales-Gil et al. (2013). This work, instead, is focused on strategies involving the design of suitable speed profiles and the optimisation of operational times within timetable in an energy saving perspective.

Regarding the eco-driving profiles, first of all, it is necessary to introduce the reference scenario, indicated as the *Time Optimal (TO)* scenario, which consists in considering the movement of the convoy in the case of maximum performance. It foresees a first part in which the train adopts the maximum acceleration value in order to reach the maximum speed (*acceleration phase*), a second part at constant speed (*cruising phase*) and, finally, there is a braking phase until the convoy draws to a halt (*deceleration phase*). For the sake of simplicity, we will refer to a motion diagram of the trapezium type (jerk value equals $+\infty$), represented in figure 2.18.

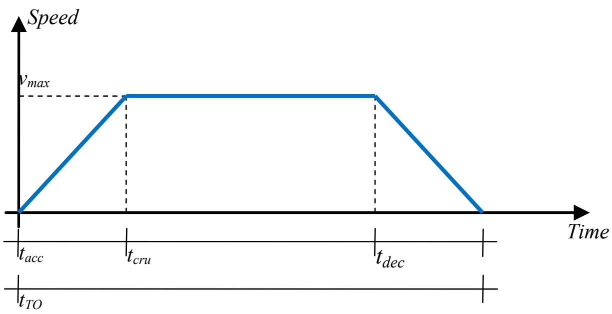


Figure 2.18 Speed profile in the case of a *Time Optimal (TO)* strategy

The total travel time between two successive stops (i.e. stations or red signals), in this case, may be calculated as follows:

$$t_{TO} = t_{acc} + t_{cru} + t_{dec} \tag{2.56}$$

where t_{TO} is the travel time in the case of *TO* strategy; t_{acc} is the time duration of the acceleration phase; t_{cru} is the time duration of the cruising phase; and t_{dec} is the time duration of the deceleration phase.

This condition of maximum performance corresponds to the minimum travel time and the maximum energy consumption. In this context, two different eco-driving strategies can be adopted, which consist respectively in:

1. inserting, between the cruising and the braking phases, a further stage, which is the so called *coasting phase*, during which the convoy moves by inertia (figure 2.19);
2. reducing the value of maximum speed (figure 2.20).

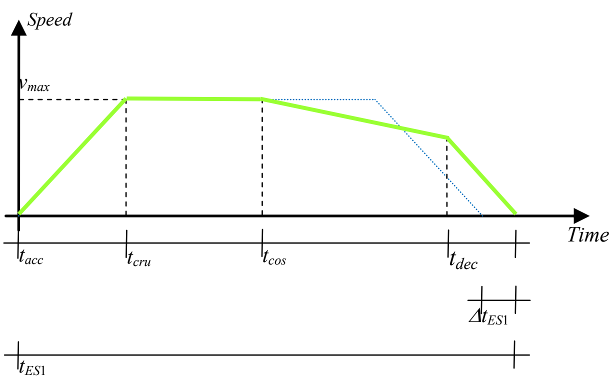


Figure 2.19 Speed profile in the case of Energy Saving (ES) strategy 1

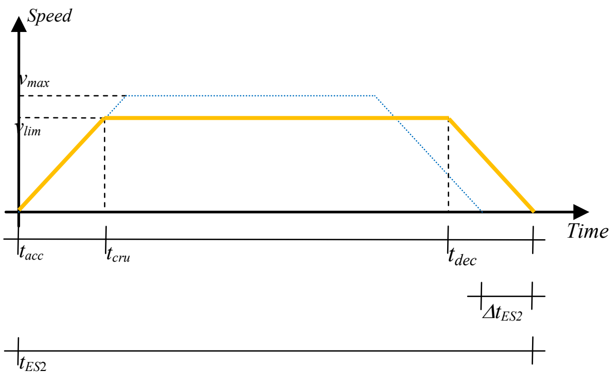


Figure 2.20 Speed profile in the case of Energy Saving (ES) strategy 2

The first strategy requires reporting to the train the switching points for the coasting phase; while the second one is more straightforward to implement, since it requires simply communicating a different speed limit. Therefore, the technological level of the rail system may affect the choice between these two approaches.

However, the total travel time between two successive stops, in both cases, increases. In particular, for the first strategy. the total travel time can be expressed as follows:

$$t_{ES1} = t_{acc} + t_{cru} + t_{cos} + t_{dec} \tag{2.57}$$

$$t_{ES1} = t_{TO} + \Delta t_{ES1} \tag{2.58}$$

where t_{ES1} is the travel time in the case of the first *ES* strategy; t_{cos} is the time duration of the coasting phase; Δt_{ES1} is the increase in travel time in the case of the first *ES* strategy with respect to *TO* strategy.

While, in the case of the second strategy, it can be calculated as follows:

$$t_{ES2} = t_{acc} + t_{cru} + t_{dec} \tag{2.59}$$

$$t_{ES2} = t_{TO} + \Delta t_{ES2} \tag{2.60}$$

where t_{ES2} is the travel time in the case of the second *ES* strategy; Δt_{ES2} is the increase in travel time in the case of the second *ES* strategy with respect to *TO* strategy.

In order to derive the increase in travel time, let

$$E(\Delta t) = \int_0^{\Delta t} dE(\tau) = \int_0^{\Delta t} P(\tau) \cdot d\tau = \int_0^{\Delta t} T(v(\tau)) \cdot v(\tau) \cdot d\tau \tag{2.61}$$

be the mechanical kinetic energy E required to move a rail convoy during time interval Δt. In particular, $dE(\tau)$ is the increase in kinetic energy at time τ; τ is the generic time instant; $P(\tau)$ is the instantaneous power at time τ; $d\tau$ is the generic infinitesimal time interval; $v(\tau)$ is the instantaneous speed at time τ; $T(\cdot)$ is the tractive effort (i.e. tractive force) at rail wheels which depends on instantaneous speed $v(\cdot)$.

Therefore, by imposing the constancy of the section length, the increase in travel time for the first strategy can be formulated as follows:

$$\Delta s = \int_0^{t_{TO}} v_{TO}(\tau) \cdot d\tau = \int_0^{t_{ES1}} v_{ES1}(\tau) \cdot d\tau \tag{2.62}$$

where Δs is the track length between the two successive stops analysed; $v_{TO}(\cdot)$ is the speed profile in the case of *TO* strategy, as shown in figure 2.18; $v_{ES1}(\cdot)$ is the speed profile in the case of the first *ES* strategy as shown in figure 2.19.

Similarly, for the second strategy:

$$\Delta s = \int_0^{t_{TO}} v_{TO}(\tau) \cdot d\tau = \int_0^{t_{ES2}} v_{ES2}(\tau) \cdot d\tau \tag{2.63}$$

where $v_{ES2}(\cdot)$ is the speed profile in the case of the second *ES* strategy, as shown in figure 2.20.

Hence, eco-driving policies are based on the adoption of speed profiles which are distant from those at maximum performance and, thus, provide a longer travel time. This implies that they are feasible only if there is an extra time availability on a given line service. This time is generally known as *reserve time.* In order to clarify this concept, it is worth analysing the different time rates which concern the timetable design phase. In particular, as already said, this task involves the computation of running times between two stops, dwell times at stations for the boarding/alighting process, buffer times and layover times.

Buffer times are generally set up during the design phase in order to address possible delays or, simply, eventual fluctuations which can occur during the service, given the stochasticity of the phenomenon being examined. It is sufficient to think, for instance, that inevitably every train driver drives in his own way, but even the very same train driver might drive in two different ways on two different days. Obviously, the lower the level of automation, the higher the relevance of the stochastic nature of the involved factors. The layover time is a time spent by the convoy at the terminus. The minimum layover time is represented by the inversion time and, eventually, by the time required for possible shunting activities. Moreover, there could be an additional time interval that goes between when the convoy is physically ready to undertake the run in the opposite direction and when it can effectively depart according to the timetable indications. However, in certain cases, with the term layover time is indicated exclusively this further time rate, while the inversion time is computed in the cycle time. For the sake of completeness, also synchronisation times, for making available transfer options for passengers, can be taken into account. Hence, the above-mentioned extra time availability could involve running time reserve, dwell time reserve, buffer time and eventual time exceeding the layover time at the terminus. These times are properly scheduled during the timetable design phase by increasing the minimum times required for the service. For example, as to travel time, the International Union of Railways (UIC) suggested increasing the minimum travel time by a percentage of 3-8 %. Obviously, the possibility of exploiting these extra times, for implementing such energy saving strategies, is subject to the preservation of timetable stability and service quality. Therefore, the identification of an analytic framework for quantifying in a reliable manner the timetable rates involved in the implementation of energy saving strategies, as well as the definition of an optimisation model which takes into account the trade-off between eco-driving profiles and passenger needs, turn out to be fundamental. Moreover, it is worth noting that, in a rail context, *ES* strategies are commonly implemented between two successive stops; while, in a metro context, the most suitable approach consists in examining the whole outward and return trip, given the fact that the service is frequency-based (Cepeda et al., 2006; Fu et al., 2012; Nuzzolo et al., 2012), which means that the parameter to be respected is the headway between two successive convoys, rather than a timetable, generally unknown to users. Therefore, in the case of metro systems, the energy saving strategies are implemented by considering arrival and departure times at the terminus, rather than at each station.

However, according to the literature, these techniques can be applied separately, by addressing individually the design of energy-efficient driving profiles (see, for instance, Chuang et al., 2008; Yang et al., 2012) and the optimisation of operational times within the timetable (see, for instance, Albrecht et al., 2002; Lancien and Fontaine, 1981) or, more frequently, in an integrated framework. In this context, Li and Lo (2014) proposed a train control approach, based on an optimisation model, which combines energy-efficient timetables and speed profiles. In particular, the procedure is characterised by a dynamic layout, since it provides a dynamic adjustment of the cycle time on the basis of travel demand changes, in order to minimise the energy consumption; moreover, a linear approximation method is implemented with the aim of dealing with a convex optimisation problem, whose resolution is performed by means of the Kuhn–Tucker conditions. Scheepmaker and Goverde (2015) developed a nested optimisation framework in which, by starting from the planned total running time, energy-efficient speed profiles are derived. More in detail, the optimal cruising speed is defined by means of the outer loop of the Fibonacci algorithm (Mathews and Fink, 2004; Siegler, 1987); while, in the inner loop, the bisection method computes, for the given cruising speed, the optimal switching points of the coasting phase. Moreover, different distributions of running time supplements are tested and compared in terms of service punctuality and energy consumption. Sicre et al. (2010) devised a simulation-based optimisation procedure in which the simulation model provides the most energy-efficient driving profile, by computing energy consumption Pareto curves for each stretch, and the optimisation tool allocates the total running reserve time available in the most efficient way among the different stretches. The proposed simulation technique deals with a manual driving mode and, specifically, allows to carry out a large variety of manual driving strategies by combining different sections of holding speed with different coasting windows. Feng et al. (2017) enriched the common optimisation framework, which combined energy control strategies with a suitable design of operational times, by performing the estimation of dwell times at stations as function of the number of passengers involved in the boarding/alighting process. By considering dwell time as a flow-dependent factor, rather than a fixed value, clearly, a more realistic computation of dwell time itself can be carried out; but, most importantly, dwell time margin, which plays a key role within the implementation of energy saving strategies, can be derived with a high degree of accuracy. Moreover, both manual (Acikbas and Soylemez, 2008; Lukaszewicz, 2000; Wong and Ho, 2004a) and automatic (Carreno, 2017; de Cuadra et al., 1996; Domínguez et al., 2012) driving systems have been investigated in the literature.

As already touched upon, the most common methodologies for analysing these strategies are simulation-based techniques. In this context, Zhao et al. (2015) developed a multi-train simulator and incorporated it into an optimisation framework whose aim is to minimise the trade-off between energy consumption and delay penalty. Additionally, both exhaustive and metaheuristic approaches are compared to optimise train operations such as enhanced brute force, ant colony optimisation and genetic algorithm. Moreover, Sicre et al. (2012) developed an off-line eco-driving design model based on simulation tasks, whose aim is defining manual energy-efficient profiles, in terms of easily interpretable and executable commands for the driver, and implemented a genetic algorithm as optimisation search technique. In particular, the proposed approach takes into account also passenger satisfaction and considers very detailed parameters such as maximum number of commands, minimum separation between commands and minimum speed of arrival at stations. De Martinis et al. (2014), and De Martinis and Weidmann (2015) merged a speed profile optimisation tool, based on a genetic algorithm as subroutine, with a micro-simulation model which reproduces the interactions among infrastructure, signalling system, rolling stock and timetable. In addition, the proposed methodology can be implemented for real-time rescheduling tasks by updating the timetable database information time after time. Other real-time approaches can be found in Chang and Chung (2005), Corman et al. (2009), D'Ariano and Albrecht (2006), Sheu and Lin (2011).

Regarding the adopted resolution methods, analytical approaches for modelling ES strategies have been proposed by Albrecht et al. (2013), Howlett et al. (2009), Khmelnitsky (2000), Kim and Chien (2011), Liu and Golovitcher (2003). However, as already pointed out, given the complexity of the matter which involves different components, whose interactions have to be modelled with a high degree of detail, also several metaheuristic techniques have been implemented, such as genetic algorithm (Ding et al., 2011; Huang et al., 2015; Keskin and Karamancioglu, 2015), ant colony optimisation (Ke et al., 2009; Lu et al., 2013b; Yan et al., 2016), simulating annealing (Kim and Chien, 2011). Furthermore, there are some contributions which combine evolutionary techniques with a fuzzy logic (Bocharnikov et al., 2007; Sicre et al., 2014) as well as with artificial neural network approaches (Acikbas and Soylemez, 2008; Chuang et al., 2008).

Several works incorporated the energy saving perspective in a multi-objective framework. Indeed, eco-driving speed profiles, generally, imply an increase in train running times and, therefore, in passenger travel times. For this reason, several authors focused on the trade-off between energy saving and passenger satisfaction (Chevrier et al., 2016; Corapi et al., 2014; Ghoseiri et al., 2004; Xu et al., 2016b; Yin et al., 2017). More in general, Cucala et al. (2012), Toletti et al. (2016), and Tonosaki et al. (2016) analysed the relation between energy-efficient strategies and stability of the planned timetable; while, Feng et al. (2014) analysed also the utilisation rate of train capacity resulting from the implementation of energy-saving strategies. Finally, Canca (2017) compared the minimum-energy timetable with those obtained by taking into account also rolling stock and other operational costs.

CHAPTER 3: PROPOSED METHODOLOGY FOR MANAGING RAIL SYSTEMS BOTH IN ORDINARY AND DISRUPTION CONDITIONS

The proposed methodological framework has been conceived as a decision-making tool for handling rail operations, both in the planning and the management phase, by duly taking into account space-time variability of travel demand and adopting a passenger-oriented perspective.

Railways represent a strategic sector for changing the balance of transport modes, given the high level of sustainability and performance which they offer, and, therefore, a valorisation of such systems is imperative. For this purpose, a two-fold action is pursued: the improving in service quality to attract users from other transport modes with greater environmental impacts (such as private cars) and, on the other hand, the reduction in energy consumption by means of suitable energy saving strategies. Therefore, in order to perform a realistic assessment, a modelling of rail operations which duly considers the interaction with travel demand, as well as relative energy consumption implications, is required.

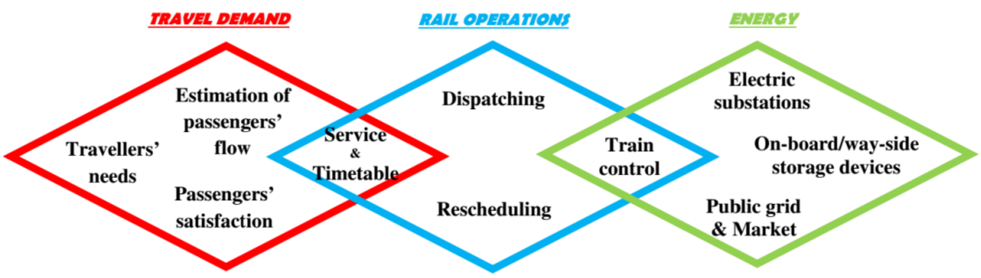


Figure 3.1 Rail operations and their interactions with travel demand and energy system

In figure 3.1, the three systems of concern (i.e. travel demand, rail operations and energy) are shown, together with the interactions existing among them.

In order to be able to implement the above described framework on a practical basis, so as to show its numerous applicative potentialities, it is necessary to make use of suitable simulation techniques, properly integrated into *ad-hoc* developed optimisation tools, which enable, on one hand, the acquisition of knowledge of the effects of any intervention, before this is carried out, and, on the other, the identification of the best resolution strategy according to the target pursued.

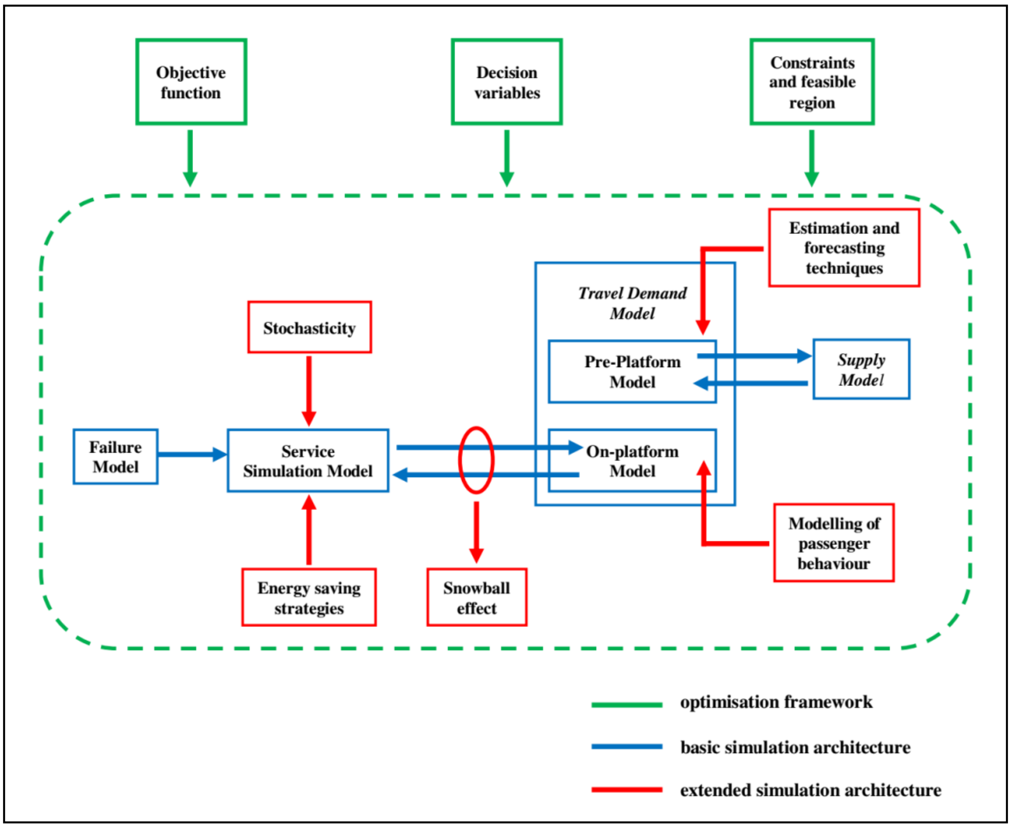


Figure 3.2 Layout of the proposed approach

In particular, the proposed simulation-optimisation integrated approach, depicted in figure 3.2, can be decomposed into three fundamental parts:

1. the optimisation framework;
2. the basic simulation architecture;
3. the extended simulation architecture;

whose details will be provided in the following.

3.1 Optimisation framework

In this section the problem of identifying the optimal intervention strategy for addressing a system failure is formalised as a bi-level multidimensional constrained optimisation problem which can be specified as follows:

$$\hat{\boldsymbol{y}} = \underset{\boldsymbol{y} \in \boldsymbol{S}_y}{arg\,min}\, Z(\boldsymbol{y}, \boldsymbol{fc}, \boldsymbol{tnp}, \boldsymbol{rnp}, \boldsymbol{td}) \tag{3.1}$$

subject to:

$$\hat{\boldsymbol{y}} = \underset{\boldsymbol{y} \in \boldsymbol{S}_y}{arg\,min}\, Z(\boldsymbol{y}, \boldsymbol{fc}, \boldsymbol{tnp}, \boldsymbol{rnp}, \boldsymbol{td}, \boldsymbol{in}^0, \boldsymbol{rs}^0, \boldsymbol{ss}^0, \boldsymbol{pt}) \tag{3.2}$$

with:

$$\begin{aligned} &Z(\boldsymbol{y}, \boldsymbol{fc}, \boldsymbol{tnp}, \boldsymbol{rnp}, \boldsymbol{td}) = \\ &= \sum_i \beta^i_{VOT} \cdot \left(\beta^i_{waiting} \cdot \sum_s \sum_p \sum_r tw^{i,r}_{s,p}(\boldsymbol{y}, \boldsymbol{fc}, \boldsymbol{tnp}, \boldsymbol{rnp}) \cdot fw^{i,r}_{s,p}(\boldsymbol{td}) + \right. \\ &\qquad \left. + \beta^i_{on-board}(\boldsymbol{td}) \cdot \sum_l \sum_r tb^{i,r}_l(\boldsymbol{y}, \boldsymbol{fc}, \boldsymbol{tnp}, \boldsymbol{rnp}) \cdot fb^{i,r}_l(\boldsymbol{td}) \right) \end{aligned} \tag{3.3}$$

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; $\boldsymbol{\Lambda}$ 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 time; $tw^{i,r}_{s,p}$ is the average user waiting time of user category i at station s, on platform p between run $(r-1)$ and run *r*; $fw_{s,p}^{i,r}$ is the number of passengers of user category *i* waiting at station *s*, on platform *p* between run *(r–1)* and run *r*; $\beta_{on-board}^{i}(\boldsymbol{td})$ 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 (i.e. it assumes a different value depending on whether the considered users are standing or sitting); $tb_l^{i,r}$ is the average 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 particular, the vector which identifies the intervention strategy (i.e. $\boldsymbol{y}$) can be viewed as made up by four components:

y_1 expressing the strategy type implemented (e.g. inversion with passengers on-board, inversion after unloading passengers, recovery on a maintenance track, waiting for a rescue means, re-routing, skipping some stops);

y_2 expressing *when* the strategy has to be implemented (e.g. as soon as possible, during the outgoing trip, during the return trip, during the layover time at the terminus);

y_3 expressing *where* the strategy has to be implemented, intended as the station where to take action;

y_4 expressing specific features of the intervention strategy (e.g. the use of a rescue vehicle or the use of a spare train).

Obviously, the possible presence of unfeasible combinations, for technical or regulatory reasons, has to be properly taken into account. For instance, the recovery on a maintenance track can be performed only in a station effectively equipped with this kind of track; the faulty convoy can change direction only in a station with points or, according to the Italian regulation, can be towed by a rescue vehicle only if empty (i.e. no passengers on-board).

Equation (3.2) defines the consistency constraint between transportation performance and travel demand flow. Its formulation requires the adoption of an articulate modelling framework, which represents the simulation architecture of the proposed approach. More details about its basic and extended structures will be provided in the following sections.

Finally, the objective function (3.3) is expressed in terms of user generalised costs, given the passenger-oriented perspective adopted in this work. In particular, only waiting and travel times are considered, since they are the only two parameters which change during the simulation of the different scenarios; while, the other attributes, which remain constant, such as the monetary cost, can be neglected.

However, a multi-objective approach can be adopted by means of more complex formulations. For example, by considering also operational costs of rail companies, objective function (3.3) becomes:

$$Z(\boldsymbol{y}, \boldsymbol{fc}, \boldsymbol{tnp}, \boldsymbol{rnp}, \boldsymbol{td}, \boldsymbol{rc}) = \beta_{UGC} \cdot UGC + \beta_{PEN} \cdot PEN + \beta_{TOC} \cdot TOC \qquad (3.4)$$

where $\boldsymbol{rc}$ is the vector of residual capacities of rail convoys. Moreover:

- *UGC* is the user generalised cost and, therefore, coincides with objective function (3.3);
- *PEN* represents the extra-cost perceived by passengers who are forced to leave the system because of a disruption event or extremely crowded conditions. In this case, indeed, the increase in waiting times can lead passengers to choose an alternative mass-transit system for reaching their destination. This term is expressed by the following equation:

$$PEN = \sum_{s} \cdot \sum_{p} \cdot \sum_{r} pl_{s,p}^{r} \cdot (optw_{s,p} + tls) \cdot \beta_{VOT} \qquad (3.5)$$

where $pl_{s,p}^{r}$ is the number of passengers 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 and change transport mode. Moreover, except in the case of an integrated fare schemes, this implies also an additional monetary cost for buying another ticket.

- *TOC* is the total operational cost incurred by rail operators for each train performing the service:

$$TOC = \sum_{r} L_r \cdot c_r \cdot ntu_r \tag{3.6}$$

where L_r is the length of the path performed by run *r* expressed in kilometres; c_r is the cost per traction unit-km; ntu_r is the number of traction units composing the run *r;*

- β_{UGC}, β_{PEN} and β_{TOC} are homogeneity coefficients which express the relative weight of the objective function terms.

Moreover, the described multi-objective framework can be further enriched by introducing other operational cost items and the evaluation of external costs, as will be shown in the following.

3.2 Basic simulation architecture

The basic simulation architecture is given by four different models which interact so as to replicate the analysed system features and model the consistency constraint between transportation system performance and travel demand flow (i.e. equation 3.2). They are: the *Service Simulation Model* (*SeSM*), the *Travel Demand Model* (*TDM*), the *Supply Model* (*SM*) and the *Failure Model* (*FM*) which is get involved when it is necessary to model perturbed conditions.

The *SeSM* provides rail system performance as function of rail infrastructure, rolling stock, signalling system, timetable and travel demand, both in ordinary and disruption conditions. It is performed by means of a microscopic synchronous rail simulation software which is able to model the service with a high level of detail.

The *TDM* is subdivided into two sub-models:

- the *Pre-platform model* (*PPM*) which provides the number of passengers arriving at station as the result of interaction with the supply model;
- the *On-platform model* (*OPM*) which simulates the dynamic interaction between rail service and travel demand occurring on platform, when a train arrives, during the boarding/alighting process. This interaction produces the so called *snowball effect*: the number of passengers on the platform influences the dwell times of trains at stations which, in turn, cause increasing delays; this implies an increase in headways which could generate more passenger flows on the platform (generally proportional to the headway increase), producing a further extension of dwell times. In particular, by considering dwell times as function of the involved flows, the snowball effect can be modelled as a a fixed-point problem. Moreover, the *OPM* duly takes into account capacity constraints of convoys and specific assumptions on passenger behaviour. Its basic structure is based on a *FIFO* (*First In-First Out*) queuing rule; however, also different priority boarding patterns can be easily implemented.

The *SM* provides performance of all transportation systems within the study area, so as to allow to model the modal split among different transport modes and compute a better estimation of the arrival rate at each station. As already mentioned, it interacts with the *PPM* by generating a further fixed-point problem as described in paragraph 2.1.6.

Finally, the *FM* provides the failure scenarios to be analysed by means of the implementation of the *RAMS (Reliability*, *Availability*, *Maintainability* and *Safety)* techniques (CENELEC, 1999) which allow to estimate the probability of failure for any element of the network (e.g. damage to a convoy, block of a track section, breakdown of a signalling system device) and calculate the effects on the rail system.

By specifying the formulation of each model and considering their interactions, equation (3.2) can be re-written as follows:

$$\begin{cases} \boldsymbol{tnp} = SM(\boldsymbol{rnp}, \boldsymbol{td}) \\ \boldsymbol{rnp} = SeSM\left(\boldsymbol{y}, FM\left(\boldsymbol{fc}, \boldsymbol{in}^0, \boldsymbol{rs}^0, \boldsymbol{ss}^0\right), \boldsymbol{td}, \boldsymbol{pt}\right) \\ \boldsymbol{td} = OPM\left(\boldsymbol{rnp}, PPM(\boldsymbol{tnp}, \boldsymbol{rnp}), FM\left(\boldsymbol{fc}, \boldsymbol{in}^0, \boldsymbol{rs}^0, \boldsymbol{ss}^0\right)\right) \end{cases} \tag{3.7}$$

where all parameters have been described before.

3.3 Extended simulation architecture

This section aims at improving the basic structure described in the previous paragraph with more detailed modelling techniques, which make the simulation more realistic and enhance the accuracy of the analysis carried out.

Firstly, the micro-simulation framework performing the *SeSM* is enriched by introducing the explicitly modelling of the stochastic nature of the involved factors, as well as the possibility of simulating the implementation of energy saving strategies.

Moreover, the dynamic interaction between on-platform flows and rail service, leading, especially in crowded contexts, to the *snowball effect*, is explicitly considered by developing a suitable tool which is able to compute dwell times as flow-dependent factors, rather than fixed values. In particular, the proposed method is based on the implementation of the *OPM* which simulates passenger behaviour on platform when a train arrives.

Furthermore, the *OPM* is enhanced by introducing the possibility of modelling different behavioural patterns for passengers during the boarding/alighting process.

Finally, also the *PPM* is improved by means of the development of travel demand estimation and forecasting (i.e. long-term evaluation) techniques which are customised to the case of rail systems.

In the following, each one of the above-mentioned improvements will be described in detail.

3.3.1 Stochastic simulation framework

In this case, the involved variables are viewed as the sum of average values and random residuals, rather than as fixed values. For this purpose, if $\boldsymbol{X}$ is the considered multivariate random variable, it has the following expression:

$$\boldsymbol{X} = \overline{\boldsymbol{X}} + \boldsymbol{\varepsilon_X} \tag{3.8}$$

where $\overline{\boldsymbol{X}}$ is a fixed vector whose elements are the mathematical expectations (i.e. first moments or means) of the elements of $\boldsymbol{X}$ (i.e. $\overline{\boldsymbol{X}} = E[\boldsymbol{X}]$) and $\boldsymbol{\varepsilon_X}$ is the random residual of $\boldsymbol{X}$, distributed according to a certain statistical rule $\boldsymbol{\Gamma_X}(\cdot)$, that is:

$$\boldsymbol{\varepsilon_X} \sim \boldsymbol{\Gamma_X}(\boldsymbol{\alpha_X}) \tag{3.9}$$

where $\boldsymbol{\alpha_X}$ is the vector of parameters of the adopted statistical distribution.

Therefore, it can be stated that:

$$\begin{aligned} &[\boldsymbol{y, fc, tnp, rnp, td, in^0, rs^0, ss^0, pt}] = \\ &= [\overline{\boldsymbol{y}}, \overline{\boldsymbol{fc}}, \overline{\boldsymbol{tnp}}, \overline{\boldsymbol{rnp}}, \overline{\boldsymbol{td}}, \overline{\boldsymbol{in}}^{\boldsymbol{0}}, \overline{\boldsymbol{rs}}^{\boldsymbol{0}}, \overline{\boldsymbol{ss}}^{\boldsymbol{0}}, \overline{\boldsymbol{pt}}] + \\ &+ [\boldsymbol{\varepsilon_y}, \boldsymbol{\varepsilon_{fc}}, \boldsymbol{\varepsilon_{tnp}}, \boldsymbol{\varepsilon_{rnp}}, \boldsymbol{\varepsilon_{in^0}}, \boldsymbol{\varepsilon_{rs^0}}, \boldsymbol{\varepsilon_{ss^0}}, \boldsymbol{\varepsilon_{pt}}] \end{aligned} \tag{3.10}$$

Clearly, the deterministic approach can be re-obtained merely by setting the vector of random residuals (i.e. $\boldsymbol{\varepsilon_X}$) equal to zero.

By implementing such approach, it is possible to model different stochastic parameters which affect the analysed system such as train performance (e.g. speed and acceleration), travel times, dwell times and delays. Moreover, the stochasticity of travel demand can be taken into account by explicitly modelling the distribution of passenger flows. Regarding the randomness of occurrence of failure events, it is worth noting that it is including into the *Failure Model* which, as stated above, implements the *RAMS* techniques.

In addition to the possibility of taking into account the random nature of single elements, a stochastic approach also allows to perform a global analysis of robustness of the recovery solutions obtained by means of a deterministic approach. In particular, a two-step procedure is proposed. Firstly, by means of deterministic microscopic simulations, the optimal intervention strategy $\hat{\boldsymbol{y}}$ and its neighbourhood $N(\hat{\boldsymbol{y}})$ are evaluated, for each failure context. The considered neighbourhood consists in all corrective actions providing objective function values close to the minimum cost (i.e. objective functions calculated in the case of strategy $\hat{\boldsymbol{y}}$). The second step consists in carrying out numerous microscopic simulations, by changing stochastically the input parameters, in order to perform a sensitivity analysis of the deterministic solution obtained (i.e. $\hat{\boldsymbol{y}}$) thus providing information about the degree of reliability ensured by it.

3.3.2 Decision support system for implementing energy saving strategies

This section deals with an analytical methodology developed for enabling an accurate computation of operational times within the timetable, so as to properly support the implementation of eco-driving strategies. Indeed, as already explained, such strategies imply an increase in running times of convoys and, therefore, they are feasible exclusively if there is a possibility of exploiting the availability of extra-time rates, properly scheduled during the timetable planning phase.

Generally, relevant time rates for the implementation of energy saving strategies can be: running time supplements, dwell time supplements and reserve times (*rt*), which can be defined as the sum of buffer times (*bt*) and layover times (*lt*), as shown by equation (3.11):

$$rt = bt + lt \tag{3.11}$$

In particular, as the proposed analytical method is explicitly designed for metro contexts, their nature of frequency-based services (i.e. the parameter to be respected is the headway between two successive convoys, rather than a timetable, generally unknown to users) has to be properly taken into account.

For this reason, considering the arrival and departure times at the terminus, rather than at each station, is the most appropriate approach. Therefore, it is possible to define reserve times rt_{ot} and rt_{rt} associated respectively to the outward trip (*ot*) and return trip (*rt*), as follows:

$$rt_{ot} = bt_{ot} + lt_{ot} \tag{3.12}$$

$$rt_{rt} = bt_{rt} + lt_{rt} \tag{3.13}$$

where lt_{ot} and lt_{rt} are the layover times associated respectively to the outward trip (*ot*) and return trip (*rt*); bt_{ot} and bt_{rt} are the buffer times respectively in the case of the outward trip (*ot*) and return trip (*rt*).

By an operational (i.e. relative to rail service) point of view, the function of time supplements is that of facing primary delays; while, buffer times are designed for minimising the so called secondary delays, since they are generated by the propagation of primary delays. Finally, the layover time is the time a train spends at the terminus. Regarding the definition of this parameter, it is worth making the following clarification. Generally, the layover time includes the time required for changing direction and making shunting or (de-)coupling operations, if any, together with an additional time rate that goes between when the convoy is physically ready to undertake the run in the opposite direction and when it can effectively depart according to the timetable indications. Specifically, in the proposed approach, since the inversion time *(it)*, including also eventual time required for shunting and (de-)coupling operations, is computed in the cycle time formulation, the layover time involves exclusively the additional time rate which is ‘wasted’ by the convoy at the terminus, waiting for the right moment to depart, in order to maintain the planned headway unaltered. In fact, if it were to depart previously, the headway would be lower than the planned value; on the contrary, if it were to depart afterwards, the headway would be higher than the planned value. Hence, the layover time (*lt*) appears as the only time resource which can be exploited without eroding other time rates designed for preserving timetable stability. Consequently, our proposal is focused on the use of such a parameter as the extra time source for the implementation of energy saving strategies. For this purpose, the *total usable reserve time* (*turt*) is introduced:

$$turt = lt_{ot} + lt_{rt} \tag{3.14}$$

In the following, the operational parameters involved in the described framework, and the relations existing among them, are formalised.

In the case of a metro system, the number of convoys required to perform the service may be calculated as:

$$NC = (CT + lt_{ot} + lt_{rt})/H \tag{3.15}$$

subject to:

$$0 \leq bt_{ot} + lt_{ot} \leq H \tag{3.16}$$

$$0 \leq bt_{rt} + lt_{rt} \leq H \tag{3.17}$$

with:

$$CT = \sum_{lot} tt_{lot} + \sum_{sot} dt_{sot} + it_{ot} + bt_{ot} + \sum_{lrt} tt_{lrt} + \sum_{srt} dt_{srt} + it_{rt} + bt_{rt} \tag{3.18}$$

where NC is the number of convoys; CT is the cycle time being calculated by means of equation (3.18); H is the headway between two successive rail convoys; tt_{lot} and tt_{lrt} are the travel times associated, respectively, to link l_{ot} and l_{rt}; l_{ot} and l_{rt} are the generic links (i.e. track sections) associated, respectively, to the outward trip (*ot*) and return trip (*rt*); dt_{sot} and dt_{srt} are the dwell times associated, respectively, to platform *sot* and *srt*; *sot* and *srt* are the generic platforms of station *s* for, respectively, the outward trip (*ot*) and return trip (*rt*); it_{ot} and it_{rt} are the inversion times (i.e. preparation times for the subsequent trip) associated, respectively, to the outward trip (*ot*) and return trip (*rt*).

By substituting (3.14) into (3.15), the following relation is obtained:

$$NC = (CT + turt)/H \tag{3.19}$$

by which it is possible to derive the *turt* as:

$$turt = NC \cdot H - CT \tag{3.20}$$

Moreover, equations (3.16) and (3.14) can be re-written as follows:

$$0 \le lt_{ot} \le H - bt_{ot} \tag{3.21}$$

$$lt_{rt} = turt - lt_{ot} \tag{3.22}$$

and allow to identify feasible values for lt_{ot} and lt_{rt}.

However, theoretically, the *turt* value could be split arbitrarily between the outward and the return trip. Hence, we may introduce a parameter α expressing the partition rate as follows:

$$lt_{ot} = \alpha \cdot turt \tag{3.23}$$

$$lt_{rt} = (1 - \alpha) \cdot turt \tag{3.24}$$

with

$\alpha \in [0;1]$.

In particular, by substituting (3.23) into (3.16):

$$0 \le bt_{ot} + \alpha \cdot turt \le H \Rightarrow 0 \le \alpha \cdot turt \le H - bt_{ot} \Rightarrow 0 \le \alpha \le \frac{H - bt_{ot}}{turt} \tag{3.25}$$

Similarly, by substituting (3.24) into (3.17):

$$0 \le bt_{rt} + (1-\alpha) \cdot turt \le H \Rightarrow 0 \le (1-\alpha) \cdot turt \le H - bt_{rt} \Rightarrow 0 \le (1-\alpha) \le \frac{H - bt_{rt}}{turt} \Rightarrow$$
$$\Rightarrow -1 \le -\alpha \le \frac{H - bt_{rt}}{turt} - 1 \Rightarrow 1 - \frac{H - bt_{rt}}{turt} \le \alpha \le 1 \tag{3.26}$$

At this point, two different cases may occur:

a) if

$$1 - \frac{H - bt_{rt}}{turt} \le \frac{H - bt_{ot}}{turt} \Rightarrow$$

$$\Rightarrow turt \le 2 \cdot H - \left(bt_{ot} + bt_{rt}\right)$$

then

$$\max\left(1 - \frac{H - bt_{rt}}{turt}; 0\right) \le \alpha \le \min\left(\frac{H - bt_{ot}}{turt}; 1\right); \qquad (3.27)$$

b) if

$$1 - \frac{H - bt_{rt}}{turt} > \frac{H - bt_{ot}}{turt} \Rightarrow$$

$$\Rightarrow turt > 2 \cdot H - \left(bt_{ot} + bt_{rt}\right)$$

then:

the solution does not exist.

Therefore, only in the case *a)* it is possible to identify a feasible set for parameter α. However, by means of the following steps, it is possible to demonstrate that the case *b)* never occurs. Indeed:

$$1 - \frac{H - bt_{rt}}{turt} > \frac{H - bt_{ot}}{turt} \Rightarrow$$

$$\Rightarrow turt > 2 \cdot H - \left(bt_{ot} + bt_{rt}\right) \Rightarrow$$

$$\Rightarrow NC \cdot H - CT > 2 \cdot H - \left(bt_{ot} + bt_{rt}\right) \Rightarrow$$

$$\Rightarrow NC \cdot H - \left(\left(tt_{ot} + tt_{rt}\right) + \left(dt_{ot} + dt_{rt}\right) + \left(it_{ot} + it_{rt}\right) + \left(bt_{ot} + bt_{rt}\right)\right) > 2 \cdot H - \left(bt_{ot} + bt_{rt}\right) \Rightarrow$$

$$\Rightarrow NC \cdot H - \left(\left(tt_{ot} + tt_{rt}\right) + \left(dt_{ot} + dt_{rt}\right) + \left(it_{ot} + it_{rt}\right)\right) > 2 \cdot H \Rightarrow$$

$$\Rightarrow \left(NC - 2\right) \cdot H > \left(tt_{ot} + tt_{rt}\right) + \left(dt_{ot} + dt_{rt}\right) + \left(it_{ot} + it_{rt}\right) \Rightarrow$$

$$\Rightarrow NC > \frac{\left(tt_{ot} + tt_{rt}\right) + \left(dt_{ot} + dt_{rt}\right) + \left(it_{ot} + it_{rt}\right)}{H} + 2 \Rightarrow$$

$$\Rightarrow \frac{(tt_{ot}+tt_{rt})+(dt_{ot}+dt_{rt})+(it_{ot}+it_{rt})}{H}+\frac{bt_{ot}+bt_{rt}}{H}+\frac{lt_{ot}+lt_{rt}}{H}>$$
$$>\frac{(tt_{ot}+tt_{rt})+(dt_{ot}+dt_{rt})+(it_{ot}+it_{rt})}{H}+2\Rightarrow$$

$$\Rightarrow \frac{bt_{ot}+bt_{rt}}{H}+\frac{lt_{ot}+lt_{rt}}{H}>2\Rightarrow$$

$$\Rightarrow \frac{bt_{ot}+lt_{ot}}{H}+\frac{bt_{rt}+lt_{rt}}{H}>2$$

which falls in contradiction with constraints (3.16) and (3.17), q.e.d.

Hence, it can be stated that it is always possible to determine a feasible set for α, which is expressed by equation (3.27).

Therefore, by properly setting α according to the specific examined context and related features in terms of energy consumption, specified for each direction (e.g. elevation profile), it is possible to identify the optimal allocation of layover times between the two terminal stations.

However, it is worth noting that the availability of a certain layover time is affected by the confidence level assumed for the computation of buffer times. Indeed, once fixed *NC*, *H* and *CT* (in terms of travel times, dwell times and inversion times), the reserve time is uniquely identified and has to be lower than *H*. Therefore, buffer times can be at most equal to the reserve time; otherwise, the solution is not feasible. In particular, only if buffer time is lower than reserve time, it is possible to have a layover time different from zero, which is clearly equal to the difference between reserve time and buffer time. Hence, in an energy saving perspective, the fact that the reserve time represents an upper bound for the buffer time and the trade-off between buffer and layover times have to be duly taken into account in the selection of the confidence level adopted for the computation of buffer times.

Other key parameters of the proposed approach are described in the following.

By considering equation (3.14), relation (3.20) becomes:

$$lt_{ot} + lt_{rt} = NC \cdot H - CT \tag{3.28}$$

while, by combining constraints (3.16) and (3.17), it is possible to obtain:

$$0 \le (bt_{ot} + bt_{rt}) + (lt_{ot} + lt_{rt}) \le 2 \cdot H \tag{3.29}$$

$$0 \le (lt_{ot} + lt_{rt}) \le 2 \cdot H - (bt_{ot} + bt_{rt}) \tag{3.30}$$

By substituting (3.28) into (3.30):

$$0 \le NC \cdot H - CT \le 2 \cdot H - (bt_{ot} + bt_{rt}) \tag{3.31}$$

$$\frac{CT}{H} \le NC \le \frac{CT + 2 \cdot H - (bt_{ot} + bt_{rt})}{H} = \frac{CT}{H} + 2 - \frac{bt_{ot} + bt_{rt}}{H} \tag{3.32}$$

Therefore, by considering that the value of *NC* has to be an integer:

$$int\left(\frac{CT}{H}\right) + 1 \le NC \le int\left(\frac{CT}{H} + 2 - \frac{bt_{ot} + bt_{rt}}{H}\right) \tag{3.33}$$

Hence, it can be stated that:

$$NC_{min} = int\left(\frac{CT}{H}\right) + 1 \tag{3.34}$$

$$NC_{max} = int\left(\frac{CT}{H} + 2 - \frac{bt_{ot} + bt_{rt}}{H}\right) \tag{3.35}$$

The previous equations allow to compute the maximum and minimum values of convoys for performing a rail service with certain features in terms of headway and cycle time.

Moreover, the time variation between the outward trip and the subsequent return trip, indicated as ΔT_{or}, may be calculated as:

$$\Delta T_{or} = \sum_{l_{ot}} tt_{l_{ot}} + \sum_{s_{ot}} dt_{s_{ot}} + it_{ot} + bt_{ot} + lt_{ot} \tag{3.36}$$

Likewise, the time variation between the return trip and the subsequent outward trip, indicated as ΔT_{ro}, may be formulated as:

$$\Delta T_{ro} = \sum_{l_{rt}} tt_{l_{rt}} + \sum_{s_{rt}} dt_{s_{rt}} + it_{rt} + bt_{rt} + lt_{rt} \tag{3.37}$$

Hence, by means of (3.36) and (3.37), equation (3.15) can be re-written as:

$$NC = \left(\Delta T_{or} + \Delta T_{ro}\right)/H \tag{3.38}$$

The time interval to achieve the regime condition may be expressed as follows:

$$\Delta T_{reg} = \left(NC - 1\right) \cdot H \tag{3.39}$$

Furthermore, the minimum headway depends on inversion times and the main features of the implemented signalling system, according to the following formulation:

$$H_{min} = \max\left\{ ts_{ot}^{inv}\,;\, ts_{rt}^{inv}\,;\, ts_{min-ss} \right\} \tag{3.40}$$

where H_{min} is the minimum value of H; ts_{ot}^{inv} is the time spacing to be respected during the inversion of the rail convoy at the final terminus of the outward trip; ts_{rt}^{inv} is the time spacing to be respected during the inversion of the rail convoy at the final terminus of the return trip; ts_{min-ss} is the minimum time spacing allowing by the implemented signalling system along the line, which has to take into account dwell times at stations and circulation rules such as the criterion of station releasing.

Obviously, the values of ts_{ot}^{inv} and ts_{rt}^{inv} depend on the infrastructure layout of the terminus. In general, they can include travel times, dwell times on inversion links (if any) and time rates related to the signalling system functioning such as, for instance, clearing times depending on train length and release times required for unlocking the block system (if the change of direction implies the passage through different block sections). As already mentioned, in certain cases, also the distance between stations can play a role in the definition of the minimum headway. Indeed, for safety reasons, especially in metro systems with a high degree of automation, in case of failure, trains have to be able to reach the subsequent station in order to provide passengers with suitable escape routes.

Lastly, it is worth pointing out that the increase in train running times generated by eco-driving strategies, besides affecting the service by an operational point of view as widely illustrated, implies an increase in passenger travel times and, therefore, a decrease in their satisfaction. For this reason, keeping faith with the passenger-oriented perspective adopted in this work, an important parameter to be considered is the user generalised cost (see equation 3.3) associated to each energy saving strategies. Therefore, the proposed multi-objective framework can be viewed as a tool for properly supporting the implementation of energy saving strategies so as to enable rail companies finding the right balance between reduction in energy consumption, timetable stability and passenger needs.

3.3.3 Modelling of the snowball effect

The snowball effect is due to the dynamic interaction between rail service and travel demand: the number of passengers on the platform influences the dwell times of trains at stations, which may cause delays; these, in turn, produce an increase in headways which generates more passenger flows on the platform providing a further extension of dwell times and, therefore, additional delays. Taking this phenomenon properly into account in the timetabling design phase is crucial to guarantee an appropriate degree of robustness of rail operations. In particular, a very critical task in order to design a stable timetable, which is able to absorb delays by avoiding disturbance propagation, is the estimation of dwell times as function of passenger flows involved in the boarding/alighting process. Therefore, this paragraph describes a simulation-based methodology for estimating dwell time as flow-dependent factors, which explicitly models passenger behaviour on platform when a train arrives as well as capacity constraints of convoys. This is possible by means of the implementation of the above described *OPM* which can implement different behavioural patterns.

In particular, in the following, two different assumptions on boarding priorities are modelled, namely the *First In – First Out* (FIFO) approach and the *Random In – First Out* (RIFO) approach, depicted respectively in figures 3.3 and 3.4.

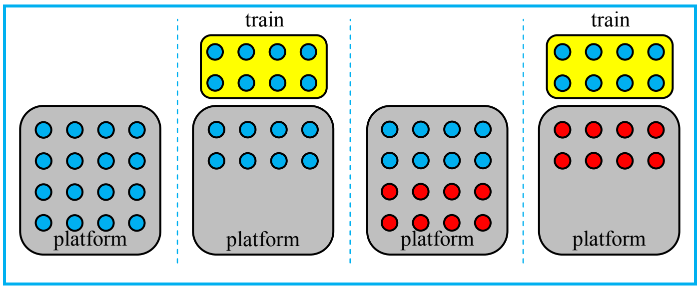


Figure 3.3 *First In – First Out* (FIFO) behavioural rule

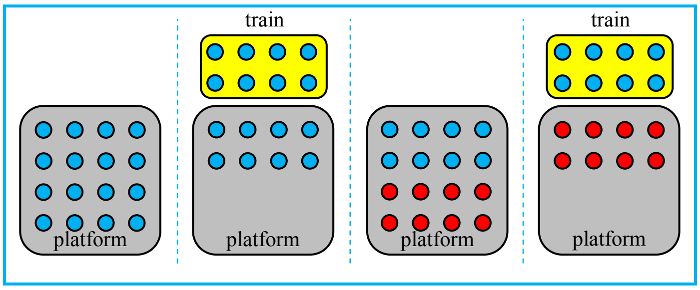


Figure 3.4 *Random In – First Out* (RIFO) behavioural rule

Specifically, a FIFO approach assumes that boarding order is related to the arrival order; this implies that a passenger may board a train only after all passengers arriving before him/her have boarded the train. On the other hand, a RIFO approach is based on the assumption that passengers waiting on the platform tend to move around by mixing with respect to their arrival order, thus altering the initial queuing pattern. In particular, we consider the maximum degree of mixing, which means that passengers are uniformly distributed on the platform with respect to the destination and arrival rates.

The assumptions adopted for both behavioural rules are set out below.

- Platforms are wide enough to host all incoming, waiting and outgoing passengers.
- The platform is uniquely determined once the run and the station have been fixed (i.e. trains travelling in the same direction always stop at the same platform).
- The dwell time of trains is constant (once the run, station and platform have been selected) and is independent of alighting and boarding flows.

- No interaction occurs on the platform among alighting, boarding and waiting passengers.
- The capacity of convoys is fixed, which means that the number of boarding passengers may be at most equal to the residual capacity.
- On-board passengers are uniformly distributed. This implies that all coaches of the same convoy have the same density and the remaining capacity is distributed uniformly among carriages. Moreover, an increase or a decrease in the number of passengers inside the train is equally distributed among coaches.
- No overlapping occurs in the train among alighting, boarding and on-board passengers, with the exception of the definition of residual capacity. This means that a different position (i.e. left or right) of platforms in subsequent stations does not influence the fluidity of passenger movements inside the coaches.

The analytical formulation of the addressed phenomenon is based on the following equations:

$$ar_{\Delta t}^{s,p,d} = \int_{\Delta t} f_{in}^{s,p,d}(\tau) \cdot d\tau \tag{3.41}$$

$$wp_r^{s,p,d} = \sum_{\Delta t=0}^{t(r,s,p)} ar_{\Delta t}^{s,p,d} - \sum_{i=1}^{r-1} bf_i^{s,p,d} \tag{3.42}$$

$$WP_r^{s,p} = \sum_d wp_r^{s,p,d} \tag{3.43}$$

$$RC_r^{s,p} = CAP_r - \sum_{i=1}^{s-1} \sum_d bf_r^{i,p,d} + \sum_{j=1}^{s} af_r^{j,p} \tag{3.44}$$

$$BF_r^{s,p} = \sum_d bf_r^{s,p,d} \tag{3.45}$$

$$BF_r^{s,p} = \begin{cases} WP_r^{s,p} & \text{if } WP_r^{s,p} \leq RC_r^{s,p} \\ RC_r^{s,p} & \text{if } WP_r^{s,p} > RC_r^{s,p} \end{cases} \tag{3.46}$$

where p is the generic platform which, according to the above assumptions, depends on run r and station s; $f_{in}^{s,p,d}(\tau)$ is the incoming passenger flow on platform p of station s, heading for destination d, in the time instant τ; Δt is the generic time interval; $ar_{\Delta t}^{s,p,d}$ is the incoming passenger flow (arrival rate) on platform p of station s, heading for destination d, during the time interval Δt; $bf_r^{s,p,d}$ is the boarding passenger flow on run r on platform p of station s, heading for destination d; t is the time interval when run r arrives on platform p of station s; $wp_r^{s,p,d}$ is the waiting passenger flow on platform p of station s, bound for destination d, when run r arrives; $WP_r^{s,p}$ is the waiting passenger flow on platform p of station s, bound for all destinations, when run r arrives; $af_r^{s,p}$ is the alighting passenger flow from run r on platform p of station s; CAP_r is the rail convoy capacity of run r; $RC_r^{s,p}$ is the residual capacity of run r when the train arrives at platform p of station s; $BF_r^{s,p}$ is the boarding passenger flow on run r on platform p of station s, bound for all destinations.

In particular, equation (3.41) expresses the arrival flow at a platform as the sum of incoming passengers; equation (3.42) provides waiting flows as the difference between arrival and boarding flows; equation (3.43) expresses the waiting flow bound for all destinations as the sum of waiting flows heading for each destination d; equation (3.44) provides the residual capacity as the rail convoy capacity minus the boarding flow plus the alighting flow; equation (3.45) expresses the boarding flow to all destinations as the sum of boarding flows to each destination d; finally, equation (3.46) calculates the significance of capacity constraints by expressing boarding flows as a function of waiting flows and residual capacities. Specifically, the last equation simulates the following phenomenon: if the waiting flow is at most equal to the residual capacity, all passengers are able to board the first arriving train; otherwise, only some are able to board, while the remaining passengers have to wait for the next trains.

Moreover, the FIFO approach can be modelled as follows:

$$x_r^{s,p}: \int_0^{x_r^{s,p}} \sum_d f_{in}^{s,p,d}(\tau) \cdot d\tau - \sum_{i=1}^{r-1} \sum_d bf_i^{s,p,d} = RC_r^{s,p} \tag{3.47}$$

$$bf_r^{s,p,d} = \int_0^{x_r^{s,p}} f_{in}^{s,p,d}(\tau)\cdot d\tau - \sum_{i=1}^{r-1} bf_i^{s,p,d} \tag{3.48}$$

where, in over-saturated conditions (i.e. when residual capacity is lower than the number of passengers waiting), it is necessary to calculate the time instant $x_r^{s,p}$ which satisfies equation (3.47) and allows the boarding flow to be calculated by means of equation (3.48).

On the contrary, equations modelling the RIFO approach are:

$$\alpha_r^{s,p} = \begin{cases} 1 & \text{if } WP_r^{s,p} \le RC_r^{s,p} \\ RC_r^{s,p} / WP_r^{s,p} & \text{if } WP_r^{s,p} > RC_r^{s,p} \end{cases} \tag{3.49}$$

$$bf_r^{s,p,d} = \alpha_r^{s,p,d} \cdot wp_r^{s,p,d} \tag{3.50}$$

where it is necessary to calculate the rate $\alpha_r^{s,p}$ which satisfies equation (3.49) and allows boarding flow to be calculated by means of equation (3.50).

It is worth noting that, in the case of under-saturated conditions, where the residual capacity is higher than the number of passengers waiting on platform and, therefore, all passengers are able to board the first arriving train, the two described approaches coincide. In particular, in this case, $x_r^{s,p}$ is equal to the arrival time of run *r* (i.e. time *t*) and $\alpha_r^{s,p}$ is equal to 1.

However, whatever the rule implemented, the proposed methodology for estimating dwell time as function of passenger flows, considers a threefold interaction:

1. as soon as the train arrives, passengers move towards the door they prefer;
2. when the capacity constraint of the single door is reached, passengers start moving to the next doors in the same coach;
3. once also the capacity constraint of the coach is reached, passengers move towards the other coaches which attract flow proportionally to their available capacity.

Some additional remarks on the above-mentioned levels of interaction are provided below.

Firstly, regarding the choice of the preferred door, it is assumed that it is made by passengers on the basis of their exit position (e.g. stairs or elevator) in the alighting stop, so as to minimise the walking distance at their own destination station (Kunimatsu et al., 2012). Indeed, especially commuters, who have experience of the system, know this information and exploit it to their advantage. This assumption is adopted only for initialising the loading algorithm which then converges according to the congestion level. In fact, in the cases of low-crowding conditions, each passenger is able to board quickly through the preferred door, without affecting the dwell time duration; while, in the cases of high-crowding conditions, since users aim to minimise the boarding time, their choices are dictated by the congestion of doors and coaches. Hence, different assumptions (e.g. distributing passengers uniformly on the platform) can be equivalently adopted without affecting simulation results. Moreover, given the lack of freedom of movement for passengers on-board, especially in crowded contexts, it can be stated that the door chosen to board the train will be the same as that to alight from. Regarding the other two interaction levels, they reflect the fact that, once the boarding through the preferred door has been missed, the aim of passengers is to get on the train as rapidly as possible, trying to remain close to the first favourite door. Therefore, users firstly try to board at least in the same coach but, if also this is not possible due to the high congestion, they are forced to settle for getting in the emptier coaches, independently of their position with respect to the exit position in the alighting stop.

Moreover, as already mentioned, the proposed methodology duly takes into account the capacity constraint of the entire convoy, so as to compute the number of passengers forced to remain on platform waiting for the next train in case the maximum capacity value is reached. This is a key issue for making simulation results accurate. Indeed, stranded passengers will be involved in the loading process of the following trains, affecting the relative dwell times, and, therefore, they cannot be neglected. Hence, the explicit simulation of capacity constraints has a crucial role in the proposed procedure. Moreover, since in this way it is possible to model the propagation of delays, the suggested method, with few adjustments, would be adopted also for addressing the implementation of rescheduling tasks in perturbed conditions.

The inputs required are passenger flows, station configurations and a function expressing the dependence of dwell times on the number of passengers at the most loaded door.

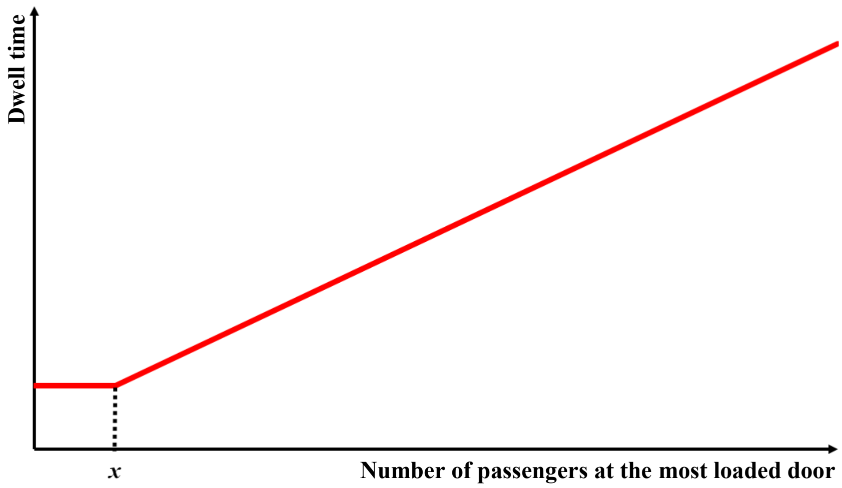


Figure 3.5 Dwell time calibration function

Qualitatively, such a function is always characterised by the same pattern (depicted in figure 3.5): it presents constant values of dwell time for low numbers of passengers, until a certain flow threshold (i.e. x), and, then, the dwell time increases as the number of passengers rises. However, since the factors involved in the definition of this function vary from case to case, it has to be properly calibrated according to the specific analysed context.

Hence, by replicating the boarding/alighting phase as a threefold interaction process, it is possible to estimate the number of passengers at the most loaded door and, consequently, the dwell time, for each simulated run and station, by means of the previously described function. In addition, given the accuracy of the implemented micro-simulation technique, it is possible to carry out the crowding level within each coach (figure 3.6).

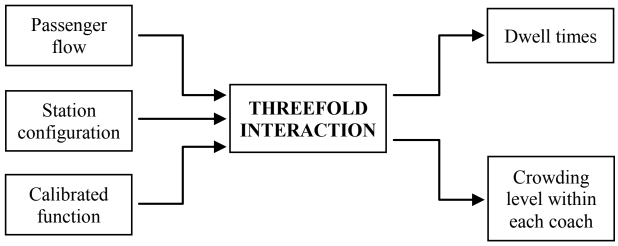


Figure 3.6 Simulation architecture

This is a very useful information, not only for estimating passenger comfort on-board, but also for supporting decision tasks of train operating companies as, for instance, the definition of a proper fleet composition in order to meet travel demand requirements. Moreover, Intelligent Transportation Systems (ITS) could be implemented with the aim of assisting passengers during the boarding/alighting process (e.g. by suggesting them the best position to be taken along the platform or which coach should be preferred, according to the crowding conditions on the approaching train). In this way, it would be possible to make boarding operations smoother thus reducing dwell times.

Analytically, the snowball effect generated by the dynamic interaction between headways and dwell times is modelled by means of a fixed-point problem formulation.

For this purpose, let

$$\boldsymbol{dwt} = \vartheta(\boldsymbol{td}) \tag{3.51}$$

be a function which expresses the dependence of dwell times on the number of boarding/alighting passengers, where $\boldsymbol{dwt}$ and $\boldsymbol{td}$ represent, respectively, dwell time and travel demand vectors. Obviously, function $\vartheta(\cdot)$ has to consider that there is a threshold value of boarding/alighting passenger flow below which the dwell time is constant.

Likewise, let

$$\boldsymbol{hd} = \psi(\boldsymbol{dwt}) \tag{3.52}$$

be the relation providing headways (i.e. vector $\boldsymbol{td}$) as function of dwell times (i.e. vector $\boldsymbol{dwt}$), by means of the simulation performed by the *SeSM*.

Since the frequency of a metro rail service strongly affects the congestion level on the platform, assuming within a short time interval the arrival rate of passengers at station *s* as constant, the travel demand (i.e. the number of passengers waiting on the platform) at each station *s* may be calculated as:

$$td_{s,d,r} = upf_{s,d,r} \cdot hd_{r,s} \tag{3.53}$$

where $td_{s,d,r}$ is the number of passengers arriving at the platform of station *s* for travelling towards destination *d* during the time interval between run (*r*–1) and run *r*; $upf_{s,d,r}$ is the arrival rate of passengers at platform of station *s* for travelling towards destination *d* during time interval between run (*r*–1) and run *r*; $hd_{r,s}$ is the headway between run (*r*–1) and run *r* at station *s*. Obviously, $td_{s,d,r}$ and $hd_{r,s}$ express, respectively, the component of vector $\boldsymbol{td}$ and vector $\boldsymbol{hd}$; while the arrival rate $upf_{s,d,r}$ is provided by the *PPM* for each station and each time interval between two successive runs.

Hence, equation (3.53) may be expressed in vector notation as:

$$\boldsymbol{td} = \gamma(\boldsymbol{hd}) \tag{3.54}$$

Therefore, by combining the equations above:

$$\begin{cases} \boldsymbol{dwt} = \vartheta(\boldsymbol{td}) \\ \boldsymbol{td} = \gamma(\boldsymbol{hd}) \\ \boldsymbol{hd} = \psi(\boldsymbol{dwt}) \end{cases} \tag{3.55}$$

or similarly:

$$\boldsymbol{dwt} = \vartheta(\gamma(\psi(\boldsymbol{dwt}))) \tag{3.56}$$

Specifically, $\vartheta(\cdot)$ represents the above mentioned function to be calibrated according to the analysed context (figure 3.5), or its indirect formulations; $\gamma(\cdot)$ has been already made explicit in equation (3.53); $\psi(\cdot)$, on the contrary, cannot be expressed in a closed form, since it represents a system of differential equations which requires being solved numerically, by means of a suitable simulation software.

It is worth pointing out that, in general, it is not possible to state that higher flows (in terms of arrival rates) necessarily imply higher dwell times. Indeed, the above described analytical framework findings show that, actually, dwell time in a station is a function of the arrival rate in that station, the arrival rates in the previous stations and the framework of travel demand (in terms of alighting flows). In particular, equation (3.51) shows the direct dependence of dwell times on travel demand (i.e. the higher the travel demand, the higher the dwell time), equation (3.53) (or, equivalently, equation 3.54) shows the direct dependence of travel demand on headway which, by means of equation (3.52), in turn depends on dwell times in previous stations. This corroborates the importance of adopting proper simulation techniques in order to capture the development of the complex cooperation and negotiation process among passengers during the boarding/alighting process.

According to the theory of the fixed-point problem, system of equations (3.55) represents a *compound fixed-point problem* in which it is necessary to find a dwell time vector which provides a headway vector which produces travel demand on platform which, in turn, generates the initial dwell time vector. The conditions ensuring the existence and the uniqueness of the solution of a fixed-point problem, described in paragraph 2.1.6, can be easily extended to the compound fixed-point problem, as shown by Cascetta, 2009.

In particular:

1) the functions involved in system of equations (3.55) (i.e. $\vartheta(\cdot)$, $\gamma(\cdot)$ and $\psi(\cdot)$) are continuous;

2) the compound function $\vartheta(\gamma(\psi(\cdot)))$ is defined in the set $\boldsymbol{S_{dwt}} = \{\boldsymbol{dwt}\}$ (where $\boldsymbol{dwt} = [dwt_1, ..., dwt_i, ..., dwt_n]^T$) with values in the set $\boldsymbol{T} = \vartheta(\gamma(\psi(\boldsymbol{S_{dwt}}))) \subseteq \boldsymbol{S_{dwt}}$;

3) the definition set $\boldsymbol{S_{dwt}}$ is:

- nonempty:

  $\boldsymbol{S_{dwt}} \neq \varnothing$ since $\forall i \; \exists dwt_i \geq 0$

- compact:

  $dwt_i \in [0; \max(dwt_i)] \; \forall i$

- and convex:

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

Therefore, Brouwer's theorem (Brouwer, 1912) is satisfied and, hence, it is possible to state that the snowball effect does not evolve indefinitely, but converges towards an equilibrium state.

On the contrary, conditions ensuring the uniqueness of the solution of a fixed-point problem, provided by Banach's theorem (Banach, 1992), are not fulfilled, since not all involved functions satisfy the properties of monotonicity. Clearly, this affects the selection of the resolution method to be adopted. In particular, as already mentioned, fixed point problems are generally solved by means of the *MSA* algorithm (described in paragraph 2.4), whose convergence is ensured by Blum's theorem (Blum, 1954). However, in this case, it cannot be applied, since the uniqueness of the solution cannot be demonstrated. Therefore, it is necessary to find a numerical evidence for assuring the convergence of the algorithm or rely on alternative resolution procedures, as will be shown in paragraph 4.4 where the proposed framework will be implemented in the case of a real metro context.

3.3.4 Travel demand estimation

This paragraph addresses the travel demand estimation problem, starting from the traditional techniques proposed in the literature and customising them to the specific features of rail operations. In particular, the procedure of concern is the aggregate estimation which consists in updating/adjusting a prior known O-D matrix by means of aggregate data such as passenger counts.

The fact that the flows to be collected are related to passengers, rather than to vehicles, leads to a first issue to be properly addressed, that is the category of passengers to be detected for each specific assessment. Indeed, different types of passenger flows are involved in the analysis of a rail system such as flows at turnstiles, boarding or alighting flows, waiting flows and on-board flows. This entails a *spatial* problem related to 'where' to count passengers. Specifically, if we were to count passengers at the turnstiles, there would be a resulting degree of uncertainty as to users' direction. Alternatively, another option is to obtain information on each single gate, but, in this case, the measure would not take into account if any and how many passengers are not able to board the train due to overcrowding. The latter information would be available, on the contrary, if the calculation is carried out on the platform. Additionally, there is a *temporal* problem to be considered, which lies in the difficulty of identifying a suitable degree of aggregation because of the discontinuous fruition which is offered by rail service. It is this discontinuity, which, for instance, compromises the take-over at turnstiles because of the gap between the moment when the users' passage is recorded and the moment when users achieve the platform and, therefore, are actually able to board the arriving train. Hence, in the light of the above, it appears obvious that, according to the target, it is necessary to design and execute the counting phase adequately and in the most reliable way.

Differently from sample surveys, which are quite complex and have high costs, counts do not require excessive expenses and can be obtained automatically. The use of automatic devices allows to perform the counting phase in an easier and more efficient manner; however, it is not exempt from incidents. First of all, in the event of a problem with the equipment, or parts of it, the whole measurement process would be compromised. For example, if we were to estimate the distribution of users on the platform with the intention to perform detection at gates and one of the gate detectors is damaged, this would make detection at gates useless, with the consequent loss of the information on the entire platform. Moreover, the possible presence of internal interchanges between lines makes the system not perfectly enclosed. Finally, also episodes of evasion which, unfortunately, occur in some circumstances, could distort the outcome.

Therefore, in the following, two methodological frameworks, which duly take into account the above-mentioned points, are illustrated. Specifically, the first one concerns an analytical procedure for extending passenger counts by means of the calibration of suitable space-time functions which allow to reduce the sampling rate without compromising estimation accuracy. The second proposal consists in a long-term estimation technique, as a support tool for performing cost-benefit analyses, which allows to properly model changes in travel demand due to demographic and transportation system variations in a wide time period (i.e. several decades).

3.3.4.1 Analytical methodology for extending passenger counts

The relevance of the proposed approach lies in the fact that the greater the number of detected data, the greater the accuracy of travel demand estimations, but also the cost and times which will incur. Hence, the necessity of finding a fair compromise between survey costs and estimation accuracy is imperative. In this context, the presented proposal is based on the development of an analytical procedure aimed at reducing the number of data to be collected, without significantly affecting estimation accuracy. This is possible by identifying some space-time relations properly calibrated for providing flows values such that minimise the error in the aggregate estimation of the O-D matrix and, therefore, in the computation of system performance made by assigning it to the analysed network. In other words, the parameter to be minimised is the gap between assignment results obtained by implementing, on one side, the O-D matrix adjusted with detected flows and, on the other, the O-D matrix adjusted with flows provided by the calibrated functions (or, alternatively, with a mixed-flow data set, i.e. made up partly with detected flows and partly with analytical flows). Indeed, the O-D matrix is not the ultimate outcome desired, but only a means for enabling an assessment of system performance, thanks to its assignment to the network, as explained in paragraph 2.1.6.

In particular, these analytical relations express boarding and alighting flows depending on the station (space component) and the time period (time component) considered. Therefore, the basic assumption is that spatial correlation (i.e. the correlation among different stations) and temporal correlation (i.e. the correlation among different time periods) of passenger flows are generally not null. Moreover, it is worth pointing out that the proposed approach has a purely descriptive nature, without any explicit assumption on user behaviour.

Specifically, the developed analytical procedure foresees the phases set out below.

Firstly, it is necessary to investigate the system under examination in detail, by collecting information concerning number and location of stations, their layout in terms of platforms and gates, rolling stock features, operating hours and timetables. This preliminary phase allows to plan 'when' and 'where' passengers flows must be detected, and, clearly, has to be followed by the actual execution of the designed survey campaign. The goal is to collect a proper amount of data for applying statistical analyses described in the following. In particular, the ideal condition is represented by the possibility of performing an exhaustive counting (i.e. the adoption of a sampling rate equal to 1) so as to identify a reference scenario which can be considered as the '*absolute truth*'.

In a metro system, an exhaustive survey consists in acquiring boarding and alighting flows in all stations for each direction and for each considered time period. However, this is possible only for small-size networks or by means of laboratory experiments on real-size synthetic networks, as proposed by Marzano et al. (2009).

The collected data can be considered, for illustrative purposes, organised as shown in figure 3.7.

		Station					
		st_1	st_2	…	st_i	…	st_m
Time period	tp_1	$f_{1\text{-}1}$	$f_{1\text{-}2}$	…	$f_{1\text{-}i}$	…	$f_{1\text{-}m}$
	tp_2	$f_{2\text{-}1}$	$f_{2\text{-}2}$	…	$f_{2\text{-}i}$	…	$f_{2\text{-}m}$
	…	…	…	…	…	…	…
	tp_n	$f_{n\text{-}1}$	$f_{n\text{-}2}$	…	$f_{n\text{-}i}$	…	$f_{n\text{-}m}$

Figure 3.7 Surveyed data (i.e. real surveyed data)

The second phase consists in simulating the adoption of a certain sampling rate, lower than 1, by hiding (i.e. assuming not detected) some data and, thus, obtaining a partial data set to be analysed (figure 3.8).

		Station					
		st_1	st_2	…	st_i	…	st_m
Time period	tp_1	$f_{1\text{-}1}$	*missing*	…	$f_{1\text{-}i}$	…	*missing*
	tp_2	*missing*	$f_{2\text{-}2}$	…	*missing*	…	$f_{2\text{-}m}$
	…	…	…	…	…	…	…
	tp_n	$f_{n\text{-}1}$	*missing*	…	$f_{n\text{-}i}$	…	*missing*

Figure 3.8 Partial set of surveyed data (i.e. simulated surveyed data).

It is worth noting that, the criterion adopted, in the simulation phase, for choosing the data to be assumed ‘not detected’ has a key role, since, by moving to real applications of the proposed approach, it replicates, somewhat, the decision-making process aimed at selecting (according to the adopted sampling rate) which data to be acquired and which ones to be neglected. Therefore, a certain degree of uniformity in space and time has to be pursued, so as to make the following interpolation steps as more accurate as possible.

Once the data have been properly collected, analysed and processed, a first statistical analysis can be performed, which is based on a mono-dimensional approach. Specifically, it involves the partial data set identified in the previous phase and consists in determining the class of functions (e.g. linear, quadratic, cubic, polynomial) which best describes the simulated survey data. This procedure is indicated as mono-dimensional because the involved functions are defined in a R^2 space, where the abscissa is the sequence of stations or the time periods and the ordinate is the surveyed flow (as shown in figure 3.9).

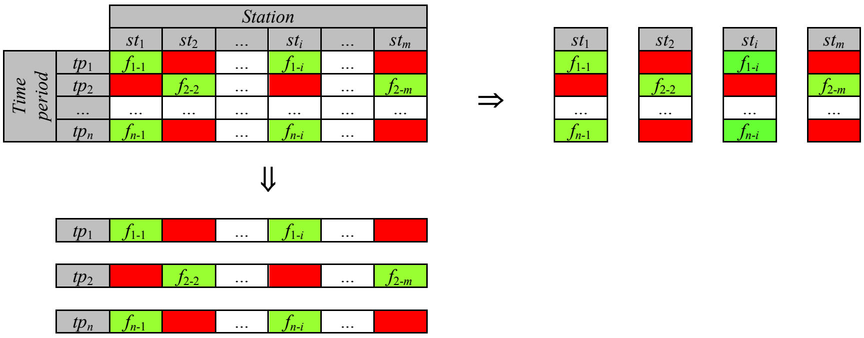


Figure 3.9 Organisation of data for mono-dimensional analyses

As already mentioned, the flows of concern are related to boarding and alighting passengers in both directions for each station and time period detected; therefore, assuming n_{fc} as the number of function classes to be analysed, it can be stated that it is necessary to calibrate and validate n_f mono-dimensional functions. Specifically:

$$n_f = n_{fc} \cdot 2 \cdot (n_{st} \cdot 2) \cdot n_{tp} \tag{3.57}$$

where n_{st} is the number of the stations (multiplier 2 for considering outgoing and return trips separately) and n_{tp} is the number of time periods considered. The quantity ($n_{st} \times 2$) is further multiplier 2 for taking into account both boarding and alighting flows.

The goodness of fit of each class of function has to be properly evaluated. For this purpose, generally, the simplest and the most frequently used parameter is the coefficient of determination, $\Re^2$ expressed as follows:

$$\Re^2 = \left(\sum_i \left(f_i - \bar{\varphi} \right)^2 \right) \Big/ \left(\sum_i \left(\varphi_i - \bar{\varphi} \right)^2 \right) \tag{3.58}$$

with:

$$\bar{\varphi} = \sum_i \varphi_i / n \tag{3.59}$$

where φ_i is the i-th simulated survey data (i.e. known value of figure 3.9), n is the number of simulated survey data φ_i, $\overline{\varphi}$ is the mean of data φ_i and f_i is the i-th value assumed by the calibrated function.

However, since the value of $\Re^2$ tends to increase with the introduction of additional predictors, usually, to penalise this effect, it is appropriate to calculate also the adjusted $\Re^2$ (indicated as $\overline{\Re}^2$), by means of the following equation:

$$\overline{\Re}^2 = \Re^2 - \left(1 - \Re^2\right) \cdot p / (n - p - 1) \tag{3.60}$$

where p expresses the number of function parameters.

Once the optimal functional form has been properly identified in both dimensions, it is possible to execute a multi-dimensional statistical analysis which involves the same data set of the mono-dimensional approach and consists in specifying (according to the optimal classes of functions previously identified), calibrating and validating, with suitable statistical tests (both global and on single coefficients), four different surfaces. The number four is due to the necessity of considering two kinds of passenger flows (boarding and alighting flows) and two kinds of trips (outgoing and return trips). In particular, the stations and the time periods are the independent variables, while the surface provides the value of flow.

Obviously, according to the outcome of the specification phase, the calibration step has to be performed with different statistical techniques such as simple linear regression, multiple linear regression, polynomial regression. Moreover, whenever there is the necessity of simulating different levels of travel demand, it is possible to rely on a quantile regression technique.

The last phase consists in comparing the application results obtained by using the whole set of the survey data (considered as the *absolute truth*) and those using the data of calibrated space-time surfaces, properly put together with the data of calibration subsets. Specifically, within this framework, three different data sets may be obtained for comparison: only the calibration subset (already shown in figure 3.8), the calibration subset extended by replacing missing data with function data (depicted in figure 3.10) and only function data for all values (indicated in figure 3.11).

		Station					
		st_1	st_2	...	st_i	...	st_m
Time period	tp_1	$f_{1\text{-}1}$	*function*	...	$f_{1\text{-}i}$	...	*function*
	tp_2	*function*	$f_{2\text{-}2}$	...	*function*	...	$f_{2\text{-}m}$
	...	...	...	...	...	...	...
	tp_n	$f_{n\text{-}1}$	*function*	...	$f_{n\text{-}i}$	...	*function*

Figure 3.10 Subset extension by means of function data

		Station					
		st_1	st_2	...	st_i	...	st_m
Time period	tp_1	*function*	*function*	...	*function*	...	*function*
	tp_2	*function*	*function*	...	*function*	...	*function*
	...	...	...	...	...	...	...
	tp_n	*function*	*function*	...	*function*	...	*function*

Figure 3.11 Function data for all values

Therefore, it is possible to implement an aggregate estimation of travel demand, by adjusting a prior-known O-D matrix according to the four data sets identified. In this way, four different O-D matrices can be derived and assigned to the network, so as obtained objective function values for each one of the four analysed cases. Hence, it is possible comparing assignment results obtained by means of the whole set with those provided by means of the other three data sets, in order to evaluate which one of them produced an outcome closer to that of the reference scenario. In particular, a small variation in the objective function with respect to the '*absolute truth*' confirms the ability of the calibrated surfaces of capturing the space-time variations of travel demand and, therefore, their usefulness in allowing a reduction of the data to be acquired during the survey phase, without prejudicing the analysis accuracy.

This implies the possibility of cutting the budget to be allocated for the survey phase but, this is not the only benefit. Indeed, such a procedure allows to analyse also networks which, due to their complexity, do not enable the achievement of a reasonable sampling rate, because this would result as uneconomic. Moreover, the possibility of replacing some missing data by means of analytical functions offers the chance of rectifying eventual inconveniences due to a failure in the detection equipment very smoothly. This is very useful, for instance, in the cases in which the lack of even a single information can compromise all other measures such as when, given the target of reconstructing the flows on platforms by means of gate counts, the data of a gate are lost.

The proposed approach lends itself to several improvements. Firstly, for the identification of stations, different spatial reference systems, such as curvilinear abscissa and polar coordinates, rather than a simply sorting technique (i.e. sorting them according to the train route), can be implemented. Moreover, the relations to be calibrated can be enhanced with additional explanatory variables, such as interchanges with other public transit systems, possibility of parking etc. In particular, the stepwise regression can be adopted by implementing a forward selection, a backward elimination or a combination of them. A forward selection consists in starting with no variables in the model and, progressively, adding predictors which satisfy a certain fit criterion; on the contrary, a backward elimination consists in starting with all candidate variables and, progressively, deleting predictors whose explanatory power is not relevant for the fit accuracy of the model. Both techniques proceed until no further improvements can be reached. The fit criterion in the analysed framework is related to the ability of the model in reproducing detected flows. Finally, the proposed analytical approach, with proper adjustments, could be implemented for performing a multi-seasonal passenger flows estimation. The idea behind is to make a model relative to a certain time period (e.g. holidays), representative of another one (e.g. working days), by means of the introduction of conversion coefficients to be properly calibrated so as to capture the eventual correlation between travel demand patterns in different time periods. The goal is ambitious and requires more detailed evaluations, above all for verifying whether the same functional form can fit different time periods or not and, consequently, the necessity of introducing some behavioural assumptions (e.g. by adopting conversion functions rather than simple coefficients). The challenge is still open; however, the above described proposal somewhat further confirms the power of such an analytical tool for managing passenger counts in the case of travel demand estimation techniques.

3.3.4.2 A long-term evaluation of travel demand

As already pointed out, the evaluation of travel demand has a key role in any assessment concerning transportation systems. In particular, the estimation of passenger flows, in current and future conditions, is required in the case of a cost-benefit analysis related to each kind of long-term measure such as infrastructural interventions (new lines or modification of existing lines), fleet improvements (partial or complete replacement of rolling stock) and signalling system modifications (replacement or upgrade of trackside and on-board equipment). Indeed, such an analysis cannot be separated from the computation of travel demand in terms of potential or expected passengers with related characteristics (i.e. starting and arrival stations, adopted time slot, trip duration, etc.). Moreover, users and their needs represent a fundamental element in an economic evaluation and, therefore, their standpoint cannot be neglected.

Additionally, in order to evaluate and compare different intervention scenarios within a cost-benefit analysis, the estimation demand model has to be elastic at least at the level of modal choice (in the case of transportation system modifications) and trip generation (in the case of demographic changes). For this purpose, it is necessary to ensure an accurate representation of the current situation and a reliable prediction of future conditions, as well as the modelling of travel demand as a random variable (i.e. not only average values but also their distributions have to be considered).

Therefore, this paragraph presents a comprehensive procedure for evaluating travel demand in contexts where the long-term estimation is the major requirement. Specifically, it is based on the use of different Italian data sources; however, a generalisation to different contexts can be simply achieved.

The first step is based on the use of data from national census, reported in the *ISTAT* (*Italian National Institute of Statistics*) database, which provide revealed information (i.e. related to behaviour actually occurring in the days prior to the survey) concerning mobility choices in terms of origin, destination, daily time period and transport mode. More specifically, these data are structured as follows:

- the considered trips are the systematic ones (i.e. for work or school purposes) during the average working day;
- origins and destinations are expressed in terms of municipalities;
- daily times are indicated as the morning peak hour (i.e. from 7.30 to 9.29) and the rest of the day;
- only outward trips are provided, since trips are generally bidirectional (i.e. from home to the workplace and return).

Clearly, it is necessary to extract from the entire database only the information relative to the study area, with the aim of identifying data concerning internal trips (i.e. with origin and destination both in the study area) and exchange trips (i.e. with the origin or the destination in the study area).

In particular, in order to increase the examined dataset and, therefore, meet the need of considering a certain distribution for travel demand values, it is necessary taking into account statistics from, at least, two decades (i.e. data from the 2001 and 2011 Italian censuses).

In the *ISTAT* database, it is also possible to find historical information as well as projections relative to population data (according to three different variation rates: minimum, average, maximum) which are crucial for making the demand elastic at level of trip generation, as will be shown shortly. In particular, demographic forecasts are performed by means of the *cohort component method* which considers death, births and migration as factors of concern. Generally, according to this approach, the population expected to be alive at the end of the projection period is obtained by multiplying base census population, of a given age group, by a certain survival rate; while, the number of births taking place in the projection period is computed by multiplying an age-specific fertility rate by the number of women in their reproductive years. Finally, it is necessary to add the number of net migrants. This is possible by means of a two-step procedure: firstly net migration rates are determined and then multiplied by the surviving population.

The second phase relies on data computed by mobility observatories (such as, for instance, *AudiMob – Observatory on the Italian mobility behaviour*) which provide additional useful information such as total daily regional trips, rates of trips during morning peak hours, rates of trip chains (i.e. trips with intermediate destinations) and regional modal split. These data allow to derive non-systematic trips during the average working day, categorised according to origin and destination municipality, time period (i.e. peak hour or rest of the day), adopted transport mode and reference year (i.e. 2001 or 2011).

At this point, it is necessary to project systematic and non-systematic trips from the census year to a successive period. Generally, historical or forecasted data can be adopted, according to the target period: historical data until a year before the current year and forecasted data for the current year and successive years. However, since the following phases foresee an adjustment of the O-D matrices with passenger counts, in this stage, it is assumed that only historical data are exploited for the projection. In particular, it is necessary to adopt an increase or decrease rate equal to population variation (i.e. a variation in α% of population in municipality A provides a variation in α% of all trips with origin in A). This allows to make the trip generation model elastic.

In the following step, it is necessary to convert data concerning systematic and non-systematic trips into travel demand matrices related to all-day trips and in which the origin and destination are expressed in terms of stations of the rail line in question, rather than in terms of municipalities.

Therefore, firstly, round trips from outward trips have to be carried out as follows:

$$OD_{rt}^{i,m} = OD_{ot}^{i,m} + \left[OD_{ot}^{i,m}\right]^{\mathrm{T}} \tag{3.61}$$

where $OD_{rt}^{i,m}$ is the origin-destination matrix related to round trips (*rt*) throughout the day associated to purpose *i* (i.e. systematic or non-systematic) and mode *m*; $OD_{ot}^{i,m}$ is the origin-destination matrix related to outward trips (*ot*) all day associated to purpose *i* and mode *m*; $\left[OD_{ot}^{i,m}\right]^{\mathrm{T}}$ is the transposed matrix of $OD_{ot}^{i,m}$.

Then, for switching from trips expressed in terms of origin and destination municipalities to trips expressed in terms of origin and destination stations, it is necessary to develop a regional network model which, by means of the implementation of a minimum path approach, allows to match each municipality to each station. In particular, the basic rules followed for municipalities within the study area are: if there are no stations in the municipality, we associate the nearest station; if there is only one station in the municipality, we of course associate that station; finally, if there are two or more stations in the municipality, we hypothesise some distribution coefficients (for instance, related to turnstile counts). On the contrary, for municipalities without the study area, which are involved in the exchange trips, it is necessary to analyse the presence of interchange stations with the line under examination, if any.

The phase that follows consists in adjusting origin-destination matrices associated to the rail mode (*r*) according to aggregated information, represented by turnstile counts. Obviously, the matrices involved in the correction procedure have to be referred to the same year in which traffic counts have been carried out. Moreover, since this kind of counts is generally aggregated in a daily scale, we propose to correct the all-day matrices and then modify initial matrices by adopting the same variation rates. This implies assuming the total travel demand as constant and considering differences as due to a different modal split. Hence, the following equations have to be implemented:

$$OD_{rt}^{r} = \sum_{i} OD_{rt}^{i,r} \tag{3.62}$$

$$\overline{OD_{rt}^{r}} = \arg\min_{x \geq 0} Z\left(d_1\left(x \; ; \; OD_{rt}^{r}\right); d_2\left(\Lambda(x) \; ; \; f^{r}\right)\right) \tag{3.63}$$

$$\delta_j = \overline{d_{j,rt}^{r}} \Big/ d_{j,rt}^{r} \quad \text{with } \overline{d_{j,rt}^{r}} \in \overline{OD_{rt}^{r}} \text{ and } d_{j,rt}^{r} \in OD_{rt}^{r} \tag{3.64}$$

$$\overline{OD_{h}^{i,r}} = \left\{ \overline{d_{j,h}^{i,r}} \; : \; \overline{d_{j,h}^{i,r}} = d_{j,h}^{i,r} \cdot \delta_j \right\} \qquad \forall h \in [ph \; ; \; ad] \quad \forall i \in [s \; ; \; ns] \tag{3.65}$$

$$d_{j,h}^{i} = \sum_{m} d_{j,h}^{i,m} \tag{3.66}$$

$$\overline{OD_{h}^{i,m}} = \left\{ \overline{d_{j,h}^{i,m}} \; : \; \overline{d_{j,h}^{i,m}} = d_{j,h}^{i,m} \cdot \left(d_{j,h}^{i} - \overline{d_{j,h}^{i,r}} \right) \Big/ \sum_{m \neq r} d_{j,h}^{i,m} \right\} \\ \forall m \neq r \quad \forall h \in [ph \, ; ad] \quad \forall i \in [s \, ; ns] \tag{3.67}$$

where OD_{rt}^{r} is the origin-destination matrix related to all-day round trips (*rt*) associated to mode *r* (i.e. rail mode), $\overline{OD_{rt}^{r}}$ is the correction of matrix OD_{rt}^{r}; *x* is the variable expressing in the optimisation problem (3.63) the generic value of matrix OD_{rt}^{r}; $Z(\cdot)$ is the objective function to be minimised; d_1 is a function which expresses the distance between matrix *x* and the a-priori estimation of matrix OD_{rt}^{r}; $\Lambda(\cdot)$ is the assignment function which provides passenger flows associated to origin-destination matrix *x*; f^{r} is the vector of turnstile counts; d_1 is a function which expresses the distance between flows obtained by assigning matrix *x* and flows provided by turnstile counts (i.e. f^{r}); δ_j is the variation rate of travel demand associated to origin-destination *j*; $\overline{d_{j,rt}^{r}}$ is the generic element of matrix $\overline{OD_{rt}^{r}}$; $d_{j,rt}^{r}$ is the generic element of matrix OD_{rt}^{r}; $\overline{OD_{h}^{i,m}}$ is the corrected origin-destination matrix in the time period *h*, for purpose *i* by using mode *m*; $\overline{d_{j,h}^{i,m}}$ is the generic element of matrix $\overline{OD_{h}^{i,m}}$ associated to origin-destination *j*; $d_{j,h}^{i,m}$ is the a-priori estimation of trips in the case of origin-destination *j*, in the time period *h*, for purpose *i* by using mode *m*; $d_{j,h}^{i}$ is the a-priori estimation of trips in the case of origin-destination j, in the time period h, for purpose i by using all transportation mode.

It is worth noting that since variable i, expressing the purpose of the trip, may assume s for systematic and ns for non-systematic and variable h, expressing time period, may assume ph for the morning peak hour and ad for all day, it is possible to state that, in the case of corrected matrices, equation (3.61) becomes:

$$\overline{OD_{rt}^{i,m}} = \overline{OD_{ad}^{i,m}} + \left[\overline{OD_{ad}^{i,m}}\right]^{\mathrm{T}} \tag{3.68}$$

where $\overline{OD_{rt}^{i,m}}$ is the corrected origin-destination matrix related to all-day round trips (rt) associated to purpose i and mode m.

In particular, equation (3.67) expresses the necessity of properly re-calibrating matrices related to the other considered transport modes, arising from the variation in the rail matrix due to the updating procedure.

At this stage, corrected matrices have to be extended to one or more analysis periods (generally the projection horizon is several decades), by applying historical and/or forecasted demographic variation rates according to the already mentioned principles.

The following step aims to make demand elastic at least at modal choice level. For this purpose, it is necessary to specify, calibrate and validate a suitable choice model, as described in paragraph 2.3 with regard to the disaggregate estimation techniques. Specifically, it is required to:

- specify a utility formulation and a probability choice model such as:

$$V_{j,h}^{i,m} = V_{j,h}^{i,m}\left(\beta_k^m\right) = \sum_k \beta_k^m X_{k,j,h}^{i,m} \tag{3.69}$$

$$p_{j,h}^{i}[m] = p_{j,h}^{i}[m]\left(V_{j,h}^{i,m}\left(\beta_k^m\right)\right) = p_{j,h}^{i}[m]\left(\beta_k^m\right) \tag{3.70}$$

  where $V_{j,h}^{i,m}$ is the utility associated to mode m in the case of purpose i during the time period h for travelling between the origin-destination j; β_k^m is the parameter associated to k-th attribute of the mode m ; $X_{j,h}^{i,m}$ is the k-th attribute associated to mode m in the case of purpose i during the time period h for travelling between the origin-destination j; $p_{j,h}^{i}[m]$ is the probability of choosing mode m for travelling between the origin-destination j in the case of purpose i during the time period h;

- calibrate the values of parameters β_k^m by means of the following optimisation problem:

$$\hat{\beta}_k^m = \arg\max_{\beta_k^m}\left(ln\, L\left(\hat{\beta}_k^m\right)\right) = \sum ln\left(p_{j,h}^{i}[m]\left(\hat{\beta}_k^m\right)\right) \tag{3.71}$$

where $\hat{\beta}_k^m$ is a calibrated value of parameter β_k^m; $L(\cdot)$ is the likelihood function to be maximised;

- validate the results by means of suitable statistical tests.

The explicit procedure for computing the variation of $p_{j,h}^{i}[r]$, due to the implementation of the design scenario to be evaluated, is set out below. For the sake of simplicity, the following assumptions are adopted:

1) the purpose and the origin-destination pair, as well as the time period of the trip, are pre-fixed;
2) the considered design intervention operates only on the rail mode;
3) the adopted random utility model is the Multinomial Logit;
4) parameter θ of the Multinomial Logit is assumed included into parameter β_k^m.

Therefore, let:

$p'[r]$ be the choice probability of the mode r, *before* the implementation of the intervention;

$p''[r]$ be the choice probability of the mode r, *after* the implementation of the intervention;

$p'[m]$ be the choice probability of the mode *m*, *before* the implementation of the intervention;

$p''[m]$ be the choice probability of the mode *m*, *after* the implementation of the intervention;

$V'[r]$ be the utility associated to the mode *r*, *before* the implementation of the intervention;

$V''[r]$ be the utility associated to the mode *r*, *after* the implementation of the intervention;

$V'[m]$ be the utility associated to the mode *m*, *before* the implementation of the design intervention

$V''[m]$ be the utility associated to the mode *m*, *after* the implementation of the intervention.

According to assumption 2):

$$V'[r] \neq V''[r] \tag{3.72}$$

$$V'[m] = V''[m] = V[m] \quad \forall m \neq r \tag{3.73}$$

Furthermore, according to assumptions 3) and 4):

$$p'[r] = \frac{exp(V'(r))}{\sum_{m \neq r} exp(V(m)) + exp(V'(r))} \tag{3.74}$$

$$p''[r] = \frac{exp(V''(r))}{\sum_{m \neq r} exp(V(m)) + exp(V''(r))} \tag{3.75}$$

According to the probability theory, which states that the sum of probabilities in the sample space is equal to 1, the following relations can be derived:

$$\sum_{m \neq r} p'[m] = 1 - p'[r] \tag{3.76}$$

$$\sum_{m \neq r} p''[m] = 1 - p''[r] \tag{3.77}$$

By combining equations (3.74) and (3.76):

$$\frac{p'[r]}{\sum_{m \neq r} p'[m]} = \frac{exp(V'[r])}{\sum_{m \neq r} exp(V[m])} = \frac{p'[r]}{1 - p'[r]} \tag{3.78}$$

Similarly, by combining equations (3.75) and (3.77):

$$\frac{p''[r]}{\sum_{m \neq r} p''[m]} = \frac{exp(V''[r])}{\sum_{m \neq r} exp(V[m])} = \frac{p''[r]}{1 - p''[r]} \tag{3.79}$$

Moreover, a variation coefficient γ is introduced. In particular, it is defined as the ratio between the exponential functions of the utility associated to the rail mode, before and after the implementation of the design alternative:

$$\gamma = \frac{exp(V'[r])}{exp(V''[r])} = exp(V'[r] - V''[r]) \tag{3.80}$$

According to equations (3.78) and (3.79):

$$\gamma = \frac{\dfrac{p'[r]}{1 - p'[r]}}{\dfrac{p''[r]}{1 - p''[r]}} \tag{3.81}$$

By manipulating equation (3.81):

$$\gamma = \frac{p'[r]}{1 - p'[r]} \cdot \frac{1 - p''[r]}{p''[r]} \Rightarrow \frac{1 - p''[r]}{p''[r]} = \gamma \frac{1 - p'[r]}{p'[r]} \Rightarrow$$

$$\Rightarrow \frac{1 - p''[r]}{p''[r]} \cdot p''[r] = \gamma \frac{1 - p'[r]}{p'[r]} \cdot p''[r] \Rightarrow 1 - p''[r] = \gamma \frac{1 - p'[r]}{p'[r]} \cdot p''[r] \Rightarrow$$

$$\Rightarrow \gamma \frac{1 - p'[r]}{p'[r]} \cdot p''[r] + p''[r] = 1 \Rightarrow p''[r] \cdot \left(\gamma \frac{1 - p'[r]}{p'[r]} + 1 \right) = 1 \Rightarrow$$

$$\Rightarrow p''[r] = \frac{1}{\gamma \cdot \frac{1-p'[r]}{p'[r]} + 1} \tag{3.82}$$

By substituting (3.80) into (3.82):

$$p''[r] = \frac{1}{1 + exp\left(V'[r] - V''[r]\right) \cdot \frac{p'[r]}{1-p'[r]}} \tag{3.83}$$

Therefore, noting the utilities associated to the rail mode, before and after the intervention, as well as the choice probability related to the rail transport before the intervention, it is possible to derive the adjusted choice probability associated to the rail transport (i.e. the probability of choosing rail transport after the intervention). Clearly, consequently, also probabilities associated to the other considered modes have to be properly updated, taking into account the previous framework.

The last phase of the proposed procedure aims at defining hourly matrices being consistent with the corrected matrices, that is:

- in the time period 7.30-9.29, the hourly travel demand can be derived by dividing by 2 the peak hour origin-destination matrix $\overline{OD_{ph}^{i,r}}$;
- in the morning period (for instance, until 13.30), excluding the peak hour period already analysed, the hourly travel demand can be derived by dividing the outward matrix minus the peak hour matrix (i.e. $\overline{OD_{ad}^{i,r}} - \overline{OD_{ph}^{i,r}}$) by suitable coefficients (for instance, obtained from previous flow studies);
- in the afternoon and evening period (for instance, from 13.30 onwards), hourly demand may be derived by dividing the transposed of the outward matrix $\overline{OD_{ad}^{i,r}}$ by suitable coefficients.

It is worth noting that the developed procedure makes use of all previously described methodologies (see paragraph 2.3) for estimating and forecasting travel demand, by properly integrating them with each other in a comprehensive theoretical framework. Indeed, the use of data from national census represents a direct estimation of travel demand. Moreover, the possibility of considering three different levels of demographic variation allows to meet the requirement of stochasticity. On the other hand, recourse to data from turnstile counts, in order to update the initial O-D matrices according to the surveyed flows, constitutes an aggregate estimation technique. Finally, the specification, calibration and validation of a suitable modal choice model represent a disaggregate estimation of travel demand. Furthermore, by means of projections to future analysis periods, through both real and estimated data, forecasting techniques are implemented.

By assigning the hourly matrices, identified in the last step, to the network, it is possible estimating economic, social, financial and environmental effects associated to each alternative scenario to be analysed. In particular, the performance indexes proposed for carrying out such an evaluation are expressed by the following objective function, which considers the costs of public administration, passengers and society:

$$Z(\boldsymbol{y}, \boldsymbol{fc}, \boldsymbol{tnp}, \boldsymbol{rnp}, \boldsymbol{td}, \boldsymbol{rc}) = \beta_{UGC} \cdot UGC + \beta_{NOC} \cdot NOC + \beta_{EC} \cdot EC \tag{3.84}$$

where β_{UGC}, β_{NOC} and β_{EC} are homogeneity coefficients which express the relative weight of the objective function terms; *NOC* is the net operational cost (i.e. the part of operational costs not covered by ticket revenues); *UGC* is the user generalised cost which, clearly, has to be computed for all involved modes; *EC* is the environmental cost referred to the whole transportation system.

The first term can be computed as follows:

$$NOC = TOC - TR \tag{3.85}$$

where *TOC* is the total operational cost, depending on the reference regulation adopted by national and regional governments for funding mass-transit transport systems, *TR* represents the ticket revenues, depending on fare policies and user mobility choices.

In particular, in Italy, the funding regulation relative to the public transport sector is based on a contractual rate, known as standard cost, at which the government pays the service company according to transport supply; in addition, a constraint on service effectiveness, expressed in terms of the ratio between ticket revenues and operational costs, has to be respected. Within this framework, alternatively to equation (3.6), the *TOC* can also be specified as:

$$TOC = C_{train-km} \cdot train - km \tag{3.86}$$

with

$$train - km = \sum_{i} \sum_{\Delta t} L_i \cdot \varphi_{i,\Delta t} \cdot T_{\Delta t} \tag{3.87}$$

$$\sum_{\Delta t} T_{\Delta t} = 8{,}760 \text{ hours} = 1 \text{ year} \tag{3.88}$$

where $C_{train\text{-}km}$ is the standard cost (expressed in Euros per train-km); $train - km$ is the unit of measurement adopted to quantify the supply service; L_i is the length (expressed in kilometres) of line i; $\varphi_{i,\Delta t}$ is the service frequency (expressed in trains per hour) of line i during time interval Δt; $T_{\Delta t}$ is duration (expressed in hours) of time interval Δt.

While, *TR* can be derived by means of the following equation:

$$TR = \sum_{j} \sum_{l} \sum_{\Delta t} \frac{tc_j}{n_{l,j}} \cdot f_{l,\Delta t} \tag{3.89}$$

where tc_j is the revenue associated to ticket type j; $n_{l,j}$ is the number of trips made by user category l by using ticket j; $f_{l,\Delta t}$ is the passenger flow of category l during time interval Δt.

The second term, *UGC*, is given by the sum of user generalised costs associated to all analysed modes, that is:

$$PGC = RUG + MTUC + RC \tag{3.90}$$

with:

$$RUG = T_{ae} + T_w + T_{ob} + T_t + C_m \quad (3.91)$$

$$MTUC = T_{ae} + T_w + T_{ob} + T_t + C_m \quad (3.92)$$

$$RC = +T_w + T_{ob} + C_m \quad (3.93)$$

where *RUG* is user cost on the analysed rail system; *MTUC* is user cost on mass-transit systems, apart from the analysed rail system; *RC* is user cost on the road system; T_{ae} is the access and egress time, T_w is the waiting time, T_{ob} is the on-board time, T_t is the transfer time; C_m is the monetary cost. Obviously, each one of these temporal rates, as well as the monetary cost, have to be derived according to the specific considered mode.

Finally, *EC* can be calculated, following the approach proposed by Gallo et al. (2011b), as:

$$EC = ec_{km} \cdot \sum_{\Delta t} \sum_{a} fc_{a,\Delta t} \cdot L_a \quad (3.94)$$

where ec_{km} is the environmental cost (expressed in Euros per kilometre) associated to each vehicle in the road system (i.e. car or truck); $fc_{a,\Delta t}$ is the traffic flow associated to road link *a* during time interval *Δt*; L_a is the length (expressed in kilometres) of road link *a*.

However, it is worth pointing out that the proposed performance indexes are intended to be illustrative and not limiting. Indeed, it is necessary to properly design the objective function to be evaluated according to the specific intervention to be analysed and the related relevant impacts.

3.4 Concluding remarks

Given the complexity of the proposed methodology and the high computational times involved, it is clear that it cannot be implemented in real-time approaches and, in fact, it is conceived for a different decision-making process. The idea behind consists in the creation of a dynamic database which, for each possible intervention strategy, related or not to a specific failure event, provides the identification and the quantification of relevant impacts on each part of the analysed system. In our proposal, the considered targets are user generalised costs as well as operational costs and energy consumption; however, it is clear that additional information can be simply carried out by developing specific simulation frameworks. Therefore, by making such a database available to dispatchers, obviously for each specific network context considered, two important benefits could be achieved. Firstly, they could be fully aware of the consequences of their own decisions, thus facing the perturbed conditions in an appropriate way, never opting again for the *non-intervention strategy* which is, even now, the most frequently measure adopted. Moreover, in this way, response times could be made comparable with real-time rescheduling approaches, without, however, the computational effort they require.

The main drawback is represented by the possibility that the specific conditions to be addressed are not included in the database yet. For this reason, it results fundamental to rely on a dynamic framework, able to be upgraded with additional information which could be referred to as event preceding its creation (e.g. by means of time-series data) or subsequent thereto (e.g. by means of learning algorithms). In particular, the implementation of a properly designed algorithm, based on feature learning techniques allowing the database to both learn new notions and use them to perform specific tasks, would be very useful. Indeed, in this case, the database could be upgraded when the system is not in operation (e.g. during the night for a day-time service) with further information from the previously performed service and, thus, draw upon more and more up-to-date data. Furthermore, the dynamic nature of the database allows to obtain, under the same upgrade level, different information for different time periods (i.e. peak and off-peak hours during the day, working and non-working days during the week and different seasons during the year), by properly taking into account travel demand time variations, whose significance for an accurate system evaluation has already been widely explained. Obviously, the information collected in the database could be useful also in ordinary conditions (e.g. in the case of the implementation of energy saving measures) as well as for supporting further designed phases (e.g. the optimal allocation of recovery tracks).

By way of conclusion, in this chapter, a decision support tool has been presented, with the aim of enabling a well-rounded evolution in dispatchers' decision-making process, both in ordinary and perturbed conditions. In particular, in the second case, the aim is twofold: identifying, on one side, measures for preventing the rise of potential disturbances (i.e. *preventive action*) and, on the other, selecting the optimal intervention strategies for properly addressing the rescheduling process required after a failure (i.e. *corrective action*).

An overview of the above described database and related features are shown in figure 3.12.

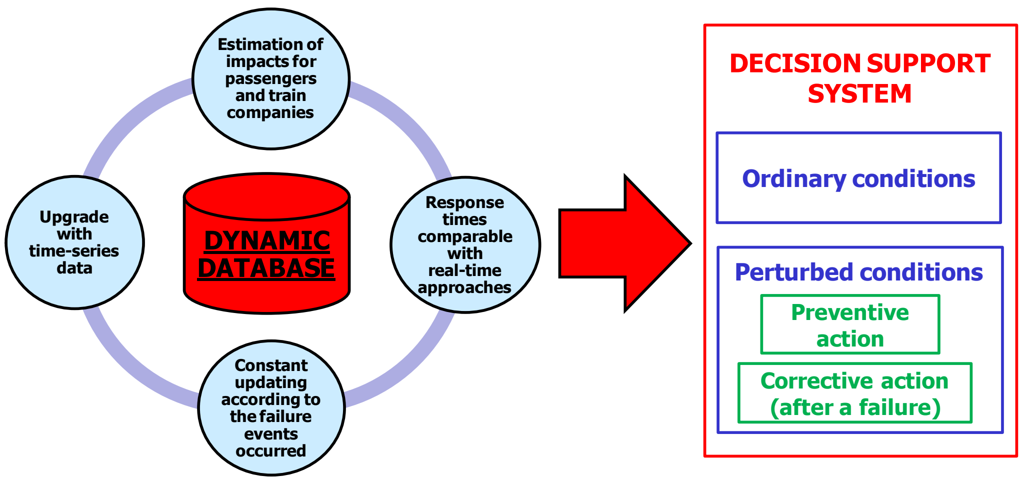


Figure 3.12 Architecture of the proposed decision support system

CHAPTER 4: APPLICATIONS TO REAL NETWORK CONTEXTS OF THE PROPOSED APPROACH

Our proposal consists in generating a dynamic database as a decision-making tool for a well-rounded management of rail systems. It has been conceived as a decision support system for handling both ordinary and perturbed conditions, as well as planning tasks. Besides rescheduling actions, very important ordinary management tasks are related to the implementation of energy saving strategies; moreover, in rail contexts, one of the most important planning phases is the timetabling process which requires an accurate evaluation of the involved operational parameters. However, each planning task, both in the case of short (e.g. fare policies) and long term measures (e.g. doubling of the track), requires an estimation of travel demand, in current and/or future conditions, as input. In particular, this chapter aims to demonstrate the suitability of the presented methodology for the above mentioned managerial issues by applying it in the case of real network contexts.

4.1 Case studies

In this paragraph, the description of the two analysed network contexts is provided, by properly putting in evidence their features and the existing differences.

The first case study is represented by *Line 1* of the Naples metro system (figure 4.1), which is operated by ANM transport company and winds through 18 stations, by connecting the high density suburbs with the city centre. It is about 18 kilometres long and mostly underground with two completely separate tunnels, one per direction, except for the stretch between Piscinola and Colli Aminei. The infrastructure layout is quite complex, because of the hilly terrain which requires the adoption of steep slopes and low radii of curvature. In a rescheduling view, it is worth noting that stations are equipped with different servicing facilities. In particular, four stations (i.e. Piscinola, Colli Aminei, Medaglie d'Oro and Garibaldi) are equipped both with points and recovery tracks; while, two (i.e. Vanvitelli and Dante) only with points. Moreover, just one depot is available for rail service and it is situated nearby Piscinola station.

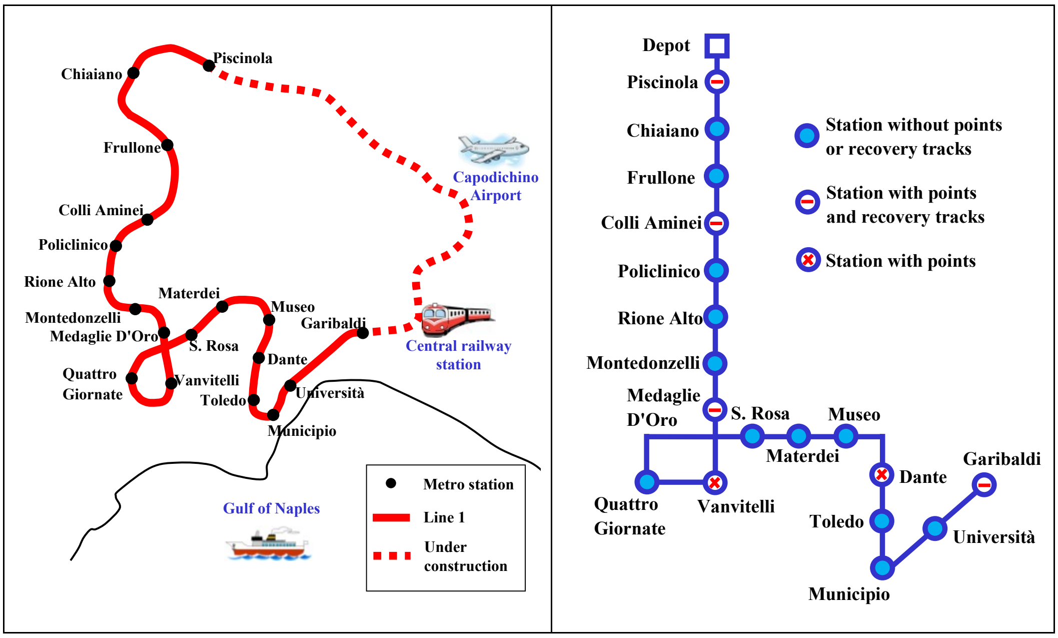


Figure 4.1 Line 1 of the Naples metro system

As to the implemented signalling system, the spacing between two consecutive convoys along the line is dictated by a station-to-station logic, which means that a convoy cannot leave a station if the following one is occupied by another train. However, stations are equipped with home signals for facing eventual emergency situations. The routing of trains within stations is ruled by electric interlocking devices (i.e. based on a relay technology), coupled with the auxiliary ATIS (Audio-frequency Transmission and Interlocking System), between Piscinola and Dante; while, electronic interlocking systems (i.e. based on a software technology) are implemented from Dante to the end of the line. Finally, regarding the on-board signalling equipment, the following systems are installed: continuous *ATP* (*Automatic Train Protection*), discontinuous *ATP* and *ATO* (*Automatic Train Operation*). In particular, the ATP system provides cab-signalling functions, supervision functions and intervention functions such as the activation of the emergency brake. On the other hand, the ATO system has the aim of allowing a partial or full automation of rail operations.

Rolling stock consists in trains composed by two-carriage modular elements (i.e. traction units) which can be coupled up to a maximum of three, by reaching a capacity of 1296 passengers. More in detail, each traction unit has a capacity of 432 passengers (120 sitting and 312 standing).

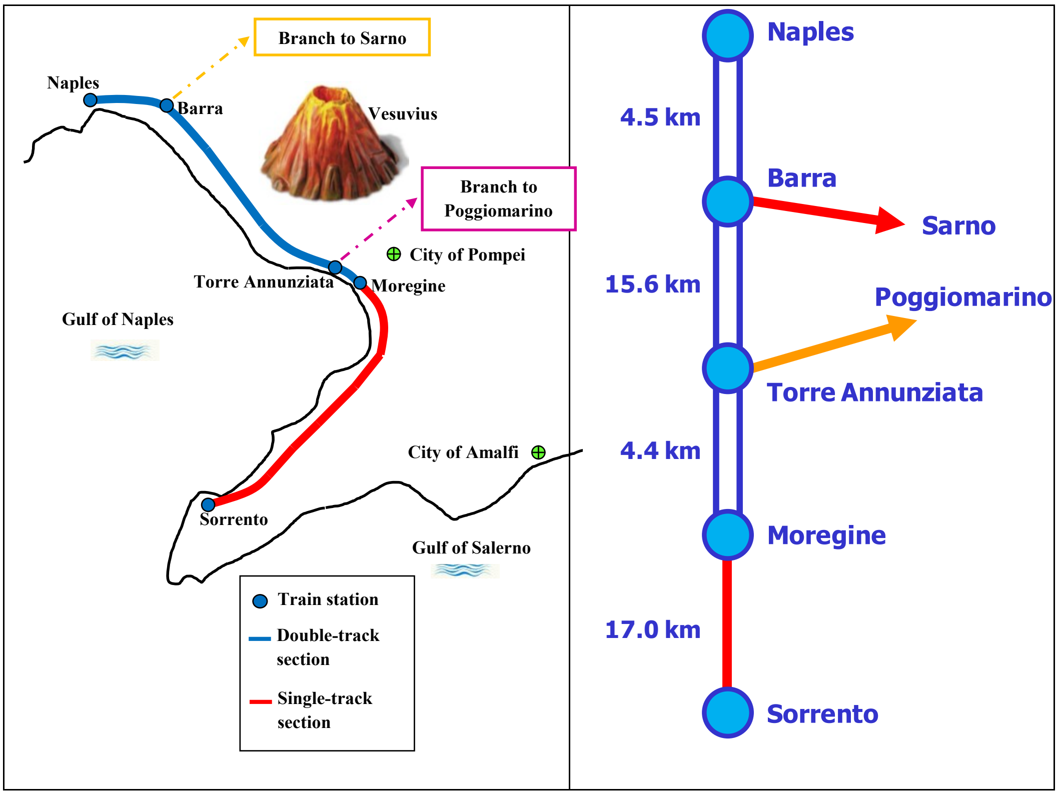


Figure 4.2 Naples-Sorrento regional line

The second case study is represented by the *Naples-Sorrento line* (figure 4.2) which is one of the six lines belonging to the Circumvesuviana regional railway, operated by EAV transport company. Circumvesuviana network is a narrow-gauge railway which serves the metropolitan area of Naples in southern Italy. It has 97 stations and is about 142 kilometers long. Specifically, the Naples-Sorrento line connects the regional capital Naples with the Sorrento peninsula, a very famous tourist area, known all over the world for its natural beauty. It is 41.5 kilometers long and can be decomposed into a first part, 24.5

km long, between Naples and Moregine, based on a double-track framework and a second part, 17.0 km long, between Moregine and Sorrento, based on a single-track framework. Moreover, in Barra and Torre Annunziata, there are the junctions respectively for Sarno and Poggiomarino. Hence, between Naples and Torre Annunziata, there is an overlapping of different lines.

The spacing between two successive convoys along the line is dictated by the Italian cab signalling system, known as BACC (Blocco Automatico a Correnti Codificate); while, the on-board signalling equipment is represented by the Italian *ATP* system, known as *SCMT* (Sistema di Controllo Marcia Treno). Finally, the interlocking systems are based on a relay technology and field elements are operated and controlled electrically by means of dedicated buttons. Trains operating on the line are made up of three indivisible carriages, each of which offers a maximum capacity of 450 passengers (48 sitting and 402 standing), for a total capacity of 1350 passengers.

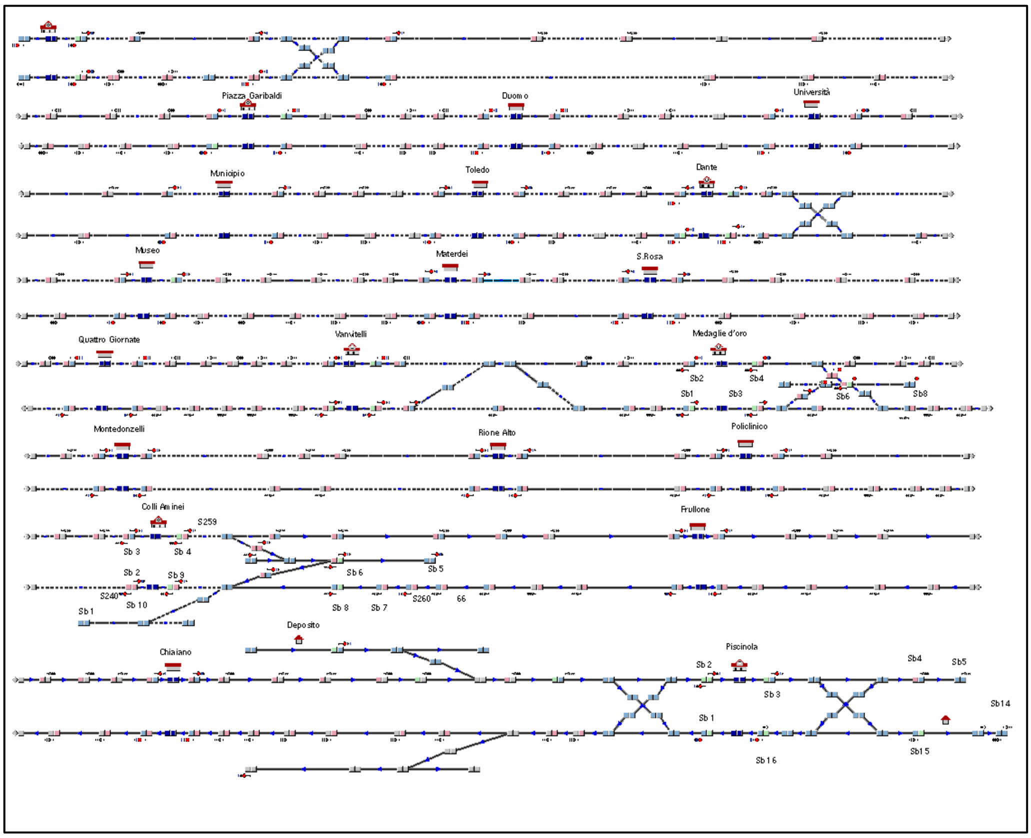


Figure 4.3 OpenTrack network representation: Line 1

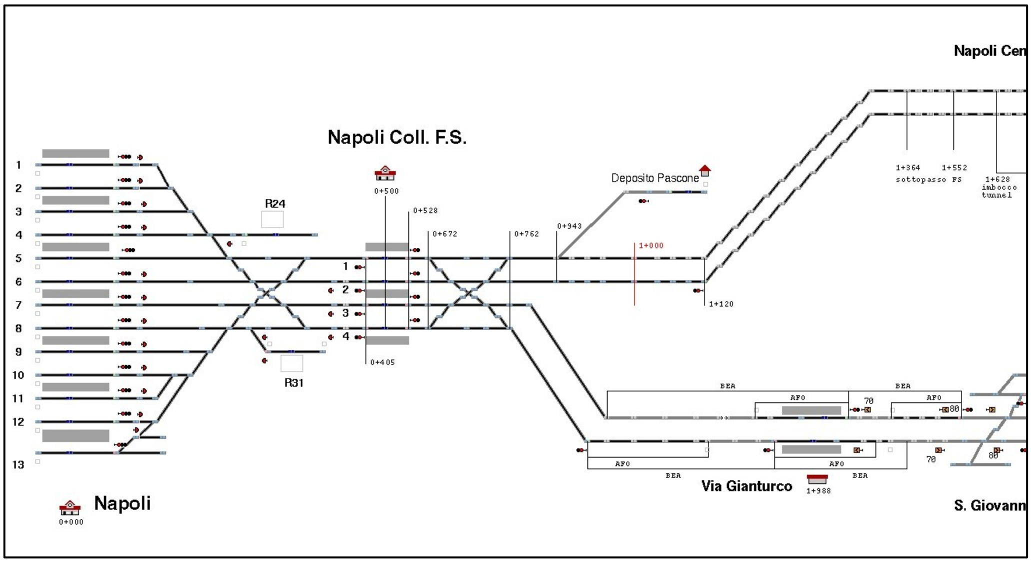


Figure 4.4 OpenTrack network representation: extract of the Naples-Sorrento line

In both cases, infrastructure, signalling system and rolling stock features need to be properly modelled within the *SeSM* which is performed by the micro simulation software OpenTrack. In particular, figures 4.3 and 4.4 show the analysed networks as being depicted in OpenTrack.

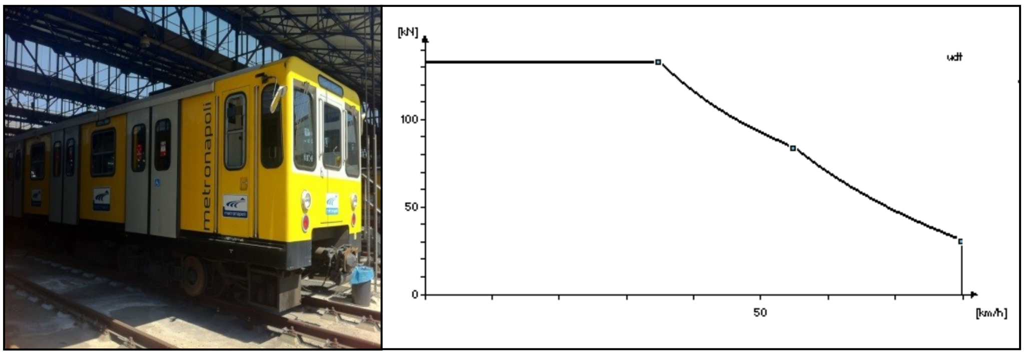


Figure 4.5 Tractive effort/velocity of Line 1 train.

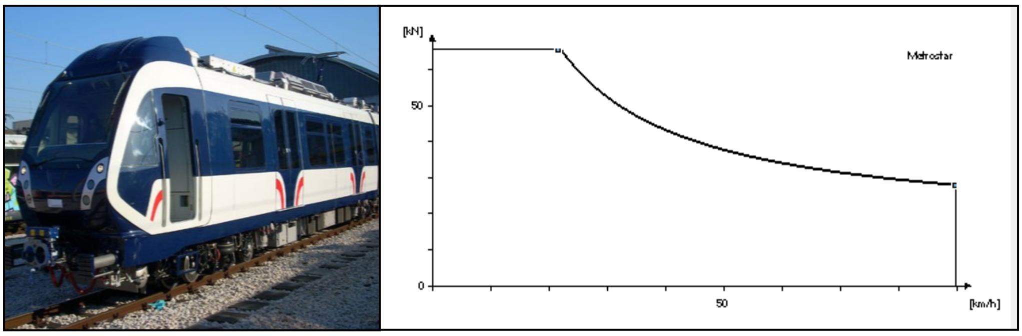


Figure 4.6 Tractive effort/velocity of the Naples-Sorrento line train.

Clearly, given the passenger-oriented perspective adopted in this work, as well as the importance of considering interactions between travel demand and rail service widely discussed in the previous chapter, the capacity offered by the convoys is crucial for our analysis. However, in order to calibrate the *SeSM* faithfully to the reality, also mechanical and traction features of rolling stock have to be accurately modelled (e.g. adherence load, maximum speed, maximum tractive effort, rotation mass factor). By way of example, figures 4.5 and 4.6 show the tractive/effort diagram of trains operating, respectively, on Line 1 and on the Naples-Sorrento line. It is worth noting that, although trains operating on the Naples-Sorrento line can reach theoretically a maximum speed of 90 km/h, the short distance between two consecutive stations rarely allows them to effectively reach such a speed.

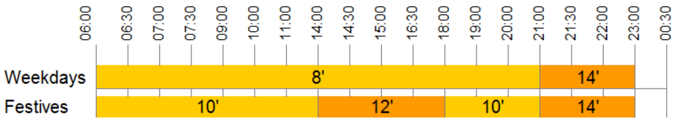


Figure 4.7 Timetable of Line 1.

	D	DD	D	D	D	DD	D	D	D	D	D	D	D	D	D	D	DD	D	D
		*	FER																
	11	13	1013	15	17	19	21	23	25	27	29	31	33	35	37	39	41	43	45
NAPOLI P.Nolana	6.09	6.40	6.44	7.09	7.39	8.11	8.39	9.09	9.39	10.09	10.39	11.09	11.39	12.09	12.39	13.09	13.41	14.09	14.39
Napoli Garibaldi	6.11	6.42	6.46	7.11	7.41	8.13	8.41	9.11	9.41	10.11	10.41	11.11	11.41	12.11	12.41	13.11	13.43	14.11	14.41
Via Gianturco	6.13		6.48	7.13	7.43		8.43	9.13	9.43	10.13	10.43	11.13	11.43	12.13	12.43	13.13		14.13	14.43
S. Giovanni	6.15		6.50	7.15	7.45		8.45	9.15	9.45	10.15	10.45	11.15	11.45	12.15	12.45	13.15		14.15	14.45
BARRA	6.17		6.52	7.17	7.47		8.47	9.17	9.47	10.17	10.47	11.17	11.47	12.17	12.47	13.17		14.17	14.47
S. Maria del Pozzo	6.19		6.54	7.19	7.49		8.49	9.19	9.49	10.19	10.49	11.19	11.49	12.19	12.49	13.19		14.19	14.49
S. Giorgio	6.21	6.49	6.56	7.21	7.51	8.20	8.51	9.21	9.51	10.21	10.51	11.21	11.51	12.21	12.51	13.21	13.50	14.21	14.51
S. Giorgio-Cav.Bronzo	6.23		6.58	7.23	7.53		8.53	9.23	9.53	10.23	10.53	11.23	11.53	12.23	12.53	13.23		14.23	14.53
Portici Bellavista	6.24		6.59	7.24	7.54		8.54	9.24	9.54	10.24	10.54	11.24	11.54	12.24	12.54	13.24		14.24	14.54
Portici Via Libertà	6.26		7.01	7.26	7.56		8.56	9.26	9.56	10.26	10.56	11.26	11.56	12.26	12.56	13.26		14.26	14.56
Ercolano Scavi	6.28	6.52	7.02	7.28	7.58	8.23	8.58	9.28	9.58	10.28	10.58	11.28	11.58	12.28	12.58	13.28	13.53	14.28	14.58
Ercolano-Miglio D'Oro	6.29		7.04	7.29	7.59		8.59	9.29	9.59	10.29	10.59	11.29	11.59	12.29	12.59	13.29		14.29	14.59
Torre del Greco	6.31	6.55	7.06	7.31	8.01	8.25	9.01	9.31	10.01	10.31	11.01	11.31	12.01	12.31	13.01	13.31	13.55	14.31	15.01
Via S. Antonio	6.33		7.08	7.33	8.03		9.03	9.33	10.03	10.33	11.03	11.33	12.03	12.33	13.03	13.33		14.33	15.03
Via del Monte	6.34		7.09	7.34	8.04		9.04	9.34	10.04	10.34	11.04	11.34	12.04	12.34	13.04	13.34		14.34	15.04
Via dei Monaci																			
Villa delle Ginestre	6.36		7.12	7.36	8.06		9.06	9.36	10.06	10.36	11.06	11.36	12.06	12.36	13.06	13.36		14.36	15.06
Leopardi	6.38		7.13	7.38	8.08		9.08	9.38	10.08	10.38	11.08	11.38	12.08	12.38	13.08	13.38		14.38	15.08
Via Viuli																13.39		14.39	15.09
Trecase	6.40		7.15	7.40	8.10		9.10	9.40	10.10	10.40	11.10	11.40	12.10	12.40	13.10	13.40		14.40	15.10
TORRE A. Olponti a.	6.42	7.00	7.17	7.42	8.12	8.31	9.12	9.42	10.12	10.42	11.12	11.42	12.12	12.42	13.12	13.42	14.01	14.42	15.12
TORRE A. Oplonti p.	6.43	7.01	7.18	7.43	8.13	8.32	9.13	9.43	10.13	10.43	11.13	11.43	12.13	12.43	13.13	13.43	14.02	14.43	15.13
Pompei S. Villa Misteri	6.47	7.05	7.22	7.47	8.17	8.36	9.17	9.47	10.17	10.47	11.17	11.47	12.17	12.47	13.17	13.47	14.06	14.47	15.17
Moregine	6.48	7.06	7.23	7.48	8.18		9.18	9.48	10.18	10.48	11.18	11.48	12.18	12.48	13.18	13.48		14.48	15.18
Ponte Persica	6.49		7.24	7.49	8.19		9.19	9.49	10.19	10.49	11.19	11.49	12.19	12.49	13.19	13.49		14.49	15.19
Pioppaino	6.51	7.09	7.26	7.51	8.21	8.40	9.21	9.51	10.21	10.51	11.21	11.51	12.21	12.51	13.21	13.51	14.10	14.51	15.21
Via Nocera	6.54	7.11	7.29	7.54	8.24	8.42	9.24	9.54	10.24	10.54	11.24	11.54	12.24	12.54	13.24	13.54	14.12	14.54	15.24
Castellammare	6.57	7.15	7.31	7.57	8.27	8.46	9.27	9.57	10.27	10.57	11.27	11.57	12.27	12.57	13.27	13.57	14.16	14.57	15.27
Pozzano			7.35	8.01	8.31	8.49	9.31	10.01	10.31	11.01	11.31	12.01	12.31	13.01	13.31	14.01	14.19	15.01	15.31
Scrajo			7.39	8.05	8.35	8.53	9.35	10.05	10.35	11.05	11.35	12.05	12.35	13.05	13.35	14.05	14.23	15.05	15.35
Vico Equense	7.06	7.23	7.40	8.06	8.36	8.54	9.36	10.06	10.36	11.06	11.36	12.06	12.36	13.06	13.36	14.06	14.24	15.06	15.36
Seiano	7.08	7.24	7.42	8.08	8.38	8.55	9.38	10.08	10.38	11.08	11.38	12.08	12.38	13.08	13.38	14.08	14.25	15.08	15.38
Meta	7.11	7.27	7.46	8.11	8.41	8.58	9.41	10.11	10.41	11.11	11.41	12.11	12.41	13.11	13.41	14.11	14.28	15.11	15.41
Piano	7.13	7.28	7.48	8.13	8.43	8.59	9.43	10.13	10.43	11.13	11.43	12.13	12.43	13.13	13.43	14.13	14.29	15.13	15.43
S. Agnello	7.15	7.30	7.50	8.15	8.45	9.01	9.45	10.15	10.45	11.15	11.45	12.15	12.45	13.15	13.45	14.15	14.31	15.15	15.45
SORRENTO	7.17	7.32	7.52	8.17	8.47	9.03	9.47	10.17	10.47	11.17	11.47	12.17	12.47	13.17	13.47	14.17	14.33	15.17	15.47

Figure 4.8 Extract of the timetable of the Naples-Sorrento line.

It is worth noting that, within the *SeSM*, another important element to be modelled is the planned timetable, which is easily accessible to all users, for instance, on web-sites of train operating companies or on departure/arrival boards in the stations (figure 4.7 and 4.8). This is a chance to underline a relevant difference between metro systems and regional services which has to be duly taken into account in the simulation task. Specifically, the peculiarity of metro contexts lies in their nature of frequency-based systems, which means that the target consists in respecting a planned headway, rather than a specific departure/arrival time at each station which, generally, is even unknown to users.

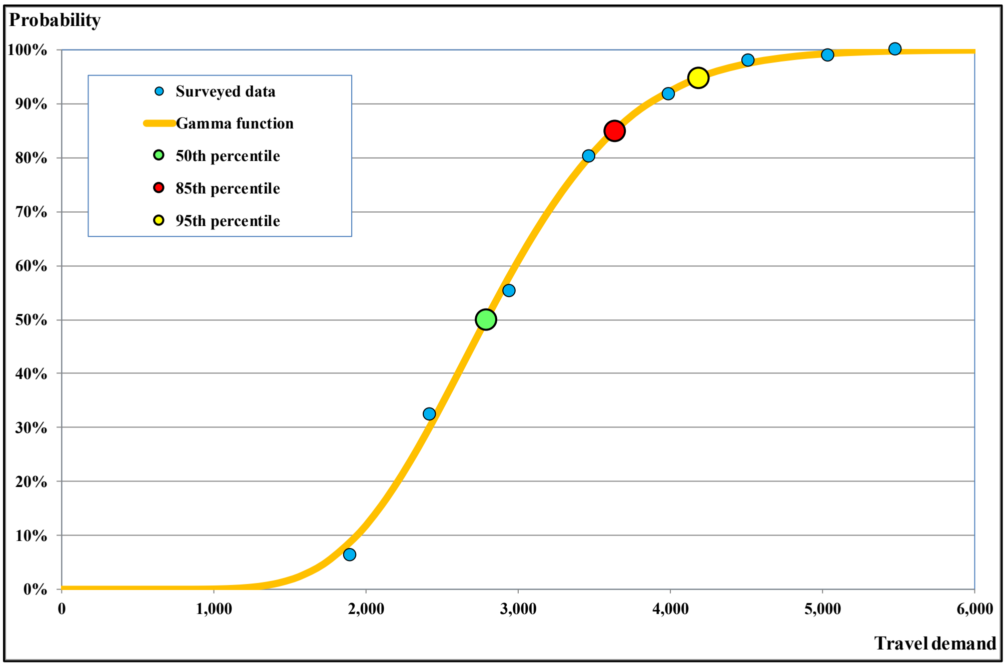


Figure 4.9 Different travel demand levels in the case of Line 1

The last factor to be specified, for each one of the analysed context, is the definition of travel demand as O-D matrices expressed in terms of rail stations to be assigned to the network. Also in this case, the different nature of the analysed systems has to be properly considered, since in metro systems the demand to be estimated refers to an urban context; while, suburban journeys have to be evaluated for regional services. In particular, travel demand implemented for Line 1 has been carried out as shown in Ercolani et al. (2014):

starting from surveyed data, the function which better fits them is selected and, thus, different demand levels (i.e. percentile values) can be identified (figure 4.9). However, it is worth pointing out that this is a preliminary result, simply used as input for comparing different rescheduling strategies. While, in the case of applications focused on statistical procedures for handling travel demand flows according to peculiarities of rail systems, this initial estimation will be improved and properly enhanced for showing the effectiveness of the proposed methodology. On the other hand, the travel demand used for the analysis of unconventional rescue strategies in the case of the Naples-Sorrento line is that estimated according to the procedure which will be described in paragraph 4.5.2.

Once all features related to the involved components (i.e. infrastructure, signalling system, rolling stock, timetable and travel demand) have been properly modelled, it is possible to reproduce the current scenario which represents the starting point for each further evaluation.

4.2 Rescheduling applications

In terms of rescheduling actions, the proposed methodology aims to identify the best intervention strategies for facing disruption conditions in a passenger-oriented perspective. However, as already explained, by enriching the objective function, additional targets can be evaluated.

The case-study analysed in this section is Line 1 of the Naples metro system and the considered failure scenario consists in supposing that a breakdown occurs to the continuous ATP system of a convoy at Chiaiano station, during the morning peak-hour. In this case, the faulty train can rely exclusively on the discontinuous ATP system and, therefore, it is forced to travel at a maximum speed of 45 km/h (i.e. the speed value dictated by the position of the balises for the discontinuous train protection). Obviously, this reduction in performance represents a bottleneck for the whole service. Given the layout of the line, the tested intervention strategies are based on the following options:

- continuing the service as far as a station equipped with a recovery track and driving the train onto the maintenance track, just after unloading passengers on the platform;
- continuing the service as far as a station equipped with points and driving the train to the depot (located next to Piscinola station) by changing direction, just after unloading passengers on the platform;
- recovering the damaged train on a maintenance track or at the depot, with or without the use of a spare train for completing the service for the rest of the day.

According to the optimisation framework described in paragraph 3.1, intervention strategies can be formalised by means of a vector $\boldsymbol{y}$ whose components are:

y_1 representing the strategy type implemented, that is:

- 1 = recovery on a maintenance track
- 2 = changing direction in a station with points

y_2 representing the time when the strategy is implemented, that is:

- 1 = during the outgoing trip
- 2 = during the return trip

y_3 representing the station where the strategy is implemented, that is:

- 1 = Colli Aminei
- 2 = Medaglie d'Oro
- 3 = Vanvitelli
- 4 = Dante
- 5 = Terminus (i.e. Garibaldi during the outgoing trip or Piscinola during the return trip)

y_4 representing the use of a spare train, that is:

- 1 = no spare train
- 2 = spare train

However, the following combinations of values are not feasible:

- changing direction during the return trip, that is: $y_1 = 2$ and $y_2 = 2$;
- changing direction at the terminus, that is: $y_1 = 2$ and $y_3 = 5$;
- recovering the train at a station without a maintenance track, that is: $y_1 = 1$ and $y_3 = 3$, and $y_1 = 1$ and $y_3 = 4$.

$$\boldsymbol{y} = (y_1, y_2, y_3, y_4)$$

Variables	Values				
	1	2	3	4	5
y_1	Recovery	Changing direction			
y_2	Outward	Return			
y_3	Colli Aminei	Medaglie D'Oro	Vanvitelli	Dante	Terminus
y_4	Na spare train	Spare train			

Terminus definition		
y_2	y_3	y_{3^*}
Outward	Terminus	Garibaldi
Return	Terminus	Piscinola

No feasibility combinations		
y_1	y_2	y_3
Changing direction	Return	
Changing direction		Terminus
Recovery		Vanvitelli
Recovery		Dante

Comb.	Variable values				Feas.	Strategy number	Variable meaning				
	y_1	y_2	y_3	y_4			y_1	y_2	y_3	y_3*	y_4
1	1	1	1	1	1	1	Recovery	Outward	ColliAminei		No spare train
2	2	1	1	1	1	2	Changing direction	Outward	ColliAminei		No spare train
3	1	2	1	1	1	3	Recovery	Return	ColliAminei		No spare train
4	2	2	1	1	0		Changing direction	Return	ColliAminei		No spare train
5	1	1	2	1	1	4	Recovery	Outward	MedaglieD'Oro		No spare train
6	2	1	2	1	1	5	Changing direction	Outward	MedaglieD'Oro		No spare train
7	1	2	2	1	1	6	Recovery	Return	MedaglieD'Oro		No spare train
8	2	2	2	1	0		Changing direction	Return	MedaglieD'Oro		No spare train
9	1	1	3	1	0		Recovery	Outward	Vanvitelli		No spare train
10	2	1	3	1	1	7	Changing direction	Outward	Vanvitelli		No spare train
11	1	2	3	1	0		Recovery	Return	Vanvitelli		No spare train
12	2	2	3	1	0		Changing direction	Return	Vanvitelli		No spare train
13	1	1	4	1	0		Recovery	Outward	Dante		No spare train
14	2	1	4	1	1	8	Changing direction	Outward	Dante		No spare train
15	1	2	4	1	0		Recovery	Return	Dante		No spare train
16	2	2	4	1	0		Changing direction	Return	Dante		No spare train
17	1	1	5	1	1	9	Recovery	Outward	Capolinea	Garibaldi	No spare train
18	2	1	5	1	0		Changing direction	Outward	Capolinea	Garibaldi	No spare train
19	1	2	5	1	1	10	Recovery	Return	Capolinea	Piscinola	No spare train
20	2	2	5	1	0		Changing direction	Return	Capolinea	Piscinola	No spare train
21	1	1	1	2	1	11	Recovery	Outward	ColliAminei		Spare train
22	2	1	1	2	1	12	Changing direction	Outward	ColliAminei		Spare train
23	1	2	1	2	1	13	Recovery	Return	ColliAminei		Spare train
24	2	2	1	2	0		Changing direction	Return	ColliAminei		Spare train
25	1	1	2	2	1	14	Recovery	Outward	MedaglieD'Oro		Spare train
26	2	1	2	2	1	15	Changing direction	Outward	MedaglieD'Oro		Spare train
27	1	2	2	2	1	16	Recovery	Return	MedaglieD'Oro		Spare train
28	2	2	2	2	0		Changing direction	Return	MedaglieD'Oro		Spare train
29	1	1	3	2	0		Recovery	Outward	Vanvitelli		Spare train
30	2	1	3	2	1	17	Changing direction	Andata	Vanvitelli		Spare train
31	1	2	3	2	0		Recovery	Return	Vanvitelli		Spare train
32	2	2	3	2	0		Changing direction	Return	Vanvitelli		Spare train
33	1	1	4	2	0		Recovery	Outward	Dante		Spare train
34	2	1	4	2	1	18	Changing direction	Outward	Dante		Spare train
35	1	2	4	2	0		Recovery	Return	Dante		Spare train
36	2	2	4	2	0		Changing direction	Return	Dante		Spare train
37	1	1	5	2	1	19	Recovery	Outward	Capolinea	Garibaldi	Spare train
38	2	1	5	2	0		Changing direction	Outward	Capolinea	Garibaldi	Spare train
39	1	2	5	2	1	20	Recovery	Return	Capolinea	Piscinola	Spare train
40	2	2	5	2	0		Changing direction	Return	Capolinea	Piscinola	Spare train

Figure 4.10 Feasible intervention strategies

Therefore, by adopting the above mentioned formulation, 20 feasible solutions can be obtained, out of a total of 40 combinations, as shown in figure 4.10 by the green lines.

In particular, a detailed description of the 20 feasible intervention strategies, identified in addition to the *do nothing* solution (i.e. continuing the service with the faulty train for the entire day, indicated as intervention strategy 0), is shown in table 4.1.

No.	**Strategy description**
0	The faulty train continues to perform its service all day
1	The train stops at CA-Colli Aminei during its outward trip and is then driven onto the recovery track. No spare trains are considered
2	The train stops its run at CA-Colli Aminei and, after changing direction, is driven empty to the depot. No spare trains are considered
3	The train completes the outward trip and starts the return trip up to CA-Colli Aminei where it is driven onto the maintenance track. No spare trains are considered
4	The train stops at MO-Medaglie d'Oro during its outward trip and is then driven onto the recovery track. No spare trains are considered
5	The train stops its run at MO-Medaglie d'Oro and, after changing direction, is driven empty to the depot. No spare trains are considered
6	The train completes the outward trip and starts the return trip up to MO-Medaglie d'Oro where it is driven onto the maintenance track. No spare trains are considered
7	The train stops its run at VA-Vanvitelli and, after changing direction, is driven empty to the depot. No spare trains are considered
8	The train stops its run at DA-Dante and, after changing direction, is driven empty to the depot. No spare trains are considered
9	The train stops at GA-Garibaldi at the end of its outward trip and is then driven onto the recovery track. No spare trains are considered
10	The train completes the outward trip and starts the return trip up to PI-Piscinola where it is driven to the depot. No spare trains are considered
11	The train stops at CA-Colli Aminei during its outward trip and is then driven onto the recovery track. A spare train starts from PI-Piscinola to replace the faulty rolling stock for the rest of the daily operation
12	The train stops its run at CA-Colli Aminei and, after changing direction, is driven empty to the depot. A spare train starts from PI-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 CA-Colli Aminei where it is driven onto the maintenance track. A spare train starts from PI-Piscinola to replace the faulty rolling stock for the rest of the daily operation
14	The train stops at MO-Medaglie d'Oro during its outward trip and is then driven onto the recovery track. A spare train starts from PI-Piscinola to replace the faulty rolling stock for the rest of the daily operation

No.	Strategy description
15	The train stops its run at MO-Medaglie d'Oro and, after changing direction, is driven empty to the depot. A spare train starts from PI-Piscinola to replace the faulty rolling stock for the rest of the daily operation
16	The train completes the outward trip and starts the return trip up to MO-Medaglie d'Oro where it is driven onto the maintenance track. A spare train starts from PI-Piscinola to replace the faulty rolling stock for the rest of the daily operation
17	The train stops its run at VA-Vanvitelli and, after changing direction, is driven empty to the depot. A spare train starts from PI-Piscinola to replace the faulty rolling stock for the rest of the daily operation
18	The train stops its run at DA-Dante and, after changing direction, is driven empty to the depot. A spare train starts from PI-Piscinola to replace the faulty rolling stock for the rest of the daily operation
19	The train stops at GA-Garibaldi at the end of its outward trip and is then driven onto the recovery track. A spare train starts from PI-Piscinola to replace the faulty rolling stock for the rest of the daily operation
20	The train completes the outward trip and starts the return trip up to PI-Piscinola where it is driven to the depot. A spare train starts from PI-Piscinola to replace the faulty rolling stock for the rest of the daily operation

Table 4.1 Description of intervention strategies

Objective function (3.3), hereinafter referred to as objective function no. 1, has been calculated for each strategy in the case of two different travel demand levels (i.e. 50th and 85th percentiles of function depicted in figure 4.9), by adopting parameter values shown in table 4.2. Results are indicated in table 4.3.

Parameter	Value
β_{VOT}	5 €/h
β_w	2.5
β_{ob}	1

Table 4.2 Parameter values

Intervention strategies	Objective function no. 1 [€]	
	Average travel demand	High travel demand
0	657,707	898,952
1	715,562	948,900
2	715,416	948,702
3	709,942	948,323
4	714,587	950,304
5	714,440	950,106
6	710,432	948,997
7	714,363	950,725
8	712,474	951,169
9	710,533	949,269
10	709,188	947,260
11	650,936	887,697
12	650,790	887,499
13	645,358	886,721
14	646,102	883,667
15	649,832	888,930
16	645,652	887,131
17	648,334	887,565
18	648,191	889,972
19	646,105	887,875
20	644,415	885,404

Table 4.3 Values of objective function no. 1 (i.e. user generalised cost) for different travel demand levels

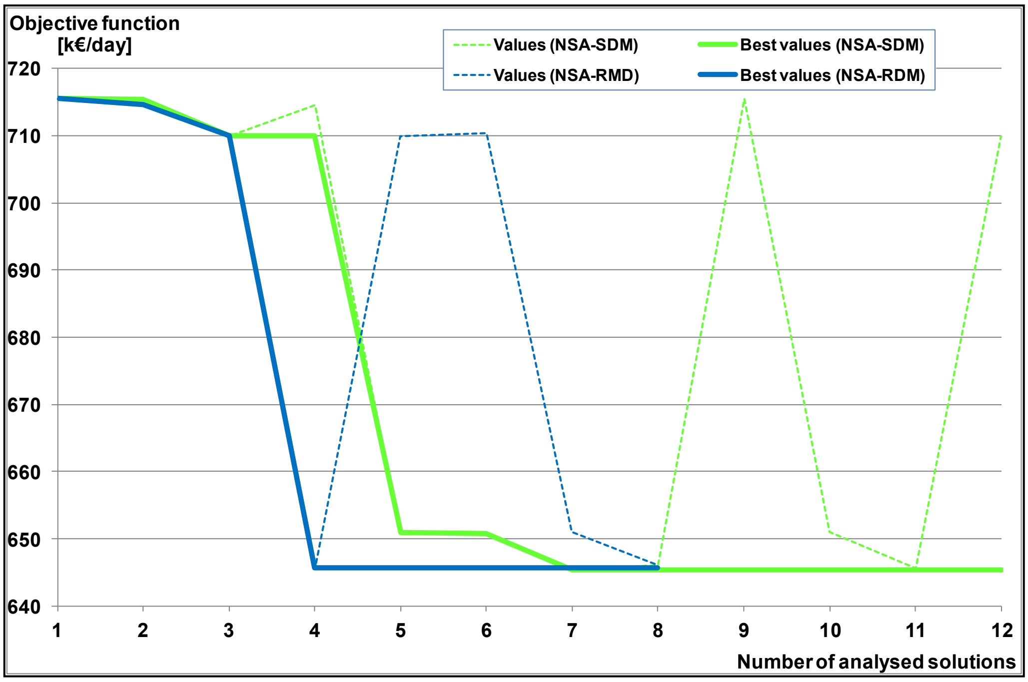
Given the simple layout of the analysed network (i.e. an isolated metro line with few points and recovery tracks), the number of feasible solutions is liable to allow the application of an exhaustive approach for solving optimisation problem (3.1). In this way, it is possible to have a frame of reference for evaluating the convenience in applying metaheuristic techniques for such problems. Therefore, in the following, a comparison between the exhaustive approach and the *NSA* method is provided. In particular, as explained in paragraph 2.7, the *Neighbourhood Search Algorithm* is a heuristic local search method for solving discrete optimisation problems, which can be implemented according to two different approaches. The *Steepest Descent Method (SDM)* consists in examining all elements of the neighbourhood and identifying the best solution (i.e. the solution with the best objective function value); while, the *Random Descent Method (RDM)* consists in randomly extracting a solution from the neighbourhood and comparing it with the current one. In particular, if the new solution is better than the current one, it then becomes the current solution; otherwise, another neighbourhood solution is randomly extracted until the neighbourhood runs out, since all solutions inside have been explored.

According to the exhaustive approach, the best solution for the average travel demand is strategy 20; while, in the case of particularly crowded days (i.e. high travel demand), strategy 14 is the one which guarantees the lower value of user generalised cost. However, in order to perform a comparison with results obtained by means of the *NSA* method, also the second and the third best strategies, for the two analysed demand levels, have been identified and provided in table 4.4.

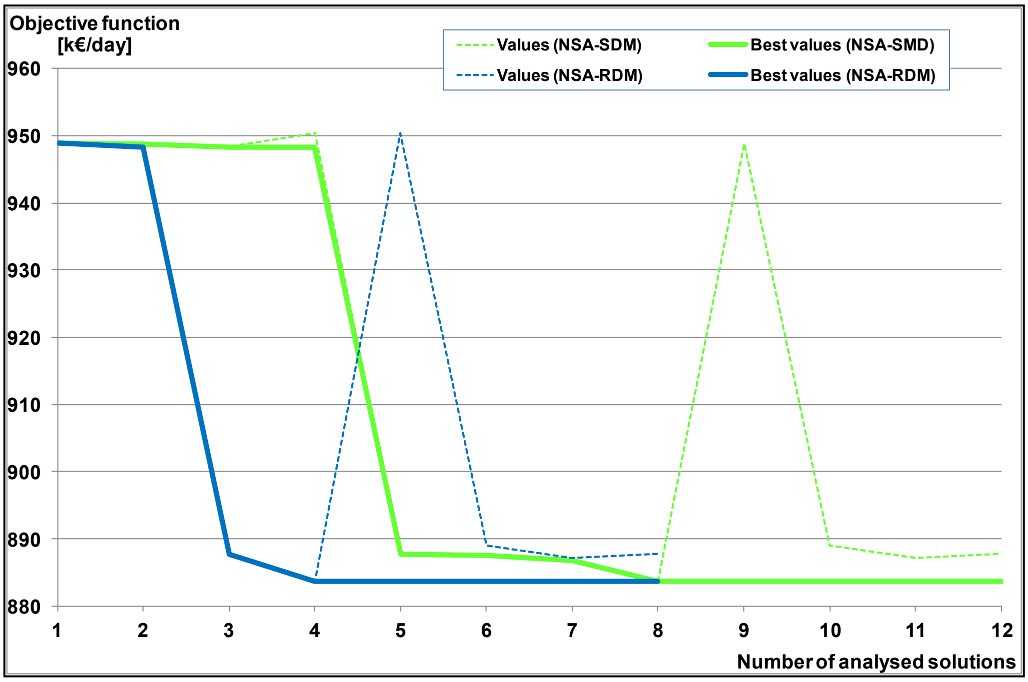
	Average travel demand		**High travel demand**	
	Intervention Strategies	**Objective function no. 1 [€]**	**Intervention Strategies**	**Objective function no. 1 [€]**
First best	20	644,415	14	883,667
Second best	13	645,358	20	885,404
Third best	16	645,652	13	886,721

Table 4.4 Exhaustive approach results

On the other hand, results obtained by implementing the two proposed variants of the NSA are compared in figure 4.11, which shows that, for both demand levels, the random approach allows to reach the same result with a lower number of iterations. Additionally, figure 4.12 contrasts the exhaustive approach with the random approach of the NSA and results show what follows. In the case of the 50th percentile, *NSA-RDM* identifies as the optimal solution strategy 13, which is a local optimum corresponding to the second best strategy according to the exhaustive approach. On the other hand, in the case of the 85th percentile, *NSA-RDM* is able to reach the global optimum (i.e. strategy 14) with a reduction of 60% in computational times with respect to the exhaustive approach.

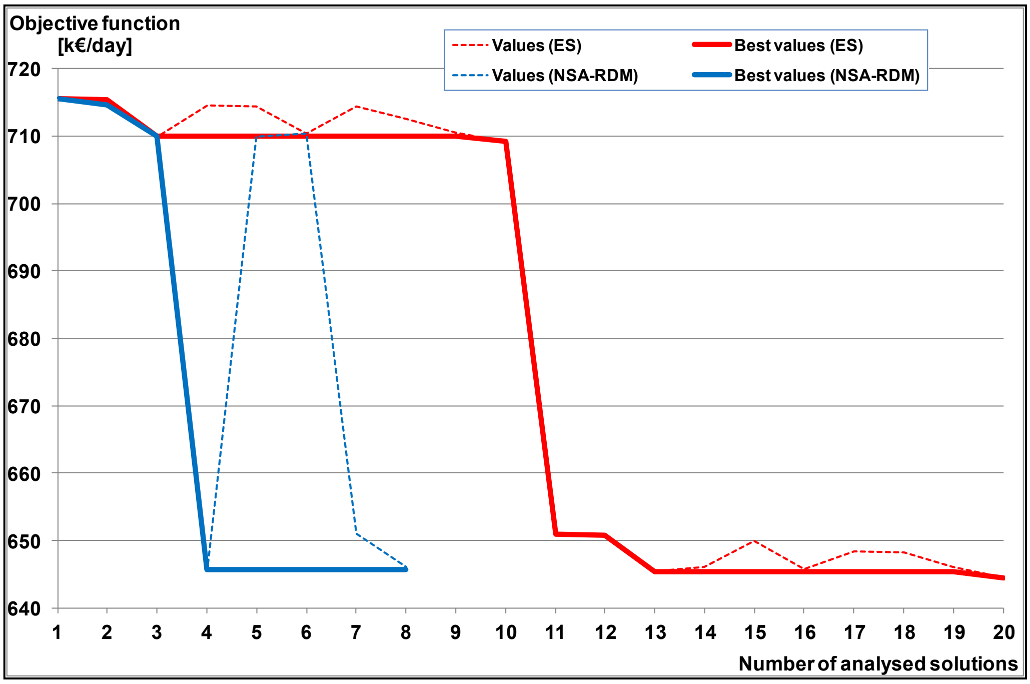


(a) 50th percentile



(b) 85th percentile

Figure 4.11 Comparison between NSA-SDM and NSA-RDM, for the two analysed travel demand levels



(a) 50th percentile

(b) 85th percentile

Figure 4.12 Comparison between the Exhaustive Approach (EA) and NSA-RDM, for the two analysed travel demand levels

Specifically, implementation of the exhaustive approach required 1.92 h; while, the use of the *NSA-RDM* provided the optimal solutions in 0.77 h. Clearly, such a reduction, which may appear negligible in the case of the simple network analysed, becomes instead very significant in the case of more complex contexts, as amply confirmed by the literature (see, for instance, D'Acierno et al., 2014).

Moreover, objective function (3.4), hereinafter referred to as objective function no. 2, corresponding to the same travel demand levels (i.e. 50th and 85th percentiles of function depicted in figure 4.9), has been calculated. In particular, the influence of different values of the components of the weight vector, i.e. β_{UGC}, β_{PEN}, β_{TOC}, was evaluated (see table 4.5). Furthermore, the parameter values adopted for calculating the objective function are indicated in table 4.6, while β_{ob} values are shown in table 4.7.

Parameter	Weight vectors	
	1	2
β_{UGC}	1	1
β_{PEN}	0.9	2.5
β_{TOC}	1	1

Table 4.5 Weights vectors

Parameter	Value
β_{VOT}	5 €/h
β_{w}	2.5
β_{ob}	see table 4.7
tls	15 minutes
c_r	18.17 €/traction unit-km

Table 4.6 Parameter values

Pax / m^2	Sitting	Standing
0	1.00	1.77
1	1.11	1.81
2	1.23	1.85
3	1.34	1.89
4	1.46	1.92
5	1.57	1.96
6	1.69	2.00

Table 4.7 Parameter β_{ob} values

The extra cost perceived by passengers (i.e. term *PEN*) was calculated by assuming that passengers decide to leave the rail system if they are forced to wait more than 20 minutes or skip two runs.

The values of objective function no. 2, for each intervention strategy, travel demand level and weight vector, are summarised in table 4.8. Specifically, for each analysed case, the optimal intervention strategy $\hat{\boldsymbol{y}}$ (i.e. red value), together with the second and the third best solutions (i.e. respectively orange and yellow values) composing its neighbourhood $N(\hat{\boldsymbol{y}})$, are identified.

Intervention strategies	Objective function no. 2 [€]			
	Average travel demand		High travel demand	
	Weight vector 1	Weight vector 2	Weight vector 1	Weight vector 2
0	810,158	823,123	1,091,408	1,175,490
1	867,133	901,799	1,148,215	1,267,760
2	867,398	902,064	1,148,428	1,267,973
3	860,941	893,028	1,146,053	1,261,216
4	866,097	900,655	1,149,438	1,268,662
5	866,361	900,919	1,149,651	1,268,875
6	861,319	893,406	1,146,615	1,261,778
7	866,253	900,629	1,150,364	1,269,627
8	864,059	897,293	1,149,833	1,266,764
9	861,096	893,183	1,146,563	1,261,726
10	860,350	892,437	1,145,153	1,260,316
11	805,454	820,973	1,083,298	1,171,854
12	804,678	820,197	1,082,470	1,171,026
13	**797,632**	**810,571**	**1,079,062**	**1,163,230**
14	799,497	815,793	**1,078,211**	1,167,662
15	803,029	818,439	1,083,090	1,171,322
16	**797,814**	**810,753**	1,079,360	**1,163,528**
17	802,080	818,341	1,082,610	1,172,291
18	801,052	815,139	1,083,248	1,169,183
19	797,944	810,883	1,079,781	1,163,949
20	**796,852**	**809,791**	**1,077,908**	**1,162,076**

Table 4.8 Values of objective function no. 2 for different travel demand levels and weight vectors

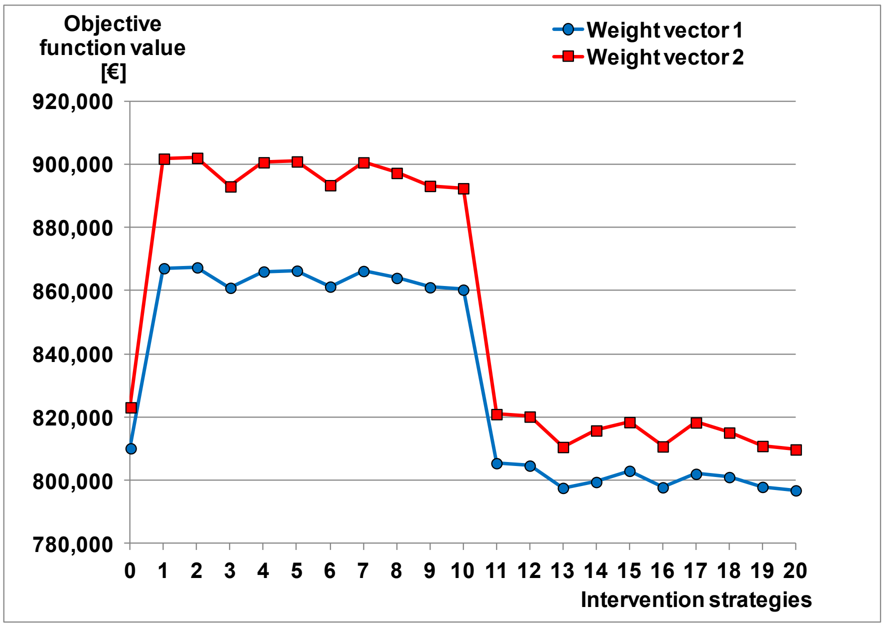


Figure 4.13 Comparison of different weight vectors for an average travel demand level

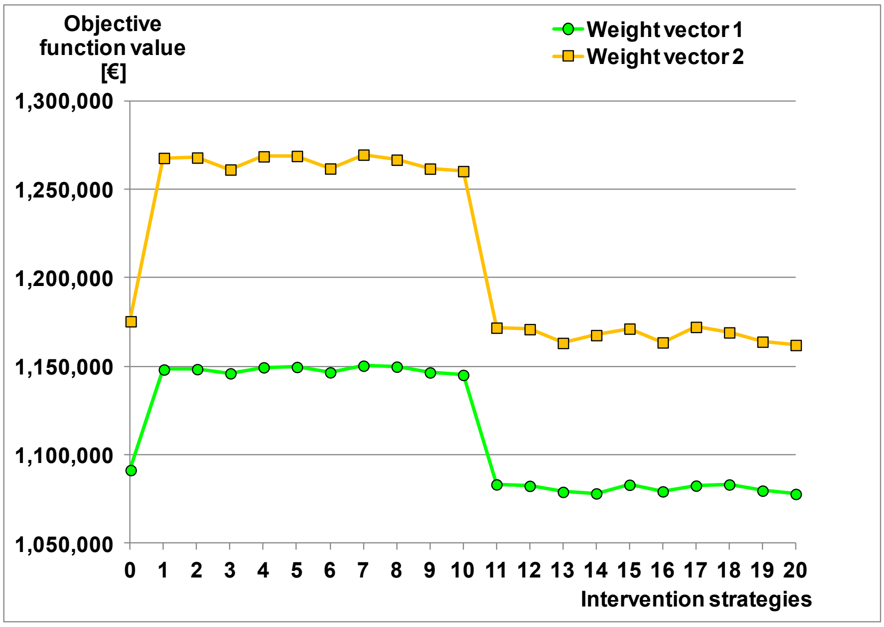


Figure 4.14 Comparison of different weight vectors for a high travel demand level

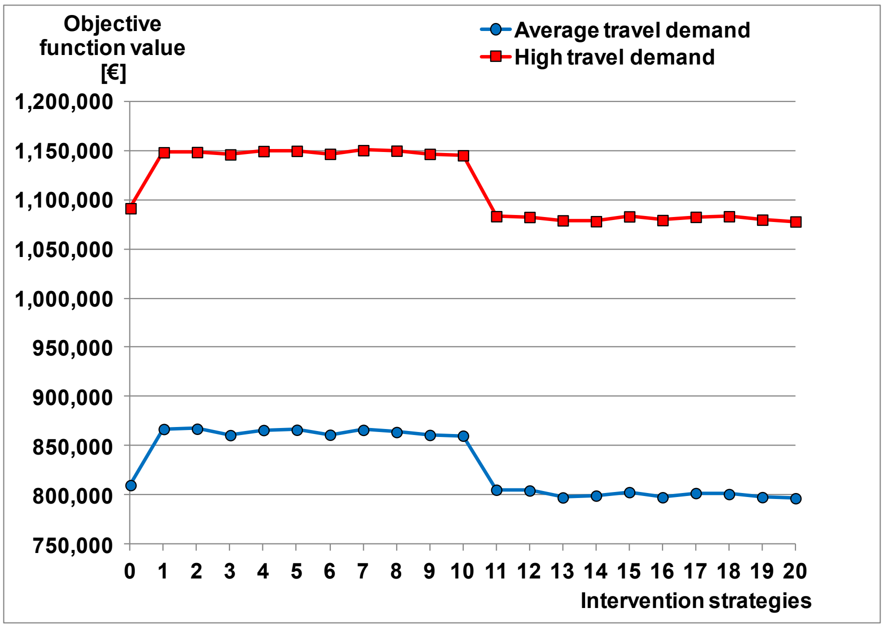


Figure 4.15 Comparison of different demand levels in the case of weight vector 1

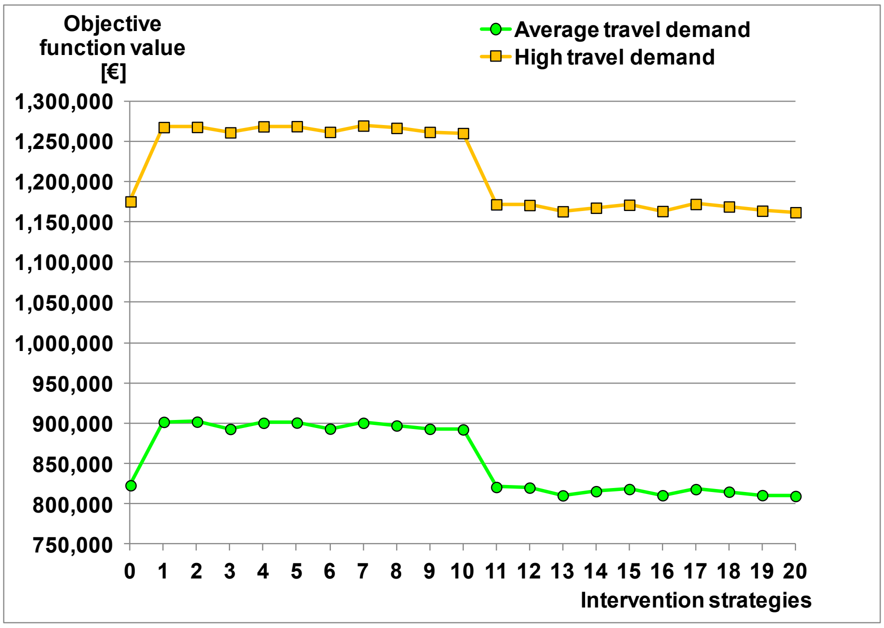


Figure 4.16 Comparison of different demand levels in the case of weight vector 2

Comparisons among optimal intervention solutions for different weight vectors are shown in figure 4.13 and figure 4.14, respectively in the case of average and high travel demand. Furthermore, comparisons among optimal intervention strategies for different travel demand levels are shown in figure 4.15 and figure 4.16, respectively in the case of weight vector 1 and weight vector 2.

To summarise, simulation results show that, in general, strategies which adopt a spare train are to be preferred; however, if this is not possible, due to the limited availability of rolling stock of the train operating company, the strategies which serve a major number of stops are desirable. Indeed, users prefer to arrive at their destination station, although with a reduced speed which clearly implies a higher on-board time, rather than to be unloaded and forced to wait the next train or to change transport mode. Letting the faulty train complete its trip, before recovering it, appears the best option also from an operational perspective. Indeed, the fact that the only depot available is near Piscinola station implies, in any case, the necessity of driving the faulty train until the terminus. In particular, it could be recovered during the operations, after unloading passengers, or at the end of service. However, in both case, this could lead to additional inconvenience. In fact, performing the recovery while on service generates additional perturbations to the ordinary conduct of rail operations; on the other hand, deciding to recover the convoy at the end of the service would require additional resources, both in terms of times and costs.

Obviously, objective function no. 2 is always higher than objective function no. 1 and, in both cases, the adoption of certain travel demand values affects the results, in confirmation of the fact that an accurate estimation of the involved passenger flows cannot be neglected in such an analysis. Furthermore, objective function no. 2 offers the possibility of emphasising one perspective rather than another, by properly setting the weight vectors. This allows to capture the trade-off between competing priorities, if any, thus adequately supporting each kind of decision-making process in an appropriate manner.

The above-mentioned outcomes have been obtained by adopting a deterministic approach and, therefore, a sensitivity analysis on the degree of robustness which they can guarantee is required. This evaluation can be performed by means of the stochastic framework described in paragraph 3.3.1. Hence, for each solution belonging to the identified neighbourhoods (see table 4.8), numerous microscopic simulations have been carried out, by changing stochastically the input parameters. Variability in acceleration, maximum speed and dwell times has been taken into account. Specifically, train performance (i.e. acceleration and speed) is modelled according to a piecewise linear distribution function: 33% of the trains are supposed to perform at 85%–90%, 33% at 90%–95%, and 34% at 95%–100%; while, dwell times at stations are defined according to a negative exponential random variable whose average is 10 seconds. Hence, in the light of the above mentioned assumptions, the objective function was computed again for several times and, then, results have been processed in order to derive the optimal intervention strategy for each case.

Table 4.9 shows the outcome of the performed sensitivity analysis. At first sight, the percentages seem to reflect what has been obtained with the deterministic approach. Indeed, strategy 20 is the one which ensures the minimum values of the objective function in three cases out of four; while, strategy 14 achieves a notable percentage only in one case. However, although stochastic results show that strategy 20 guarantees the highest level of robustness, upon a closer examination, it can be seen that it is the optimal solution at most in 43% of cases; hence, the deterministic approach misses the target in the remaining 57% of cases.

		Strategy 13	Strategy 14	Strategy 16	Strategy 20
Average travel demand level	**Weight vector 1**	30%	2%	31%	**37%**
	Weight vector 2	30%	1%	**35%**	34%
High travel demand level	**Weight vector 1**	26%	12%	24%	**38%**
	Weight vector 2	26%	6%	25%	**43%**

Table 4.9 Sensitivity analysis results

This confirms the significance of estimating the error degree of a purely deterministic procedure in order to be able to interpret the obtained results in a more accurate manner.

Besides stochasticity in service performance (e.g. speed and acceleration), planned timetable (e.g. delays, dwell time) and travel demand levels, it is worth considering also the randomness in the occurrence of a breakdown. More in detail, the proposed analysis has been performed once the location where the failure occurs has been fixed (i.e. Chiaiano station); however, the same event could show up at any other point of the network. Therefore, for properly considering this aspect, the so called *failure-strategy matrix* is introduced. It is a $(2n \times m)$ matrix, with n equal to the number of stations where the breakdown may occur and m equal to the number of strategies to be implemented.

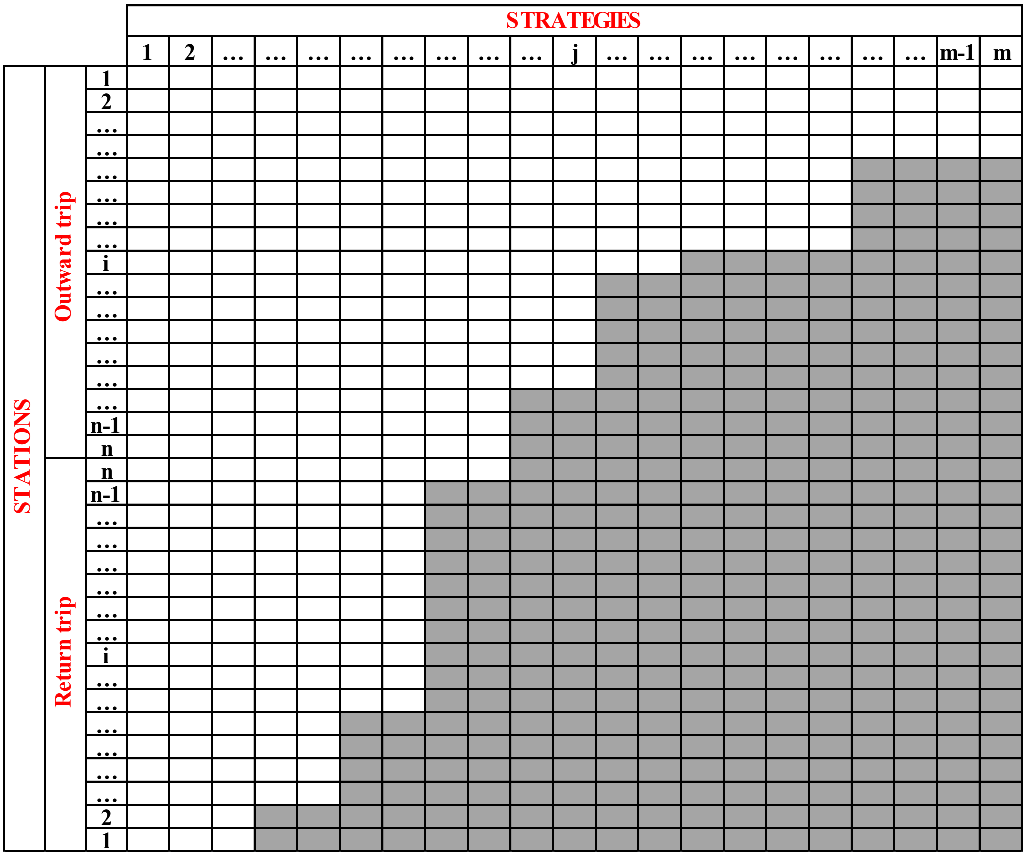


Figure 4.17 Failure-Strategy matrix

Obviously, as the train proceeds along its trip, the number of feasible strategies decreases, since the solutions involving stations already met become no longer practicable. Therefore, it can be stated that, by properly organising the structure to adopt, a sort of triangular matrix (since it is non-square) can be obtained, as shown in figure 4.17. The generic entry could be any kind of information which is possible to derive by using the proposed approach as, for instance, objective function values with the related robustness index computed by means of the above described sensitivity analysis. In this way, dispatchers would be able to evaluate also the degree of reliability offered by each solution and, given the structure of the proposed database, this results very useful. Indeed, as already stated, the main drawback is represented by the possibility that a specific condition to be addressed are not included in the database yet; however, the database could contain rescheduling contexts very similar to the one that is being faced. Thus, the additional available information on the degree of robustness, combined with the experience of dispatchers, could allow them to evaluate the transferability level of the intervention strategies listed in the database, as well as of the relative effects on rail operations. The introduced matrix is characterised by a discrete layout, since it considers only stations and not intermediate points between them; however, there is nothing to prevent the increase in spatial resolution implemented for simulating failure events, so as to enrich the adopted level of detail. Clearly, the overall view offered by such matrix structures can be similarly exploited for analysing different kinds of breakdown and any other issue of concern.

4.2.1 Evaluation of unconventional rescue strategies for managing disruption conditions

This paragraph shows an application on the Naples-Sorrento line aimed at investigating the technological feasibility of unconventional recovery strategies based on the use of operating rail convoys or bimodal rail-road maintenance vehicles (such as locotractors, diggers or catenary maintenance vehicles). In particular, the term '*unconventional*' concerns the fact that they are not allowed under the current Italian regulations. Obviously, the presence of these rescue vehicles affects the proper conduct of rail service and, therefore, influences passenger satisfaction. Hence, the challenge is to determine the intervention strategies which provide the right balance between the swiftness of rescue operations and the disturbance inflicted upon rail services during the failure management phase.

The proposed application focuses on a failure which makes the faulty convoy able to travel in non-autonomous conditions; this occurs, for instance, when the on-board traction system gets broken. In particular, the assumed failure scenario consists in considering that, during the morning peak hour, a train running from Sorrento to Naples (which represents the most loaded direction in terms of passenger flows) breaks down and, therefore, it is forced to stop at Scrajo station, where all on-board passengers have to alight. It is worth noting that, since this station between Pozzano and Vico Equense has no points, an additional issue to be considered regards the necessity of picking up passengers who were unloaded from the faulty train.

The analysed rescue strategies, obviously involving the available rolling stock of the company which operates the line, consist in the following ten scenarios:

- *Scenario 1.1*, based on the use of a diesel locomotive with a power of 260 kW, located at the Pascone depot. The diesel vehicle is driven to Scrajo where it couples to the faulty train and, after changing direction, tows it to Castellammare. Rescheduling is then required to pick up passengers who were unloaded from the faulty train at Scrajo;
- *Scenario 1.2*, similar to scenario 2.1 but based on the use of a diesel locomotive with a lower power (i.e. 74 kW);
- *Scenario 2*, based on the use of a train not operating when the failure occurs: one of the empty convoys available at Sorrento (i.e. a convoy which has completed its run from Naples to Sorrento and is ready to start its service in the opposite direction according to the planned timetable) is driven to Scrajo where it couples to the faulty train and, after changing direction, tows it to Vico Equense. Rescheduling is then

required to pick up passengers who were unloaded from the faulty train at Scrajo;

- *Scenario 3.1*, based on using two electric locotractors (i.e. maintenance vehicles) located at Castellammare. The vehicles are driven to Scrajo where they couple to the faulty train and, after changing direction, return to Castellammare station where the faulty train is recovered. Rescheduling is then required to pick up passengers who were unloaded from the faulty train at Scrajo;
- *Scenario 3.2*, similar to scenario 3.1 but, in this case, the locotractors are initially located at Vico Equense;
- *Scenario 4.1*, based on the use of a train operating in the opposite direction with respect to the faulty convoy (i.e. from Naples to Sorrento) which has already gone past Scrajo (i.e. the station where the faulty convoy has stopped) when the failure occurs. The operating train interrupts its ordinary service at Piano di Sorrento station, where it unloads passengers. It then changes direction and proceeds empty to Scrajo for coupling to the faulty train, which is finally towed to Vico Equense in order to be recovered. In this case, an additional issue needs to be addressed. Indeed, the required rescheduling is twofold: in favour of users waiting at Scraio (i.e. passengers who were on board the faulty vehicle) and Piano di Sorrento (i.e. passengers who were on board the rescue vehicle). Passengers waiting at Piano di Sorrento station are picked up by a different train from the one which unloaded them before;
- *Scenario 4.2*, similar to scenario 4.1 but, in this case, passengers waiting at Piano di Sorrento are picked up by the same train which unloaded them before;
- *Scenario 5.1*, similar to scenario 4.1 but, in this case, the rescue convoy has not yet gone past Scrajo. Therefore, it interrupts its ordinary service at Pozzano, where it unloads passengers, and proceeds empty to Scrajo

for coupling to the faulty train. The coupled vehicle changes direction and finally the faulty train is recovered at Castellammare station;

- *Scenario 5.2*, as in scenario 5.1 but, in this case, passengers waiting at Pozzano are picked up by the same train which unloaded them before;
- *Scenario 6*, based on the use of a train operating in the same direction as the faulty convoy (i.e. from Sorrento to Naples) and which precedes it. Therefore, the operating train interrupts its ordinary service at Vico Equense, where it unloads passengers, and proceeds empty to Scrajo for coupling to the faulty train. The coupled vehicle then changes direction and the faulty train is finally towed to Vico Equense in order to be recovered. At this point, the rescue vehicle may restart its ordinary itinerary from Vico Equense to Naples, clearly under a properly rescheduled timetable.

Specifically, the first three strategies (i.e. from 1.1 to 2) are termed *ordinary* because they are allowed under the current regulations. They involve shunter locomotives or non-operating trains. By contrast, the other strategies (i.e. from 3.1 to 6) are termed *unconventional* because, as already stated, they involve vehicles which are currently not allowed to be used for rescue services (see EAV 2015a; 2015b).

Scenario	User generalised cost [€]
1.1	340,116
1.2	317,855
2	336,659
3.1	355,061
3.2	375,753
4.1	345,727
4.2	313,805
5.1	275,010
5.2	274,887
6	322,489

Table 4.10 User generalised cost for each recovery scenario

Therefore, by adopting the optimisation framework described in paragraph 3.1 and implementing the travel demand obtained by means of the procedure which will be illustrated in paragraph 4.5.2, it is possible to derive the user generalised cost (i.e. objective function 3.3) for each recovery scenario. Simulation results are shown in table 4.10.

The outcome of the procedure points out that the optimal intervention strategies are those involving operating trains, which offer a considerable towing capability and are able to reach higher speeds. Moreover, the fact that such convoys are able to provide a service for passengers, immediately after completing rescue operations, makes them the most appropriate choice in order to minimise user discomfort.

By contrast, the highest user generalised costs are provided by the use of electric locotractors. Indeed, such vehicles, contrasting with a low purchase price, as well as the possibility of being located at strategic points of the line and travelling on ordinary asphalt roads (in order to minimise the time required to reach the faulty train), offer a very limited towing capability. This implies a high disturbance to the ordinary conduct of rail operations and generates great inconvenience for users.

Regarding the recovery actions based on the use of operating trains, it is worth pointing out what follows. Although such strategies present the drawback of unloading passengers, both for the faulty and the recovery trains, simulation results show that passenger waiting times on platforms, when they alight from the faulty train, are actually lower than the delays incurred if they remain on board and the system is restored by means of a diesel locomotive or an empty but distant rail convoy. However, the alteration of passenger perception stays. For this reason, it is fundamental to introduce suitable info-mobility strategies for providing information to passengers on the development of rescue operations.

Obviously, for allowing an effective utilisation of operating trains as rescue vehicles, amendments to current Italian regulations are required. However, given the high level of complexity and automation achieved in all transportation fields, it can be stated that this represents a mere formality.

4.3 Energy saving policies applications

In what follows, the analytical methodology introduced in paragraph 3.3.2, as a decision support tool for the implementation of eco-driving strategies, is applied to the Line 1 metro system.

Figure 4.18 shows the elevation profile of the line; while, figure 4.19 provides the layout of terminus stations: in both cases, a great asymmetry, which clearly influences energy consumption in the two trip directions, can be noted.

Table 4.11 provides numerical values of all operational parameters being calculated by means of microscopic simulations. In particular, travel, dwell and inversion times were obtained by adopting a deterministic method; while, for the computation of buffer times, a stochastic approach is required because of their function of recovering delays. Therefore, by implementing several stochastic simulations, taking into account the randomness of train performance, dwell times and delays, it is possible to derive the statistical distribution of all involved parameters and, thus, determine buffer times as function of an assumed confidence level.

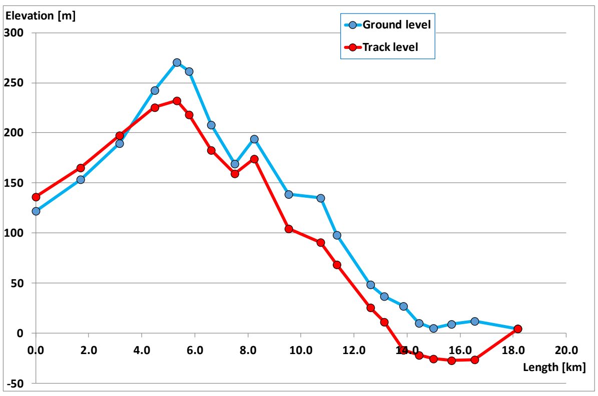


Figure 4.18 Elevation profile of Line 1

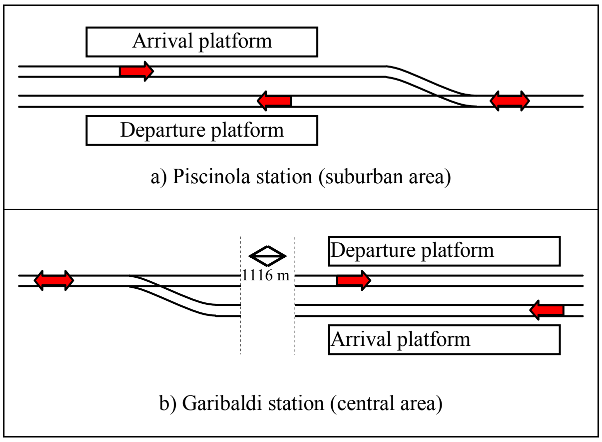


Figure 4.19 Layout of terminus stations

Parameter	Value	
	Piscinola-Garibaldi direction	***Garibaldi-Piscinola direction***
Travel distance	18.791 *km*	18.616 *km*
Total travel time	1,463 *s* [24.4 *min*]	1,485 *s* [24.8 *min*]
Total dwell time	400 *s* [6.7 *min*]	400 *s* [6.7 *min*]
Inversion time	307 *s* [5.1 *min*]	268 *s* [4.5 *min*]
Buffer time [90th percentile]	116 *s* [1.9 *min*]	103 *s* [1.7 *min*]
Buffer time [95th percentile]	131 *s* [2.2 *min*]	116 *s* [1.9 *min*]
Buffer time [99th percentile]	159 *s* [2.7 *min*]	141 *s* [2.4 *min*]
Cycle time [90th percentile]	4,542 *s* [75.7 *min*]	
Cycle time [95th percentile]	4,570 *s* [76.2 *min*]	
Cycle time [99th percentile]	4,623 *s* [77.1 *min*]	
Minimum headway	307 *s* [5.1 *min*]	
Energy consumption	279.01 *kWh*	386.21 *kWh*

Table 4.11 Operational parameters of Line 1

For this purpose, let δ^{i}_{ot} and δ^{i}_{rt} be the difference in performance between the stochastic and the deterministic travel times, respectively, in the case of outward and return trip, at *i*-th stochastic simulation. Specifically, they can be calculated as follows:

$$\delta^{i}_{ot} = \left(\sum_{lot} tt^{i,STOC}_{lot} + \sum_{sot} dt^{i,STOC}_{sot} + it^{i,STOC}_{ot} \right) - \left(\sum_{lot} tt^{DET}_{lot} + \sum_{sot} dt^{DET}_{sot} + it^{DET}_{ot} \right) \quad (4.1)$$

$$\delta_{rt}^{i} = \left(\sum_{lrt} tt_{lrt}^{i,STOC} + \sum_{srt} dt_{srt}^{i,STOC} + it_{rt}^{i,STOC} \right) - \left(\sum_{lot} tt_{lrt}^{DET} + \sum_{srt} dt_{srt}^{DET} + it_{rt}^{DET} \right) \quad (4.2)$$

where $X^{i,STOC}$ represents the value of variable X in the case of the i-th stochastic simulation; X^{DET} represents the value of variable X in the case of a deterministic simulation.

Since stochastic simulations are based on reductions in train performance, it can be stated that:

$$X^{i,STOC} \geq X^{DET} \quad \forall i \ \forall X \Rightarrow \begin{cases} \delta_{ot}^{i} \geq 0 & \forall i \\ \delta_{rt}^{i} \geq 0 & \forall i \end{cases} \quad (4.3)$$

Hence, by assuming that δ_{ot}^{i} and δ_{ot}^{i} are distributed according to a Normal (i.e. Gaussian) distribution, it is possible to calibrate function parameters (i.e. mean and variance), so as to reproduce observed data, by solving the following minimisation problems:

$$\left[\hat{\mu}_{ot}, \hat{\sigma}_{ot}^{2}\right] = \underset{\mu_{ot}, \sigma_{ot}^{2}}{argmin}\, Z_{ot}\left(\delta_{ot}^{i}, \mu_{ot}, \sigma_{ot}^{2}\right) \quad (4.4)$$

$$\left[\hat{\mu}_{rt}, \hat{\sigma}_{rt}^{2}\right] = \underset{\mu_{rt}, \sigma_{rt}^{2}}{argmin}\, Z_{rt}\left(\delta_{rt}^{i}, \mu_{rt}, \sigma_{rt}^{2}\right) \quad (4.5)$$

with

$$\sigma_{ot}^{2} \geq 0 \text{ and } \sigma_{rt}^{2} \geq 0 \quad (4.6)$$

where μ_{ot} and μ_{rt} are the means of the Normal distributions in the case of outward trip (*ot*) and return trip (*rt*); $\hat{\mu}_{ot}$ and $\hat{\mu}_{rt}$ are optimal values of μ_{ot} and μ_{rt}; σ_{ot}^{2} and σ_{rt}^{2} are the variances of the Normal distributions in the case of outward trip (*ot*) and return trip (*rt*); $\hat{\sigma}_{ot}^{2}$ and $\hat{\sigma}_{rt}^{2}$ are the optimal values of σ_{ot}^{2} and σ_{rt}^{2}; Z_{ot} is an objective function which expresses the gap between the cumulative distribution of observed values δ_{ot}^{i} and the cumulative distribution of the normal function of parameters $(\mu_{ot};\sigma_{ot}^{2})$; Z_{rt} is an objective function which expresses the gap between the cumulative distribution of observed values δ_{rt}^{i} and the cumulative distribution of the normal function of parameters $(\mu_{rt};\sigma_{rt}^{2})$.

μ_{ot}	σ_{ot}^{2}	μ_{rt}	σ_{rt}^{2}
64.114	40.803	56.922	36.247

Table 4.12 Normal distribution parameters

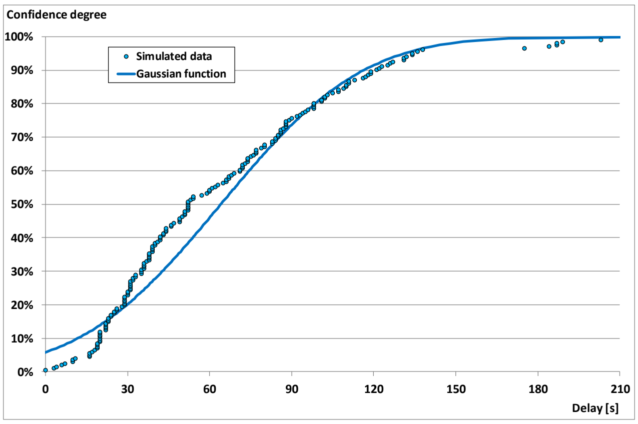


Figure 4.20 Comparison between cumulative distributions in the case of outward trip

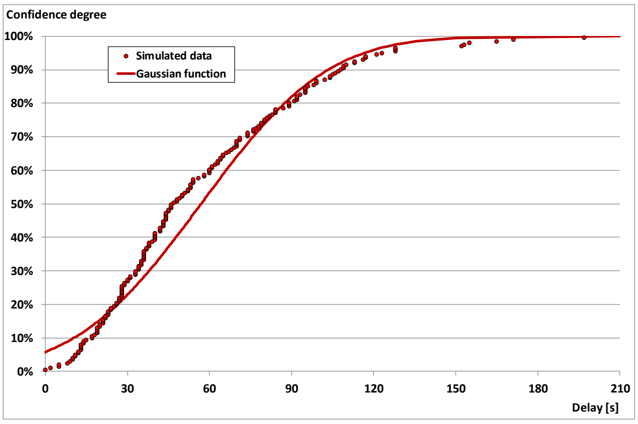


Figure 4.21 Comparison between cumulative distributions in the case of return trip

Results of the calibration phases (i.e. the solution of minimisation problems 4.4 and 4.5) are shown in table 4.12; while, comparisons between the cumulative distribution of observed values and the cumulative distribution of corresponding normal functions are proposed in figures 4.20 and 4.21.

These values were adopted for deriving buffer times in the case of confidence levels equal to 90th, 95th and 99th percentiles. Clearly, since the computation of cycle time involves buffer times, also this parameter was calculated according to the same confidence levels (see table 4.11).

The next step consists in analysing some operating schemes with the aim of comparing analytical results with those obtained by means of the micro simulation software OpenTrack, so as to validate the proposed methodology. In particular, once a value H (which has to meet at least the minimum headway requirement, indicated in table 4.11 as 5.1 min) has been fixed, it is possible to calculate the minimum and the maximum number of convoys, required to perform the service, by applying equations (3.34) and (3.35) in the case of layover times equal to 0. Then, once a feasible number of convoys has been fixed, the *turt* can be computed by means of equation (3.20). At this point, it is necessary to properly split the *turt* between the outward and the return trip, and, therefore, the feasible set for parameter α has to be computed by means of equation (3.27). In particular, within this range, the value of α (i.e. α_{opt}) which allows to obtain the lower H_{min} has been selected. Obviously, H_{min} will be different from the initial headway H. Indeed, adopting a certain value of α implies the definition of the layover times at the terminus stations and, since these times are spent by the convoy in a stop condition, the spacing allowed between trains, unavoidably, changes, as well as the minimum feasible headway according to equation (3.40). For this reason, it is necessary to perform a feasibility test of the analysed configurations which consists in verifying that the value of H_{min} is not higher than the initial headway H.

Results carried out in the case of the three considered percentiles (i.e. 90th, 95th and 99th) are provided in tables 4.13 - 4.15, where the red values indicate the unfeasible operating schemes. It is worth pointing out that analytical results present a full congruence with the simulation outcome, confirming the effectiveness of the developed analytical framework.

***H* [min]**	***NC*$_{min}$**	***NC*$_{max}$**	***NC***	***turt* [min]**	***α*$_{opt}$**	***H*$_{min}$ [min]**	***Test***
5.5	14	15	14	1.30	42.3%	5.12	OK
5.5	14	15	15	6.80	48.5%	6.87	NO
6.0	13	14	13	2.30	45.7%	5.12	OK
6.0	13	14	14	8.30	48.8%	7.62	NO
7.0	11	12	11	1.30	42.3%	5.12	OK
7.0	11	12	12	8.30	48.8%	7.62	NO
8.0	10	11	10	4.30	47.7%	5.62	OK
8.0	10	11	11	12.30	49.2%	9.62	NO
9.0	9	10	9	5.30	48.1%	6.12	OK
9.0	9	10	10	14.30	49.3%	10.62	NO
10.0	8	9	8	4.30	47.7%	5.62	OK
10.0	8	9	9	14.30	49.3%	10.62	NO
12.0	7	8	7	8.30	48.8%	7.62	OK
12.0	7	8	8	20.30	49.5%	13.62	NO
14.0	6	7	6	8.30	48.8%	7.62	OK
14.0	6	7	7	22.30	49.6%	14.62	NO
15.0	6	6	6	14.30	49.3%	10.62	OK

Table 4.13 Feasible configuration calculation in the case of 90th percentile

***H* [min]**	***NC*$_{min}$**	***NC*$_{max}$**	***NC***	***turt* [min]**	***α*$_{opt}$**	***H*$_{min}$ [min]**	***Test***
5.5	14	15	14	0.83	36.0%	5.12	OK
5.5	14	15	15	6.33	48.2%	6.87	NO
6.0	13	14	13	1.83	43.6%	5.12	OK
6.0	13	14	14	7.83	48.5%	7.62	NO
7.0	11	12	11	0.83	36.0%	5.12	OK
7.0	11	12	12	7.83	48.5%	7.62	NO
8.0	10	11	10	3.83	47.0%	5.62	OK
8.0	10	11	11	11.83	49.0%	9.62	NO
9.0	9	10	9	4.83	47.6%	6.12	OK
9.0	9	10	10	13.83	49.2%	10.62	NO
10.0	8	9	8	3.83	47.0%	5.62	OK
10.0	8	9	9	13.83	49.2%	10.62	NO
12.0	7	8	7	7.83	48.5%	7.62	OK
12.0	7	8	8	19.83	49.4%	13.62	NO
14.0	6	7	6	7.83	48.5%	7.62	OK
14.0	6	7	7	21.83	49.5%	14.62	NO
15.0	6	6	6	13.83	49.2%	10.62	OK

Table 4.14 Feasible configuration calculation in the case of 95th percentile

H [min]	NC_{min}	NC_{max}	NC	turt [min]	α_{opt}	H_{min} [min]	Test
5.5	15	15	15	5.45	47.4%	6.87	NO
5.5	15	15	16	10.95	48.7%	9.62	NO
6.0	13	14	13	0.95	35.1%	5.12	OK
6.0	13	14	14	6.95	48.0%	7.62	NO
7.0	12	12	12	6.95	48.0%	7.62	NO
7.0	12	12	13	13.95	49.0%	11.12	NO
8.0	10	11	10	2.95	45.2%	5.62	OK
8.0	10	11	11	10.95	48.7%	9.62	NO
9.0	9	10	9	3.95	46.4%	6.12	OK
9.0	9	10	10	12.95	48.9%	10.62	NO
10.0	8	9	8	2.95	45.2%	5.62	OK
10.0	8	9	9	12.95	48.9%	10.62	NO
12.0	7	8	7	6.95	48.0%	7.62	OK
12.0	7	8	8	18.95	49.3%	13.62	NO
14.0	6	7	6	6.95	48.0%	7.62	OK
14.0	6	7	7	20.95	49.3%	14.62	NO
15.0	6	6	6	12.95	48.9%	10.62	OK

Table 4.15 Feasible configuration calculation in the case of 99th percentile

The increases in total buffer time are equal to 0.47 minutes from the 90th to the 95th confidence level, equal to 0.88 minutes from the 95th to the 99th confidence level, and equal to 1.34 minutes from the 90th to the 99th confidence level. In particular, by comparing these quantities with the *turt* values the following can be stated. If the sum of layover times (i.e. *turt*) is higher than buffer time increases, the increase in total buffer time can be compensated by the reduction in total layover time and, thus, the sum of buffer and layover times (i.e. total reserve time), as well as the minimum headway, can be kept constant. On the other hand, if the *turt* is lower than the increase in buffer times, total reserve time cannot be kept constant and the configuration may be unfeasible.

4.4 Planning tasks: estimation of dwell times as flow-dependent factors

A fundamental task to be addressed in the case of rail systems is, undoubtedly, the timetable planning phase. It is aimed at carrying out a stable schedule, satisfying travel demand requirements and offering an appropriate degree of resilience, in order to be able to mitigate delays by avoiding their propagation, as well as other knock-on incidents. Within this framework, the evaluation of dwell time as a flow-dependent factor is crucial so as to obtain a reliable estimation of such parameter and adequately support the timetabling phase.

In the following, an application to the Line 1 metro system of the methodological framework described in paragraph 3.3.3 will be provided. It is worth pointing out that the choice of a metro network was no accident; indeed, because of the close distance between two successive stations, dwell times are generally comparable to travel times and, therefore, the need to perform an accurate estimation of them grows in importance.

As already shown, the reciprocal dependence between rail service and travel demand generates the snowball effects: the number of passengers on the platform influences the dwell times of trains at stations, which may cause delays; these, in turn, produce an increase in headways which generates more passenger flows on the platform providing a further extension of dwell times and, therefore, additional delays. Actually, the fact that the analysed metro context is characterised by an 8 minute headway for most of the day makes the snowball effect not so evident. Therefore, in order to verify, on one hand, the capacity of the developed methodology to capture this phenomenon and, on the other hand, the importance of estimating dwell time as function of the involved flows, in the proposed application we stressed the system by simulating denser timetables, with a decreasing value of headways between two successive convoys and a fixed planned dwell time (i.e. equal for each station and for each run), without any difference between peak hours and off-peak hours.

For each one of the considered timetables, the threefold interaction between passengers and trains is simulated and the fixed-point problem (3.56) is solved, so as to estimate dwell times for each station and for each run as flow-dependent values.

Specifically, two kinds of survey have been performed: a survey of passenger flows for estimating travel demand (see figure 4.9) and a survey of boarding/alighting flows, as well as train stop durations, for determining the passenger flow-dwell time function, depicted in figure 4.22.

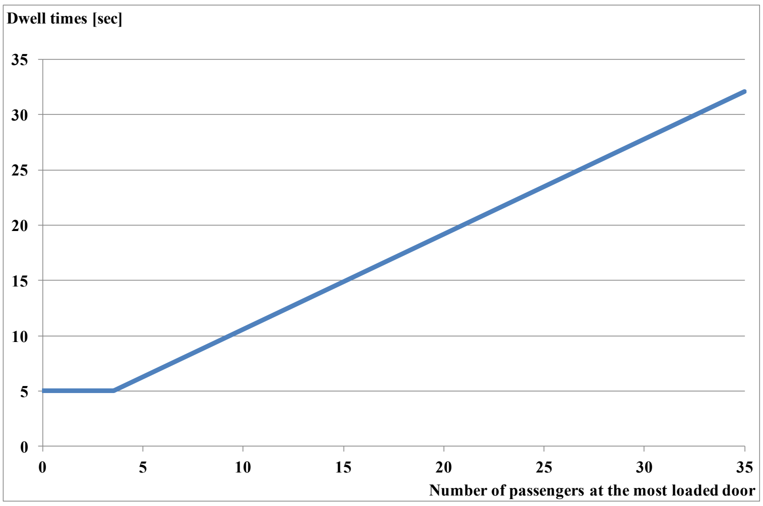


Figure 4.22 Dwell time calibration function

In particular, its analytical formulation is provided in the following equation:

$$dwt_s = \begin{cases} 5 & \text{if} \quad 0 \le td_s^{max} \le 3.5279 \\ 0.8602 \cdot td_s^{max} + 1.9653 & \text{if} \quad 3.5279 < td_s^{max} \le 2 \cdot Cap_{rc} \end{cases} \tag{4.7}$$

where td_s^{max} represents the sum of passengers boarding and alighting at the most loaded door; Cap_{rc} is the capacity of the rail coach which represents the maximum number of boarding passengers or, equivalently, the maximum number of alighting passengers. Hence, the worst case consists in a completely full coach, which first unloads all passengers and then loads them again (in this case, the number of transiting passengers is equal to $2 \cdot Cap_{rc}$). This implies that, since td_s^{max} has a maximum value, the dwell time is upper bounded.

In particular, in order to investigate the sensitivity of the proposed framework to different crowding conditions, three levels of travel demand have been considered (i.e. 85th, 90th and 95th percentiles of the distribution shown in figure 4.9).

Numerical values adopted for modelling involved capacity constraints are:

- 20 passengers per door (according to the experimental evidence during surveys);
- 216 passengers per carriage (according to the capacity of the rolling stock adopted on the line);
- wherever the capacity of the carriage was reached, the surplus is distributed in proportion to the residual capacity of the remaining wagons.

Once all required data have been collected, the resolution algorithm can be implemented. As already mentioned, generally, the procedure used for solving fixed point problems is the *MSA* algorithm (described in paragraph 2.4); however, in this case, given the nature of the involved functions, which do not satisfy the theoretical conditions ensuring the uniqueness of the solution, it is not possible to rule out that the algorithm diverges. In other words, it is not possible to demonstrate the convergence of the algorithm on a mathematical basis. Hence, a numerical evidence for assuring convergence properties has to be found or, alternatively, different resolution methods have to be implemented. Therefore, in addition to the *MSA* method, the *iterative algorithm* is analysed. In particular, an accurate assessment of their convergence properties, in the specific considered context, has been carried out.

As regards the iterative algorithm, there is no case in which its convergence can be guaranteed on a theoretical mathematical basis and this, at first, could suggest that such an approach is inappropriate as much as the *MSA* method. However, we can rely on the numerical evidence provided by Placido (2015) which implemented both resolution procedures in the case of the same real metro line (i.e. Line 1) and ascertained their convergence to the same configuration of dwell times as well.

Nevertheless, in this specific case, the fact that the *MSA* method produces decimal values at each iteration represents an additional drawback. Indeed, since the service simulation model is implemented by means of the OpenTrack software, decimal output values have to be rounded up/down to the nearest integer before being set up within the model and this generates a two-fold problem: algorithm convergence slows down and the iterative process assumes a discrete nature. In particular, the latter issue gives rise to a leak in the continuity of function $\psi(\cdot)$ indicated in equation (3.52) and, therefore, in the conditions ensuring the existence of solution.

Hence, in the light of the above, the resolution method selected is the iterative algorithm (see figure 4.23).

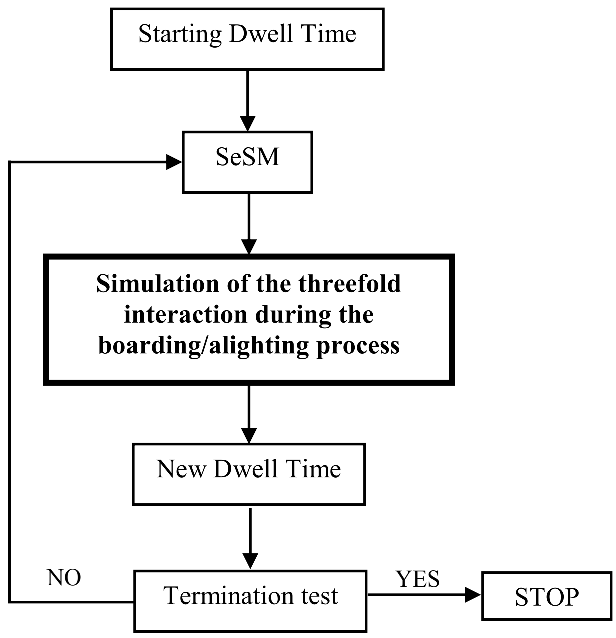


Figure 4.23 Iterative algorithm

The initialization of the algorithm occurs with a random value of the dwell time vector $\boldsymbol{dwt}^0$. Then, at the generic iteration *i*, according to the dwell time vector $\boldsymbol{dwt}^i$, headways are derived by the *SeSM*. Consequently, on the basis of the output of the *SeSM,* a new dwell time vector, $\boldsymbol{dwt}^{i+1}$, is established by simulating the threefold interaction between passengers and convoys.

The implemented termination test is:

$$\max_j \left(\frac{dwt_j^{i+1} - dwt_j^i}{dwt_j^i} \right) < 0.01 \text{ or } i > M \tag{4.8}$$

where M is a pre-fixed value indicating the maximum number of iterations (for instance, M = 1,000). If the test is verified, the algorithm stops; otherwise, the new headways are calculated. The significance of setting a termination test lies in the necessity of preventing the algorithm from performing an infinite number of iterations.

However, it is worth mentioning a downside of the proposed methodology: given the random nature of dwell times, resulting values should be considered as the expected values of dwell times needed for the boarding/alighting process and, therefore, no information concerning their statistical distribution is provided.

The iterative method has been implemented for each planned headway analysed (i.e. from 8 to 3 minutes) and for each level of travel demand (i.e. 50th, 85th and 95th percentile), amounting to a total of 18 processed scenarios. The number of runs for each planned headway (detailed for the outward and return trip) and the number of iterations required for solving the fixed-point problem, in the case of each planned headway and each travel demand level, are provided in table 4.16. It is worth noting that the convergence is reached for each analysed scenario, since the number of maximum iterations is never achieved. In particular, as can be seen, the number of iterations varies from a minimum of 6 to a maximum of 18, presenting a considerable increase with the reduction in the planned headway, due to a growing system instability. Clearly, the lower the planned headway to be analysed, the higher the number of runs and, consequently, the computational time required. However, since our methodology represents a support tool for design tasks, time is not a matter of concern.

Planned headway [min]	Number of runs		Travel demand level	Number of iterations
	Outward trips	Return trips		
8	23	19	50th	6
			85th	7
			95th	6
7	26	21	50th	6
			85th	8
			95th	6
6	30	25	50th	7
			85th	13
			95th	10
5	36	29	50th	17
			85th	10
			95th	15
4	45	36	50th	8
			85th	10
			95th	18
3	60	48	50th	8
			85th	15
			95th	12

Table 4.16 Feasible configuration calculation in the case of 99th percentile

The outcome of the procedure consists in estimating dwell times for each station and for each run. By way of illustration, tables 4.17 - 4.19 provide converging dwell times for a planned headway of 8 minutes, detailed for the three levels of travel demand analysed. Clearly, such a result has been derived also for other simulated planned headways. Moreover, for ensuring that the train is able to perform the outward trip, the return trip and the outward trip again, without being delayed, also dwell times at the first station have to be properly designed.

Dwell times present a changeable nature both along columns (i.e. the stations) and rows (i.e. the runs with a certain planned departure time associated), confirming the spatial and temporal variability by which they are affected. Furthermore, the same run at the same station could have a different dwell time on different days, according to the travel demand level at that time and, hence, the amount of flow involved in the boarding/alighting process.

Attention is drawn to the fact that values shown in the tables represent the dwell times reached at the end of the evolution of the snowball effect, which, as already pointed out, amplifies the involved quantities, and this justifies the presence of values which could be considered excessive in ordinary conditions (e.g. 90 seconds). Additionally, chaotic conditions generated by congestion, both during the boarding/alighting process and on-board, further magnify these values. Indeed, as shown by Weston (1989) and Douglas (2012), a certain number of users, which constitutes a mixed flow (i.e. some users must board and others must alight), requires a longer time, with respect to the case that they represent an unidirectional flow, for going through the same door. Moreover, the fact that users do not know exactly the position of the door in the moment in which the train is approaching induces them to walk randomly on the platform. Regarding on-board congestion, dwell times can increase also due to standing passengers close to a door or passengers who move inside the coach. This could happen because, in the analysed context, doors do not open on the same side in all stations. Therefore, results confirm that it is not possible to simply state that dwell times increase as the travel demand increases. This is due to the fact that, as already said, dwell time in a station depends on the arrival rates in that station, the arrival rates in the previous stations and the framework of travel demand (in terms of alighting flows); hence, it represents the converging value of an equilibrium procedure. This confirms the necessity of adopting suitable simulation techniques for accurately modelling such a complex and non-linear phenomenon.

The variation of dwell times between two successive runs may produce changes in headway in each station. However, the timetable was planned so as to ensure that headway was constant on average. In order to highlight this aspect, figures 4.24 - 4.26 provide, for each planned headway and for each travel demand level, the maximum and the minimum actual headways. Specifically, they are obtained by implementing the converging dwell times in the simulation model and, then, selecting maximum and minimum values among all resulting headways (i.e. for each run and for each station).

Additionally, it is shown that average values of actual headways coincide perfectly with planned ones.

Outward trip																	
STATION / RUN	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**
1	44	37	13	12	29	19	39	43	59	45	11	28	29	76	36	32	50
2	44	37	13	12	29	19	39	43	59	45	11	28	29	76	36	32	50
3	44	37	13	12	29	19	39	43	59	45	11	28	29	76	36	32	50
4	44	37	13	12	29	19	39	43	59	45	11	28	29	76	36	32	50
5	44	37	13	12	29	19	39	43	59	45	11	28	29	76	36	32	50
6	44	37	13	12	29	19	39	43	59	45	11	28	29	76	36	32	50
7	44	37	13	12	29	19	39	43	59	45	11	28	29	76	36	32	50
8	44	37	13	12	29	19	39	43	59	45	11	28	29	76	36	32	50
9	44	37	13	12	29	19	39	43	59	45	11	28	29	76	36	32	50
10	44	37	13	12	29	19	39	43	59	45	11	28	29	76	36	32	50
11	44	37	13	12	29	19	39	43	59	45	11	28	29	76	36	32	50
12	44	37	13	12	29	19	39	43	59	45	11	28	29	76	36	32	50
13	44	37	13	12	29	19	39	43	59	45	11	28	29	76	36	32	50
14	44	37	13	12	29	19	39	43	59	45	11	28	29	76	36	32	50
15	44	37	13	12	29	19	39	43	59	45	11	28	29	76	36	32	50
16	44	31	8	8	26	17	39	42	57	44	11	27	29	74	36	32	49
17	44	38	9	13	36	24	30	73	96	43	11	32	29	77	36	32	50
18	38	35	8	12	35	22	29	69	88	38	16	30	17	73	37	29	49
19	38	36	8	12	35	23	30	71	90	38	22	31	19	76	58	59	63
20	38	36	8	12	35	23	30	70	90	38	22	31	18	76	39	40	40
21	38	36	8	12	35	23	30	70	90	38	22	31	18	76	39	40	40
22	38	36	8	12	35	23	30	48	65	27	20	13	9	26	32	35	34
23	38	36	8	6	22	6	17	23	31	15	15	7	7	17	24	25	16
Return trip																	
STATION / RUN	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**
1	65	27	36	45	22	23	28	19	48	50	25	26	43	24	22	52	80
2	65	27	36	45	22	23	28	19	48	50	25	26	43	24	22	52	80
3	52	22	26	33	59	19	24	17	62	39	16	14	53	16	15	61	65
4	52	22	27	33	60	19	24	17	62	39	16	14	53	16	15	61	66
5	52	22	27	33	59	19	24	17	62	39	16	14	53	16	15	61	65
6	52	22	27	33	59	19	24	17	62	39	16	14	53	16	15	61	65
7	52	22	27	33	60	19	24	17	62	39	16	14	53	16	15	61	66
8	52	22	27	33	59	19	24	17	62	39	16	14	53	16	15	61	65
9	52	22	27	33	59	19	24	17	62	39	16	14	53	16	15	61	65
10	52	22	27	33	60	19	24	17	62	39	16	14	53	16	15	61	66
11	52	22	27	33	59	19	24	17	62	39	16	14	53	16	15	61	65
12	52	22	27	33	60	19	24	17	62	39	16	14	53	16	15	61	66
13	52	22	27	33	59	19	24	17	62	39	16	14	53	16	15	61	65
14	52	22	27	33	60	19	24	17	62	39	16	14	53	16	15	61	66
15	52	22	27	33	59	19	24	17	62	39	16	14	53	16	15	61	65
16	52	22	27	33	60	19	24	17	62	39	16	14	53	16	15	61	63
17	52	22	27	33	59	19	24	17	62	39	16	14	52	15	14	58	58
18	52	22	27	33	60	19	24	17	50	33	9	9	41	10	10	39	34
19	52	22	24	25	16	10	10	6	25	13	6	6	16	6	6	16	20

Table 4.17 Converging dwell times for a planned headway of 8 minutes - 50th percentile

Outward trip																	
STATION / RUN	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	44	35	14	18	43	17	22	52	66	55	11	35	19	65	36	34	59
2	44	35	14	18	43	17	22	52	66	55	11	35	19	65	36	34	59
3	44	35	14	18	43	17	22	52	66	55	11	35	19	65	36	34	59
4	44	35	14	18	43	17	22	52	66	55	11	35	19	65	36	34	59
5	44	35	14	18	43	17	22	52	66	55	11	35	19	65	36	34	59
6	44	35	14	18	43	17	22	52	66	55	11	35	19	65	36	34	59
7	44	35	14	18	43	17	22	52	66	55	11	35	19	65	36	34	59
8	44	35	14	18	43	17	22	52	66	55	11	35	19	65	36	34	59
9	44	35	14	18	43	17	22	52	66	55	11	35	19	65	36	34	59
10	44	35	14	18	43	17	22	52	66	55	11	35	19	65	36	34	59
11	44	35	14	18	43	17	22	52	66	55	11	35	19	65	36	34	59
12	44	35	14	18	43	17	22	52	66	55	11	35	19	65	36	34	59
13	44	35	14	18	43	17	22	52	66	55	11	35	19	65	36	34	59
14	44	35	14	18	43	17	22	52	66	55	11	35	19	65	36	34	59
15	44	35	14	18	43	17	22	52	66	55	11	35	19	65	36	34	59
16	44	35	9	14	31	13	23	55	67	60	11	31	19	66	36	34	59
17	44	50	10	16	41	18	42	64	72	47	11	31	18	68	36	34	58
18	44	45	10	16	45	32	47	84	87	48	8	31	18	79	31	32	56
19	33	42	9	16	46	34	50	82	92	50	12	32	25	74	54	60	66
20	33	42	9	16	46	34	50	82	92	50	12	32	25	74	36	41	44
21	33	42	9	16	46	34	50	82	91	50	12	32	25	74	36	41	44
22	33	42	9	16	46	34	50	54	54	33	9	14	21	44	29	36	31
23	33	42	9	7	28	7	22	29	24	19	8	8	20	14	24	31	20
Return trip																	
STATION / RUN	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	51	28	48	57	28	15	19	19	47	43	28	28	30	24	22	44	62
2	51	28	48	57	28	15	19	19	47	43	28	28	30	24	22	44	62
3	55	30	36	34	38	22	22	20	53	50	21	20	53	20	18	59	71
4	55	30	37	32	49	22	25	24	55	52	21	19	58	20	18	72	78
5	55	30	37	32	49	22	25	23	54	52	21	19	57	19	18	71	76
6	55	30	37	32	49	22	25	23	54	52	21	19	57	19	18	71	76
7	55	30	37	32	49	22	25	23	54	52	21	19	57	19	18	71	76
8	55	30	37	32	49	22	25	23	54	52	21	19	57	19	18	71	76
9	55	30	37	32	49	22	25	23	54	52	21	19	57	19	18	71	76
10	56	31	38	33	49	22	25	24	55	53	21	19	58	20	18	72	79
11	54	29	36	30	50	21	25	22	54	50	20	18	56	19	18	70	73
12	55	30	37	32	49	22	25	23	55	52	21	19	57	19	18	71	77
13	55	30	37	32	49	22	25	23	54	52	21	19	57	19	18	71	76
14	55	30	37	32	49	22	25	23	54	52	21	19	57	19	18	71	76
15	55	30	37	32	49	22	25	23	54	52	21	19	57	19	18	71	76
16	55	30	37	32	49	22	25	23	54	52	21	19	57	19	18	71	74
17	55	30	37	32	49	22	25	23	54	52	20	18	55	17	16	65	65
18	55	30	37	32	49	22	25	15	38	44	12	12	40	10	10	39	33
19	55	30	15	17	22	13	13	7	17	17	7	7	22	7	7	22	13

Table 4.18 Converging dwell times for a planned headway of 8 minutes - 85th percentile

Outward trip																	
RUN \ STATION	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**
1	51	50	16	15	35	16	24	44	59	35	10	39	19	63	37	34	58
2	51	50	16	15	35	16	24	44	59	35	10	39	19	63	37	34	58
3	51	50	16	15	35	16	24	44	59	35	10	39	19	63	37	34	58
4	51	50	16	15	35	16	24	44	59	35	10	39	19	63	37	34	58
5	51	50	16	15	35	16	24	44	59	35	10	39	19	63	37	34	58
6	51	50	16	15	35	16	24	44	59	35	10	39	19	63	37	34	58
7	51	50	16	15	35	16	24	44	59	35	10	39	19	63	37	34	58
8	51	50	16	15	35	16	24	44	59	35	10	39	19	63	37	34	58
9	51	50	16	15	35	16	24	44	59	35	10	39	19	63	37	34	58
10	51	50	16	15	35	16	24	44	59	35	10	39	19	63	37	34	58
11	51	50	16	15	35	16	24	44	59	35	10	39	19	63	37	34	58
12	51	50	16	15	35	16	24	44	59	35	10	39	19	63	37	34	58
13	51	50	16	15	35	16	24	44	59	35	10	39	19	63	37	34	58
14	51	50	16	15	35	16	24	44	59	35	10	39	19	63	37	34	58
15	51	50	16	15	35	16	24	44	59	35	10	39	19	63	37	34	58
16	51	51	11	10	26	11	24	44	59	28	10	37	19	61	37	34	58
17	51	41	15	12	46	13	37	50	55	42	10	32	18	58	37	34	57
18	51	38	36	16	42	19	40	69	55	44	8	33	18	52	25	32	55
19	39	33	29	44	50	23	43	70	70	47	12	41	26	81	32	41	58
20	39	34	11	50	43	22	43	67	63	49	12	39	26	61	19	27	36
21	39	34	11	40	39	22	46	75	70	52	12	36	26	79	19	27	36
22	39	34	11	18	51	17	51	46	31	28	9	16	22	36	12	22	23
23	39	34	10	8	31	8	24	32	12	21	9	9	22	16	12	22	23
Return trip																	
RUN \ STATION	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**
1	51	36	51	55	28	17	21	19	49	40	30	30	30	24	23	41	70
2	51	36	51	55	28	17	21	19	49	40	30	30	30	24	23	41	70
3	65	35	38	59	37	23	26	21	40	33	24	24	52	21	20	59	65
4	65	35	27	39	39	22	23	19	53	50	20	20	55	21	20	62	77
5	65	35	27	39	49	23	26	28	57	62	25	23	56	22	20	80	94
6	65	35	27	39	49	23	26	27	56	61	24	22	55	21	19	77	90
7	65	35	27	39	49	23	26	27	56	61	24	22	55	21	20	78	90
8	65	35	27	39	49	23	26	27	56	61	24	22	55	21	20	78	90
9	65	35	27	39	49	23	26	27	56	61	24	22	55	21	20	78	90
10	65	35	27	39	49	23	26	27	56	61	24	22	55	21	20	78	90
11	65	35	27	39	49	23	26	27	56	61	24	22	55	21	20	78	90
12	65	35	27	39	49	23	26	27	56	61	24	22	55	21	20	78	90
13	65	35	27	39	49	23	26	27	56	61	24	22	55	21	20	78	90
14	65	35	27	39	49	23	26	27	56	61	24	22	55	21	20	78	90
15	65	35	27	39	49	23	26	27	56	61	24	22	55	21	20	78	90
16	65	35	27	39	49	23	26	27	56	61	24	22	55	21	20	78	88
17	65	35	27	39	49	23	26	27	56	61	24	22	52	19	17	70	76
18	65	35	27	39	49	23	26	15	36	51	13	13	35	11	11	40	39
19	65	35	18	20	26	15	15	8	20	20	8	8	26	8	8	26	16

Table 4.19 Converging dwell times for a planned headway of 8 minutes - 95th percentile

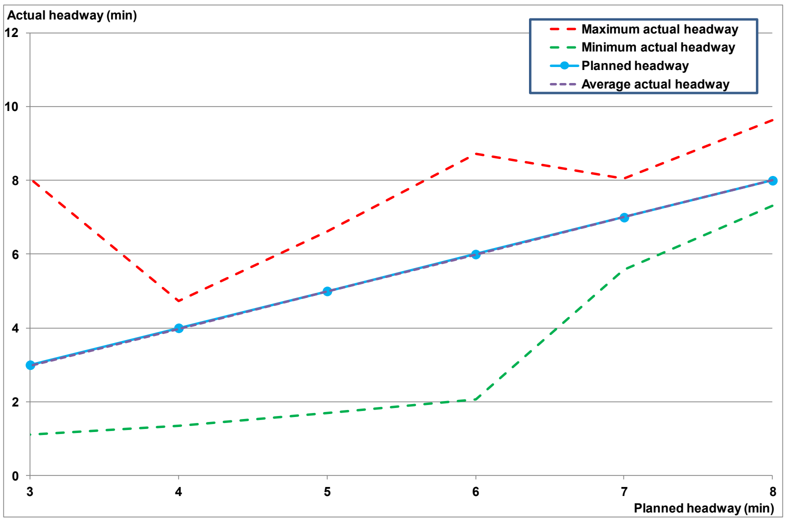


Figure 4.24 Fluctuation band of actual headway - 50th percentile

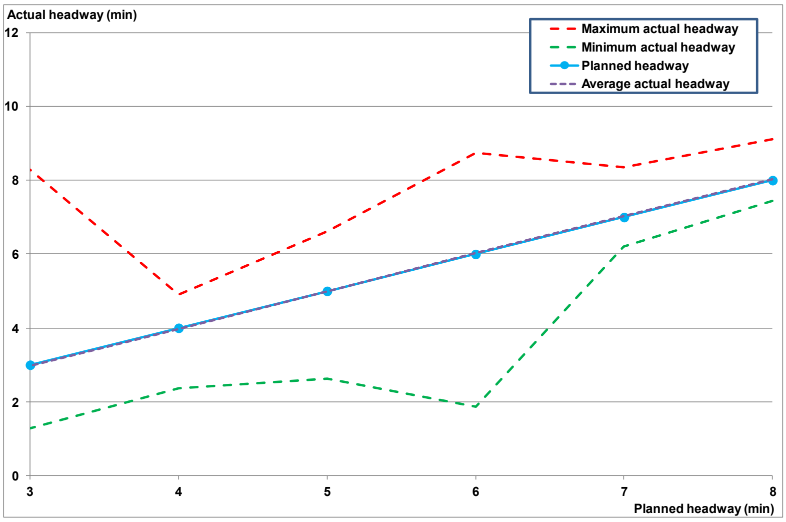


Figure 4.25 Fluctuation band of actual headway - 85th percentile

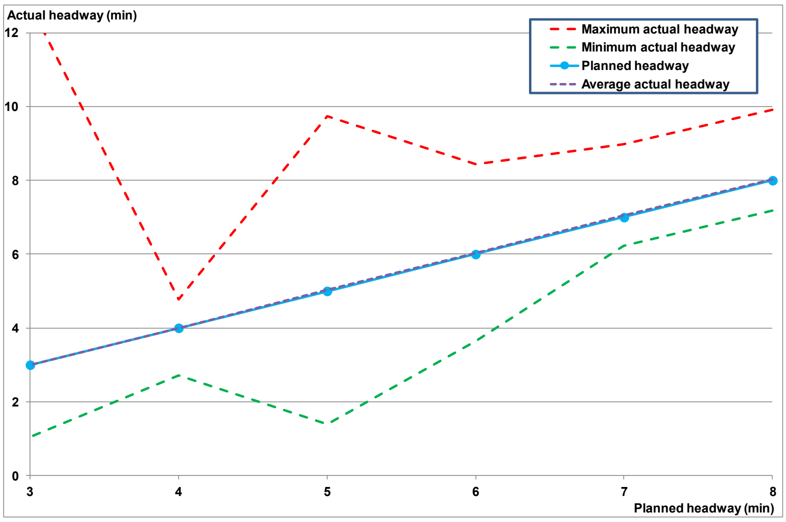


Figure 4.26 Fluctuation band of actual headway - 95th percentile

Results show that an inaccurate estimation of dwell times generates an unstable timetable and produces degradation in service performance, at the expense of users. Indeed, the fact that headways vary so pronouncedly, between low and high values, implies that some trains are very close, while others are very far apart (with an increase in both the mean and variance of passenger waiting times), as well as the presence of overcrowded trains followed by empty ones (i.e. platooning phenomenon).

In conclusion, the application to a real metro context has confirmed the ability of the proposed methodology of capturing the snowball effect, as well as its effectiveness in providing an accurate estimation of dwell times, as function of the involved flows, so as to properly support the timetable design phase.

4.4.1 A comparison between FIFO and RIFO queuing rules

The application described in this paragraph focuses on the simulation of passenger behaviour on platform, when a train arrives, in a metro context, i.e. Line 1 of the Naples metro system. The significance of this task lies in the fact that, either for planning a service or for mitigating negative impacts due to perturbed conditions, it is necessary to correctly model, beforehand, user reactions to alternative projects or rescheduling strategies.

Specifically, in the following, two different boarding priority patterns are analysed and compared, namely, the traditional FIFO approach (i.e. a passenger may board a train only after all passengers arriving before him/her have boarded the train) and the RIFO behaviour (i.e. passengers waiting on the platform tend to move around by mixing with respect to their arrival order, thus altering the initial queuing pattern). The two queuing rules above have been analysed with respect to different travel demand levels which were obtained by multiplying uniformly the average working-day demand by 15 values between 0.2 (i.e. 20%) and 3.0 (i.e. 300%). Obviously, in the case of a multiplier equal to 1.0 (i.e. 100%), the analysed demand level coincides with the average working-day demand.

In particular, in this specific application, the *KPIs* (*Key Performance Indexes*) set out below have been derived, for both behavioural approaches, by adopting the methodological framework proposed in paragraph 3.3.3.

Total On-Board Time (*TOBT*) represents the total time spent on the train by passengers during their trip. It may be formulated as:

$$TOBT = \sum_{l} \sum_{r} tb_l^r \cdot fb_l^r \qquad (4.9)$$

where tb_l^r is the average on-board time spent during run r on link l and fb_l^r is the number of passengers who travel on the rail convoy associated to run r on link l.

Likewise, *Total Waiting Time* (*TWT*) represents the total time spent by passengers on the platform waiting for a train. It may be formulated as:

$$TWT = \sum_{s} \sum_{p} \sum_{r} tw_{s,p}^{r} \cdot fw_{s,p}^{r} \tag{4.10}$$

where $tw_{s,p}^{r}$ is the average waiting time at station *s*, on platform *p* between run *(r–1)* and run *r* and $fw_{s,p}^{r}$ is the number of passengers waiting at station *s*, on platform *p* for run *r*.

Moreover, two objective functions have been calculated (i.e. *OF1* and *OF2*), which represent, respectively, the total time spent by passengers on the metro system and the total cost supported by passengers and the mass transit agency. In particular, they can be formulated according to the following expressions.

$$OF1 = \beta_{ob} \cdot TOBT + \beta_{w} \cdot TWT \tag{4.11}$$

where β_{w} (assumed equal to 2.5) is a parameter which describes user perception of the time spent waiting for trains with respect to the perception of the time spent on board, expressed by β_{ob} (assumed as unitary).

$$OF2 = UTC + NOC \tag{4.12}$$

with:

$$UTC = \beta_{VOT} \cdot (\beta_{ob} \cdot TOBT + \beta_{w} \cdot TWT) + TTC \tag{4.13}$$

$$NOC = TOC - RT \tag{4.14}$$

where *UTC* is the user total cost; *NOC* is the net operational cost; β_{VOT} is a parameter which expresses the monetary value of time (assumed equal to 5 €/h); *TTC* is the total ticket cost, that is the total expenditure incurred by passengers for the purchase of tickets; *TOC* is the total operational cost, that is the total expenditure incurred by the mass-transit agency for metro operations; *TR* is ticket revenue.

Equation (4.14) coincides with equation (3.85) and, therefore, the *TOC* can be calculated as shown by (3.6) or, alternatively, by (3.86). More in detail, since the proposed application is focused on Line 1 of the Naples metro system, whose standard cost is provided in terms of traction unit (i.e. 18.17 €/traction unit-km), the formulation here adopted is that given by equation (3.6). Moreover, it is worth pointing out that, since term *TTC* is always equal to term *TR*, in the definition of *OF2* they cancel each other out and, therefore, their calculation can be neglected.

Numerical results are shown in tables 4.20 and 4.21, and figures 4.27 - 4.32.

Travel demand multiplier	**Assigned travel demand [pax/day]**	**Unsatisfied travel demand [pax/day]**	**Total On-Board Time [h/day]**	**Total Waiting Time [h/day]**	**Objective Function no. 1 [h/day]**	**Objective Function no. 2 [€/day]**
0.2	42,411	–	9,205	3,006	16,719	228,817
0.4	84,823	–	18,410	6,012	33,439	312,413
0.6	127,234	–	27,615	9,017	50,158	396,010
0.8	169,645	–	36,820	12,023	66,877	479,607
1.0	212,056	–	46,025	15,055	83,661	563,526
1.2	253,898	569	55,130	26,249	120,751	748,975
1.4	291,860	5,018	63,580	90,230	289,156	1,590,999
1.6	319,785	19,505	69,916	160,021	469,968	2,495,058
1.8	347,807	33,894	75,874	249,388	699,344	3,641,942
2.0	372,765	51,347	81,138	393,097	1,063,880	5,464,620
2.2	386,184	80,339	84,070	476,644	1,275,680	6,523,618
2.4	397,988	110,947	86,290	550,163	1,461,697	7,453,704
2.6	407,364	143,982	87,848	601,872	1,592,528	8,107,858
2.8	417,561	176,197	89,118	721,169	1,892,040	9,605,418
3.0	427,658	208,511	90,217	817,240	2,133,319	10,811,81

Table 4.20 Simulation results in the case of the FIFO approach

Travel demand multiplier	Assigned travel demand [pax/day]	Unsatisfied travel demand [pax/day]	Total On-Board Time [h/day]	Total Waiting Time [h/day]	Objective Function no. 1 [h/day]	Objective Function no. 2 [€/day]
0.2	42,411	–	9,205	3,006	16,719	228,817
0.4	84,823	–	18,410	6,012	33,439	312,413
0.6	127,234	–	27,615	9,017	50,158	396,010
0.8	169,645	–	36,820	12,023	66,877	479,607
1.0	212,056	–	46,025	15,055	83,661	563,526
1.2	253,898	569	55,130	26,111	120,408	747,260
1.4	291,860	5,018	63,580	85,311	276,859	1,529,515
1.6	319,999	19,291	69,911	141,934	424,747	2,268,954
1.8	347,874	33,828	75,873	225,084	638,584	3,338,141
2.0	372,883	51,230	81,100	352,270	961,774	4,954,089
2.2	386,325	80,198	84,009	412,265	1,114,671	5,718,575
2.4	397,631	111,304	86,173	464,205	1,246,686	6,378,651
2.6	407,162	144,184	87,739	507,172	1,355,669	6,923,563
2.8	416,054	177,704	88,898	573,993	1,523,882	7,764,630
3.0	423,831	212,337	89,862	619,329	1,638,185	8,336,145

Table 4.21 Simulation results in the case of the RIFO approach

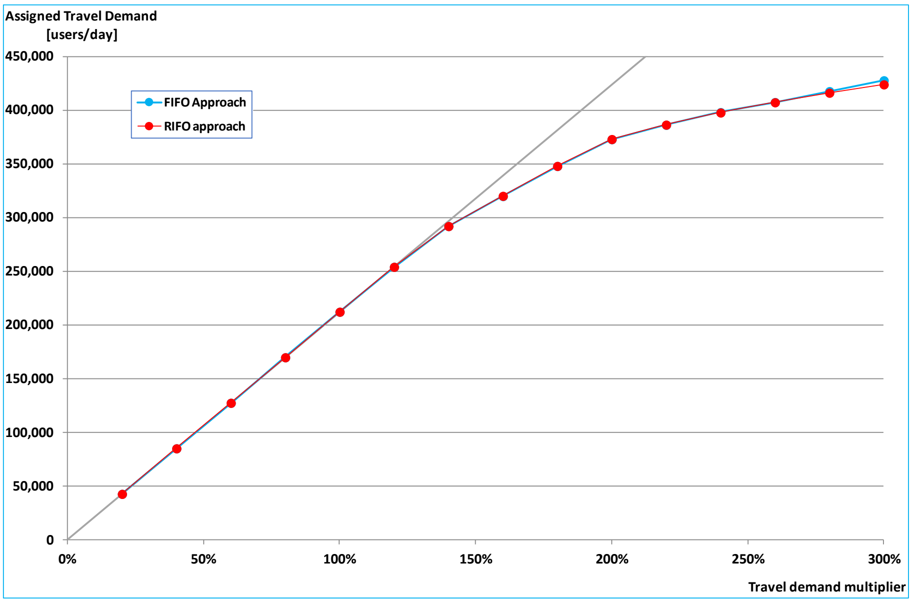


Figure 4.27 Assigned travel demand

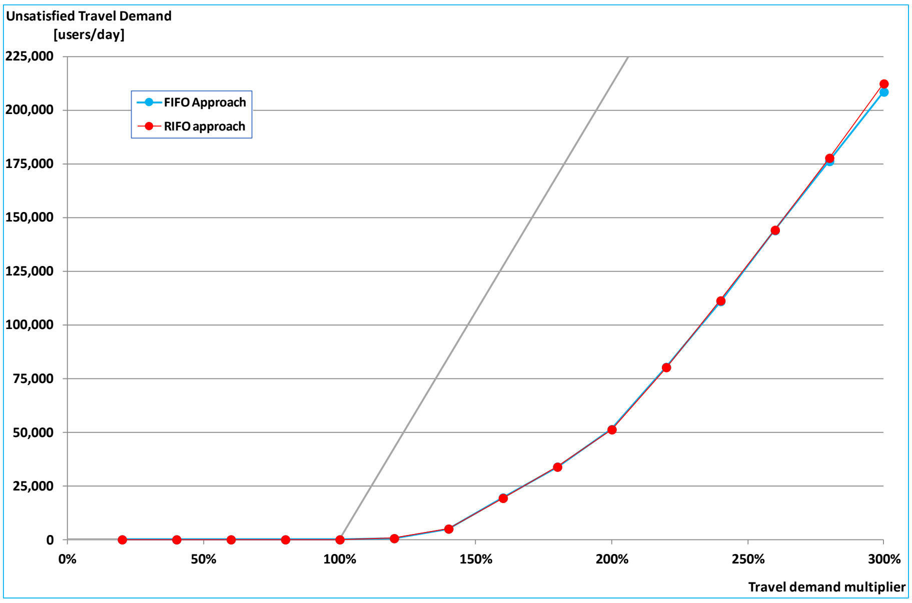


Figure 4.28 Unsatisfied travel demand

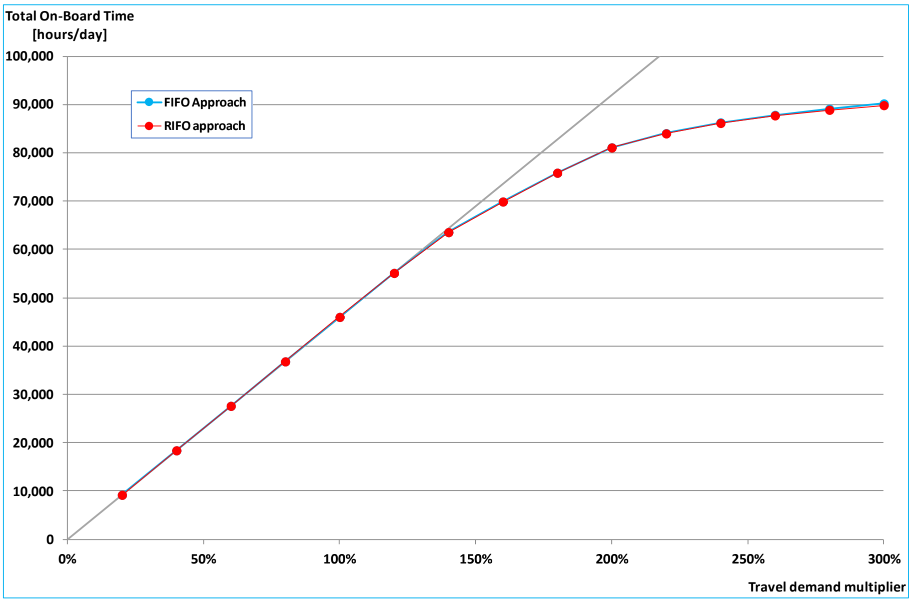


Figure 4.29 Total On-Board Time (TOBT)

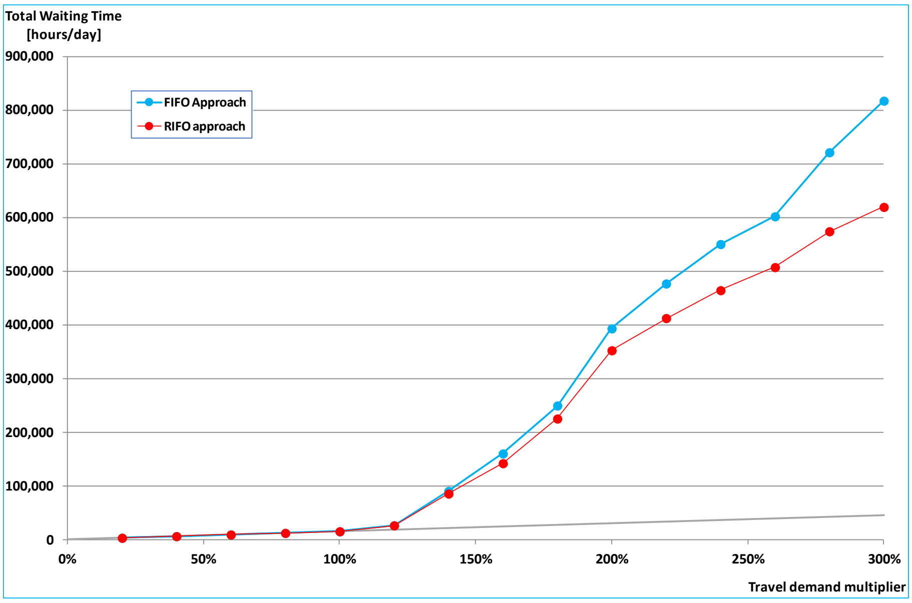


Figure 4.30 Total Waiting Time (TWT)

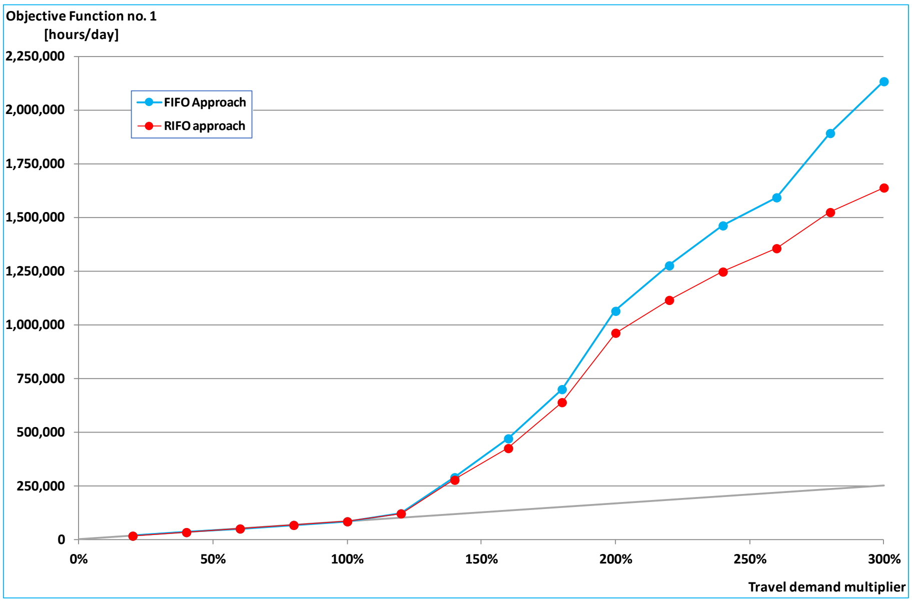


Figure 4.31 Values of objective function 1

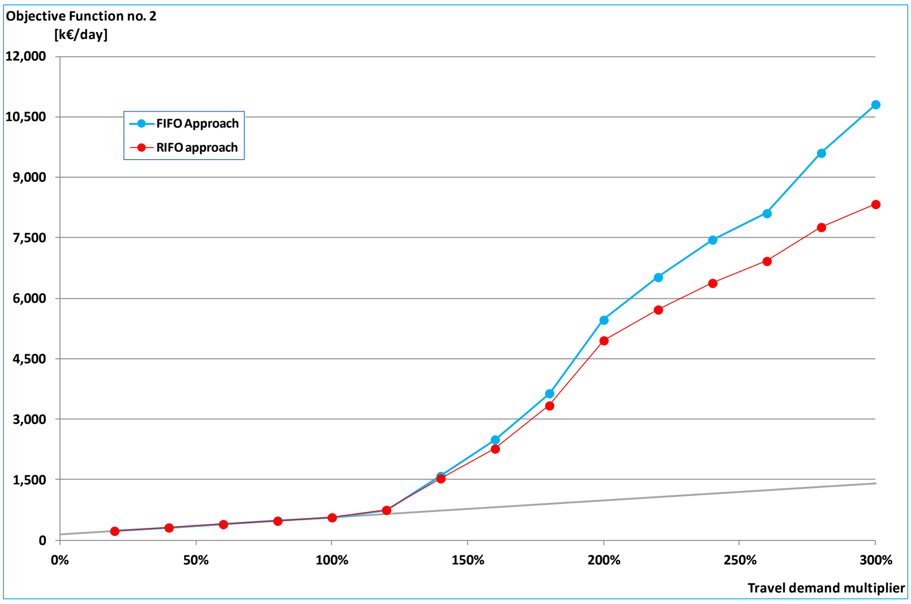


Figure 4.32 Values of objective function 2

In figure 4.27, the grey line represents the increase in travel demand according to the value of multipliers. As can be seen, in the case of multipliers lower than 100% (which represents the current demand), all passengers are able to board a train, or, in other words, all demand is assigned, and the two approaches provide the same results. While, for higher multipliers, the amount of assigned travel demand decreases, but there is no discontinuity, since the increase in passengers tends to fill the trains with a residual capacity still available. Also in this case the two approaches provide similar results.

In figure 4.28 the un-assigned demand is depicted and the grey line represents the increase in passengers with respect to the current condition. For values lower than 100%, it is null because there is no surplus in demand. Indeed, both approaches assign all passengers and provide the same results. By contrast, in the case of values higher than 100%, passengers tend to fill convoys which are not perfectly full (i.e. non-saturated trains) and, thus, values tend to develop an asymptotic behaviour with respect to the grey line where the slant asymptote is shifted as a function of the residual capacity in the case of 100%.

In figures 4.29 and 4.30, the grey line represents the increase, respectively, in *TOBT* and *TWT,* under the assumption of absence of capacity constraints, which implies that all passengers are able to board the first arriving train.

In particular, regarding the *TOTB,* the two approaches have a performance similar to the assigned travel demand, since running times of convoys are assumed as constant. Indeed, in this application, the dependence of dwell times on the number of boarding/alighting users, addressed in paragraph 3.3.3, has been neglected.

On the other hand, the *TWT* is affected by the introduction of capacity constraints, since they enable passengers to board the first arriving train and force them to wait for successive convoys. Moreover, the adoption of different queuing rules has an impact on waiting times, as can be seen by the gap between the red and blue lines in figure 4.30. Therefore, it can be stated that the increase in travel demand provides considerable increases in waiting times, which are further affected by the adopted behavioural pattern.

Finally, figures 4.31 and 4.32 show values of the two computed objective functions and the grey lines represent the objective function being calculated in the absence of capacity constraints. In particular, the performance of these functions is similar to that of the total waiting time, both in terms of increases and in terms of difference between the two approaches, since the *TWT* term is the predominant rate with respect to the others.

In conclusion, it can be stated that waiting time is the parameter mostly affected by the adopted queuing rule and that the greater the congestion level, the greater the difference in results between the two approaches. However, it is worth pointing out that the analysed queuing rules represent two extreme conditions, while, in real cases, generally, some passengers follow a FIFO approach and some others a RIFO approach. In particular, the distribution of passengers who adopt the FIFO rule with respect to those adopting the RIFO rule could be derived by means of turnstile data.

4.5 Travel demand estimation applications

The following applications are based on the methodological frameworks for handling travel demand flows described in paragraphs 3.3.4.1 and 3.3.4.2, which have been applied, respectively, in the case of the Line 1 metro system and the Naples-Sorrento regional line.

It is worth noting that, in the first case (i.e. metro system), the issue of travel demand estimation has been addressed at an urban level; on the other hand, for the regional context, the scale of the problem increases, since extra-urban (i.e. rural) trips have to be modelled. Moreover, the approach proposed in the first application is able to support short-term interventions (e.g. fare policies); differently, the procedure adopted in the second application is aimed at properly supporting economic evaluations on the feasibility of long-term projects.

Specifically, passenger flows extended with the approach proposed in paragraph 3.3.4.1 have been used for improving the preliminary estimation of travel demand related to the analysed metro system (figure 4.9); while, travel demand obtained in the case of the Naples-Sorrento line, for the current scenario (i.e. referred to the year 2016), has been implemented in the application addressing the evaluation of unconventional rescheduling strategies under perturbed conditions (see paragraph 4.2.1).

4.5.1 Calibration and validation of space-time relations representative of passenger flow data

The following application consists in identifying some mathematical relations, expressing boarding and alighting flows of passengers depending on the station (space component) and the time period considered (time component), properly calibrated for reproducing, analytically, the space-time variability of passenger flows in a metro context. The goal is to enable a decrease in the amount of data to be collected during the survey phase (which, clearly, implies reductions in related times and costs), without prejudicing analysis accuracy.

The adopted procedure is based on the phases described in paragraph 3.3.4.1, which are synthetically set out below again, for the sake of simplicity:

- designing and executing a survey campaign;
- simulating a certain sampling rate, obtained by assuming some data as not detected;
- performing a mono-dimensional statistical analysis on the partial data set for identifying the optimal functional form;
- performing a multi-dimensional statistical analysis on the partial set in order to specify, calibrate and validate one or more space-time functions;
- validating the methodology by comparing simulation results obtained by using the whole set of the surveyed data (considered as the *absolute truth*) with those produced by processing data of the calibrated space-time surfaces.

The required survey activities were implemented in July 2015 to collect data related to daily flows on an average working day in summer. It is worth pointing out that investigations were organised to detect flows for each single access (gate, stair, elevator, etc.), which were subsequently grouped according to platforms and travel directions. This implies the necessity of detecting data for all the accesses to each platform; otherwise, the distribution of the entire platform gets vitiated. Specifically, 3 time periods and 18 stations have been considered; therefore, the output of the survey phase consists in four matrices, of dimensions (3×18), whose framework is that shown in figure 3.7. The simulated sampling rate is assumed equal to 50%.

Then, the mono-dimensional statistical analysis has been performed as follows. Each one of the four surveyed matrices (for two kinds of passenger flow, i.e. boarding and alighting flows, and two kinds of trip, i.e. outgoing and return trips) has been split into 54 (i.e. 3×18) vectors and, consequently, the goodness of fit (i.e. the discrepancy between surveyed data and function data) offered by different classes of functions, with respect to each vector, has been tested.

Specifically, linear, quadratic, cubic, fourth-degree polynomial, fifth-degree polynomial, power, logarithmic and exponential functions have been evaluated. However, due to the scarcity of data along the matrix columns (i.e. there were at most two data), linear functions were adopted only in row analyses. The goodness of fit of each class of function was estimated by means of the coefficient of determination $\Re^2$, calculated as shown in equation (3.58).

The category which provided the best $\Re^2$ values, in both dimensions, was that of polynomial functions and, therefore, such a formulation has been implemented in the next step (i.e. multi-dimensional statistical analysis). In particular, the obtained relations are:

$$f_1^k(x,y) = a_1^k x^2 y^3 + a_2^k x\ y^4 + a_3^k x\ y^3 + a_4^k y^4 + a_5^k y^3 \tag{4.15}$$

$$\begin{aligned} f_2^k(x,y) = {} & b_1^k x^4 + b_2^k x^3 y + b_3^k x^2 y^2 + b_4^k x\ y^3 + b_5^k y^4 + b_6^k x^3 + b_7^k x^2 y + \\ & + b_8^k x\ y^2 + b_9^k y^3 + b_{10}^k x^2 + b_{11}^k x\ y + b_{12}^k y^2 + b_{13}^k x + b_{14}^k y + b_{15}^k \end{aligned} \tag{4.16}$$

with

$k \in \{AG\ ;\ BG\ ;\ AP\ ;\ BP\}$

where x represents the time period; y represents the sequence of stations ($y = 1$ in the case of Piscinola and $y = 18$ in the case of Garibaldi); AG represents alighting flows (A) in the case of the outgoing trip (i.e. Garibaldi direction); BG represents boarding flows (B) in the case of the outgoing trip (i.e. Garibaldi direction); AP represents alighting flows (A) in the case of the return trip (i.e. Piscinola direction); BP represents boarding flows (B) in the case of the return trip (i.e. Piscinola direction). Attention is drawn to the fact that equation (4.15) is defined in the case of $k \in \{AG\ ;\ AP\ ;\ BP\}$, while equation (4.16) is defined only in the case of $k = BG$. Hence, values f_i are computed by means of equation (4.15) or equation (4.16) on the basis of the values assumed by parameter k.

Once time-space relations have been specified, a polynomial regression has been carried out for defining the numerical values of the involved parameters (i.e. a_i and b_i) and both global and coefficient statistical tests have been derived.

The implemented global statistical tests are: $\Re^2$, shown in equation (3.58), $\overline{\Re^2}$, shown in equation (3.60), and F-test (indicated as F), formulated as follows:

$$F = \Re^2 \cdot (n-p-1) / \left(\left(1-\Re^2\right) \cdot p \right) \tag{4.17}$$

where p expresses the number of function parameters, which is equal to 5 in the case of function (4.15) and equal to 15 in the case of function (4.16). Moreover, the t-student (indicated as t) test of coefficients is performed:

$$t = \left|a_i^k\right| / \sqrt{Var\left(a_i^k\right)} \quad \text{or} \quad t = \left|b_i^k\right| / \sqrt{Var\left(b_i^k\right)} \tag{4.18}$$

where $Var(a_i^k)$, or equivalently $Var(b_i^k)$, is the i-th element of the main diagonal of variance-covariance matrix Σ, obtained as:

$$\Sigma = -\left[\partial^2 \varepsilon / \left(\partial a_i^k \, \partial a_j^k\right)\right]^{-1} \quad \text{or} \quad \Sigma = -\left[\partial^2 \varepsilon / \left(\partial b_i^k \, \partial b_j^k\right)\right]^{-1} \tag{4.19}$$

with

$$\varepsilon = \sum_i \left(f_i - \varphi_i\right)^2 \tag{4.20}$$

where φ_i is the i-th simulated surveyed data.

Function type (k)	$\Re^2$	$\overline{\Re^2}$	***F*-test**		
			***F* value**	**Threshold**	**Confidence level**
AG	0.764	0.701	12.273	8.018	99.90%
BG	0.829	0.572	3.220	3.190	94.00%
AP	0.621	0.521	6.218	5.967	99.50%
BP	0.514	0.386	4.023	4.016	97.00%

Table 4.22 Global statistical tests

Parameter	a_1^k	a_2^k	a_3^k	a_4^k	a_5^k
Value	0.0624	0.0184	–0.5150	–0.0272	0.7285
t-value	489.62	588.89	706.461	398.376	588.973
Threshold	5.077	5.077	5.077	5.077	5.077
Confidence level	99.99%	99.99%	99.99%	99.99%	99.99%

Table 4.23 Coefficient statistical tests of coefficients: *AG* condition

Parameter	b_1^k	b_2^k	b_3^k	b_4^k	b_5^k
Value	136.82	–71.15	3.82	1.05	0.16
t-value	$2.381 \cdot 10^6$	908.94	367.44	$2.892 \cdot 10^6$	$1.845 \cdot 10^8$
Threshold	6.412	6.412	6.412	6.412	6.412
Confidence level	99.99%	99.99%	99.99%	99.99%	99.99%
Parameter	b_6^k	b_7^k	b_8^k	b_9^k	b_{10}^k
Value	–557.32	362.08	–48.62	–8.59	129.16
t-value	7285.76	958.53	1220.60	5718.63	3016.26
Threshold	6.412	6.412	6.412	6.412	6.412
Confidence level	99.99%	99.99%	99.99%	99.99%	99.99%
Parameter	b_{11}^k	b_{12}^k	b_{13}^k	b_{14}^k	b_{15}^k
Value	–201.01	172.87	380.50	–970.94	2035.15
t-value	342.91	3513.05	465.52	$1.040 \cdot 10^8$	$1.063 \cdot 10^5$
Threshold	6.412	6.412	6.412	6.412	6.412
Confidence level	99.99%	99.99%	99.99%	99.99%	99.99%

Table 4.24 Coefficient statistical tests of coefficients: *BG* condition

Parameter	a_1^k	a_2^k	a_3^k	a_4^k	a_5^k
Value	–0.0552	0.0096	0.0677	–0.0585	0.7942
t-value	375.497	184.120	69.562	525.132	468.483
Threshold	5.077	5.077	5.077	5.077	5.077
Confidence level	99.99%	99.99%	99.99%	99.99%	99.99%

Table 4.25 Coefficient statistical tests of coefficients: *AP* condition

Parameter	a_1^k	a_2^k	a_3^k	a_4^k	a_5^k
Value	–0.0400	–0.0006	0.1864	–0.0142	0.1726
t-value	332.221	14.512	234.541	165.639	124.745
Threshold	5.077	5.077	5.077	5.077	5.077
Confidence level	99.99%	99.99%	99.99%	99.99%	99.99%

Table 4.26 Coefficient statistical tests of coefficients: *BP* condition

Table 4.22 provides results of the global statistical tests; while, tables 4.23 - 4.26 show results of the statistical tests of coefficients.

The last phase, aimed at validating the effectiveness of the proposed methodology, consists in comparing simulation results obtained by using the whole set of the surveyed data (considered as the *absolute truth*) with those using the data of calibrated space-time surfaces, properly put together with the data of calibration subsets. Specifically, within this framework, three different data sets may be obtained: only the calibration subset, the calibration subset extended by replacing missing data with function data and only function data for all values. Therefore, it is possible to implement an aggregate estimation of travel demand, according to the four data sets identified. In particular, the prior-known information used in the aggregate estimation is represented by the travel demand depicted in figure 4.9 and implemented in applications related to the Line 1 metro system. It is worth noting that, since the original information concerns the winter average working day, by means of the updating procedure, travel demand becomes representative of the summer period. In this way, four different O-D matrices can be derived and assigned to the network. The test framework in relation to which objective function values have been calculated, for each one of the four analysed data sets, is represented by the rescheduling assessment provided in paragraph 4.2, which considers 20 different recovery strategies. Hence, it is possible to compare assignment results (i.e. user generalised cost for each intervention action) obtained by implementing, on one hand, travel demand adjusted with the whole set and, on the other, travel demand adjusted with the other three data sets, in order to evaluate which one of them produces an outcome closer to that of the reference scenario.

The output of the analysis is reported in table 4.27 and shows that the data set which provides the smallest deviation, with respect to the results obtained with the whole set, is the one that integrates surveyed data with data from analytical relations. In particular, in this specific context, the derived space-time functions allow a 50% reduction in the amount of data to be acquired, with an accuracy reduction of less than 6% (see red value in table 4.27).

Clearly, this shows a trend; nevertheless, strictly speaking, a sensitivity analysis of the provided output with respect to the initial sampling rate (here set equal to 100%), as well as further applications to different metro contexts, are required. However, the preliminary outcome is promising and confirms the possibility, by means of suitable analytical relations, of cutting the budget to be allocated for the survey phase and investigating also networks in which the achievement of a reasonable sampling rate would result uneconomic, due to their complexity.

Intervention strategy	Partial surveyed set	Replaced missing data	Function data
0	9.73%	9.96%	26.97%
1	9.51%	0.84%	6.63%
2	9.51%	0.84%	6.63%
3	9.55%	0.97%	6.54%
4	9.49%	0.83%	6.62%
5	9.49%	0.83%	6.62%
6	9.56%	0.99%	6.56%
7	10.40%	0.04%	5.82%
8	9.52%	0.87%	6.63%
9	9.51%	0.99%	6.68%
10	9.55%	0.96%	6.54%
11	9.54%	10.21%	27.08%
12	9.54%	10.21%	27.08%
13	9.58%	10.12%	27.03%
14	9.53%	10.23%	27.09%
15	9.53%	10.23%	27.09%
16	9.60%	10.12%	27.03%
17	9.53%	10.18%	27.05%
18	9.55%	10.16%	27.04%
19	9.55%	10.18%	27.06%
20	9.59%	10.12%	27.03%
Average	9.59%	**5.71%**	17.28%
Median	10.40%	10.23%	27.09%
Minimum	9.49%	0.04%	5.82%
Maximum	9.54%	9.96%	26.97%

Table 4.27 Objective function accuracy for each different calibration set

4.5.2 A cost-benefit analysis relative to the implementation of an innovative signalling system in a regional context

Given the passenger-oriented perspective adopted in the presented work, in this section, some improving measures aimed at reducing user generalised costs are evaluated within a cost-benefit framework.

Obviously, each specific rate of passenger generalised costs is affected by certain features of rail systems: location of stops and stations affects access and egress times; the headway between two successive convoys, allowed by the travel speed and the adopted signalling system, affects waiting times; rolling stock performance and infrastructure characteristics affect travel times; layout of stations, platforms and rolling stock affects transfer times; pricing policies affect ticket costs. In particular, it is possible to identify the following intervention categories: infrastructural projects, fleet improvements, signalling system modifications, fare policies. Clearly, financing infrastructural measures require a considerable amount of resources, as well as the availability of large areas to be exploited. Therefore, such interventions frequently are not feasible, especially in high-density contexts; nevertheless, in certain cases, they could become imperative. Similarly, fleet modifications and new fare policies imply the need of additional national or regional subsidies, rarely available in the current economic conditions. The proposed application, instead, is focused on the implementation of an innovative signalling system on the Naples-Sorrento regional line.

Signalling systems are based on two paradigms: the spacing between two successive convoys and the train integrity supervision. The first one consists in imposing a minimum distance between two successive trains so that, in the case of the first train slowing or stopping, the following one is able to react safely; while, the second requirement consists in verifying the completeness of a train while it is operating. The choice of evaluating this kind of intervention is due to a twofold reason. Firstly, it allows to reduce passenger waiting times, since it provides an increase in service frequency; moreover, such measures are increasingly required in European countries, with the aim of meeting requirements dictated by the system of standards for management and interoperation of railways developed by the European Union, i.e. *European Rail Traffic Management System* (*ERTMS*). Specifically, the signalling, control and train protection criteria are provided by the so called *European Train Control System (ETCS),* which can be implemented according to three different levels: the higher the implementation level, the higher the network performance in terms of maximum speed and minimum headway between two successive convoys.

Level 1 is a cab signalling system which can be integrated with the existing signalling system, leaving fixed signals in place. In this case, movement authorities, as well as route data, are transmitted to the convoy in a discontinuous manner, i.e. when it travels over the Eurobalise beacons. Indeed, besides providing route data, Eurobalises are able to pick up signal aspects from the trackside signals, by means of the so called Lineside Electronic Unit (LEU), and transmit them to the vehicle. Level 2 is a digital radio-based system. Indeed, fixed signals are completely removed and there is a radio block center which continuously exchanges information with the train by means of the *GSM-R* technology. However, train detection and the train integrity supervision still remain in place at the trackside. In level 3, instead, also trackside equipment disappears; hence, the train integrity supervision is handed to on-board devices. Moreover, in this case, train spacing is no more based on a physical space (i.e. block sections), but it is dictated by the current operational conditions (i.e. moving block). More in detail, the radio block center is able to detect, continuously, the distance between two successive convoys, by verifying that it is, at least, equal to the braking distance. In this way it is possible to maximise the degree of capacity infrastructure utilisation and, therefore, reach very low headways, with an increase in service quality. Strictly speaking, also the ETCS level 0 has been defined. It indicates the condition in which, although rolling stock is equipped with ETCS, the infrastructure does not comply with European standards.

In terms of real applications, only Level 2 has been applied in actual railways, since the issue of on-board train integrity verification is still under research and development. Indeed, the signalling system analysed in this application can be defined as an ETCS level 3 in which, specifically, the on-board train integrity supervision is managed by a satellite technology. However, no technical details concerning this system will be provided, since the goal is to investigate its effects on rail service performance and, therefore, in terms of passenger satisfaction. A signalling system which allows a lower spacing between two successive convoys, generally, can provide two main benefits: a reduction in headways, which implies lower user waiting times, and an increase in travel speed which implies lower user travel times. However, given the infrastructure layout of the analysed regional line, characterised by stations at a very close distance (i.e. about 1.2 km), this second benefit essentially fails. Therefore, in the specific investigated context, the main contribution of the signalling improvement regards the decrease in waiting time which, for travellers, represents the most onerous rate among times and costs to be incurred for making a trip.

More in detail, the proposed methodology aims to evaluate economic and environmental effects related to a replacing intervention of signalling system, by performing a cost-benefit analysis based on a feasibility threshold approach. In this context, a key factor to be considered is represented by the involved passenger flows, in current and future conditions. For this purpose, as already stated, travel demand has to be elastic at least at the level of modal choice (in the case of transportation system modifications) and trip generation (in the case of demographic changes). In order to satisfy the above mentioned requirements, the procedure described in paragraph 3.3.4.2 can be applied, whose phases, together with the adopted Italian data sources, are synthetically set out below again, for the sake of simplicity:

- estimation of systematic trips by means of data from the national census;
- estimation of non-systematic trips by means of data from mobility observatories;
- evaluation of travel demand variation among different time periods by means of historical data from the resident population;
- development of a regional network model for transforming

O-D matrices defined in terms of municipalities into O-D matrices defined in terms of rail stations;

- update of the initial O-D matrices by means of turnstile counts, in order to reproduce surveyed flows;
- definition of travel demand in future scenarios by means of historical and/or forecasted data from the resident population (i.e. elasticity with respect to trip generation);
- specification, calibration and validation of a suitable modal choice model for providing an elastic travel demand model with respect to performance variations in the analysed transportation system;
- computation of the hourly O-D matrices to be assigned to the network;
- calculation of performance indexes.

Scenario	Description
1	Current infrastructure; current signalling system; current timetable.
2	Current infrastructure; current signalling system; current timetable for overlapping lines; maximising frequency for Naples–Sorrento line.
3	Current infrastructure; current signalling system; maximising frequency for Naples–Sorrento line, considering it a priority over other overlapping lines.
4	Current signalling system; doubling of Moregine–Sorrento section; current timetable for overlapping lines; maximising frequency for Naples–Sorrento line.
5	Current signalling system; doubling of Moregine–Sorrento section; maximising frequency for Naples–Sorrento line, considering it a priority over other overlapping lines.
6	Doubling of Moregine–Sorrento section; innovative signalling system which allows a 4 minute headway to be achieved between two successive rail convoys; maximising frequency for Naples–Sorrento line, considering it a priority over other overlapping lines.
7	Doubling of Moregine–Sorrento section; innovative signalling system which allows a 3 minute headway to be achieved between two successive rail convoys; maximising frequency for Naples–Sorrento line, considering it a priority over other overlapping lines.
8	Doubling of Moregine–Sorrento section; innovative signalling system which allows a 2 minute headway to be achieved between two successive rail convoys; maximising frequency for Naples–Sorrento line, considering it a priority over other overlapping lines.

Table 4.28 Analysed scenarios

Simulation scenarios compared within the cost-benefit framework are described in table 4.28. In particular, along with the current conditions (modelled in Scenario 1), other seven options of increasing complexity in terms of technological and monetary effort are evaluated. Moreover, since the last part of the line is characterised by a single-track layout, the feasibility of an infrastructural intervention, consisting in the doubling of the line between Moregine and Sorrento, is also investigated, so as to maximise the effects provided by the new signalling system.

Therefore, by implementing the basic simulation framework proposed in this work (with the exception of the *Failure model*), the objective function provided by equation 3.84 can be computed for each analysed scenario.

The simulation outcome in terms of objective function values in the analysed time period, detailed for minimum, average and maximum levels of demographic variation, is provided by tables 4.29 - 4.33.

In particular, parameters β_{UGC}, β_{NOC} and β_{EC} have been set equal to 1.

Scenario	Objective Function Value		
	Minimum	Average	Maximum
1	21,612,206	26,001,375	30,390,544
2	21,562,321	25,966,660	30,371,000
3	21,556,056	25,962,951	30,369,845
4	21,350,117	25,822,162	30,294,208
5	21,181,881	25,694,677	30,207,474
6	21,005,063	25,557,593	30,110,123
7	20,363,535	25,030,696	29,697,858
8	18,136,387	22,935,558	27,734,729

Table 4.29 Objective function values – year 2016

Scenario	Objective Function Value		
	Minimum	Average	Maximum
1	20,902,096	25,470,409	30,150,743
2	20,857,177	25,438,469	30,132,151
3	20,851,346	25,434,976	30,131,060
4	20,659,078	25,301,345	30,057,672
5	20,499,187	25,178,209	29,972,295
6	20,330,619	25,045,435	29,876,295
7	19,716,293	24,533,087	29,468,742
8	17,574,679	22,488,373	27,524,045

Table 4.30 Objective function values – year 2026

Scenario	Objective Function Value		
	Minimum	Average	Maximum
1	19,878,911	24,640,361	29,653,055
2	19,841,150	24,612,758	29,636,438
3	19,835,943	24,609,605	29,635,479
4	19,663,374	24,487,163	29,566,759
5	19,515,507	24,370,825	29,484,199
6	19,358,826	24,244,789	29,391,002
7	18,783,695	23,755,185	28,993,229
8	16,765,324	21,789,297	27,086,785

Table 4.31 Objective function values– year 2036

Scenario	Objective Function Value		
	Minimum	Average	Maximum
1	18,535,015	23,484,308	28,845,225
2	18,506,654	23,462,747	28,831,814
3	18,502,268	23,460,065	28,831,068
4	18,355,572	23,353,207	28,769,925
5	18,223,498	23,246,337	28,691,938
6	18,082,430	23,129,687	28,603,290
7	17,558,779	22,671,758	28,221,393
8	15,702,281	20,815,657	26,377,041

Table 4.32 Objective function values– year 2046

Scenario	Objective Function Value		
	Minimum	Average	Maximum
1	16,869,305	21,984,231	27,695,675
2	16,852,596	21,970,508	27,686,826
3	16,849,226	21,968,439	27,686,385
4	16,734,599	21,881,802	27,636,024
5	16,622,100	21,787,218	27,564,543
6	16,500,384	21,682,746	27,482,370
7	16,040,541	21,265,919	27,123,063
8	14,384,679	19,552,276	25,367,067

Table 4.33 Objective function values– year 2056

Furthermore, variations in the objective function value with respect to the non-intervention scenario (i.e. Scenario 1) are shown in tables 4.34 - 4.38.

Scenario	Objective Function Variation		
	Minimum	Average	Maximum
1	–	–	–
2	-0.06%	-0.14%	-0.23%
3	-0.07%	-0.16%	-0.26%
4	-0.32%	-0.74%	-1.21%
5	-0.60%	-1.26%	-1.99%
6	-0.92%	-1.81%	-2.81%
7	-2.28%	-3.93%	-5.78%
8	-8.74%	-12.20%	-16.08%

Table 4.34 Objective function variations – year 2016

Scenario	Objective Function Variation		
	Minimum	Average	Maximum
1	–	–	–
2	-0.06%	-0.13%	-0.21%
3	-0.07%	-0.15%	-0.24%
4	-0.31%	-0.71%	-1.16%
5	-0.59%	-1.22%	-1.93%
6	-0.91%	-1.77%	-2.73%
7	-2.26%	-3.87%	-5.67%
8	-8.71%	-12.11%	-15.92%

Table 4.35 Objective function variations – year 2026

Scenario	Objective Function Variation		
	Minimum	Average	Maximum
1	–	–	–
2	-0.06%	-0.12%	-0.19%
3	-0.06%	-0.13%	-0.22%
4	-0.29%	-0.67%	-1.08%
5	-0.57%	-1.16%	-1.83%
6	-0.88%	-1.70%	-2.62%
7	-2.23%	-3.78%	-5.51%
8	-8.65%	-11.96%	-15.66%

Table 4.36 Objective function variations – year 2036

Scenario	Objective Function Variation		
	Minimum	Average	Maximum
1	–	–	–
2	-0.05%	-0.10%	-0.15%
3	-0.05%	-0.11%	-0.18%
4	-0.26%	-0.60%	-0.97%
5	-0.53%	-1.08%	-1.68%
6	-0.84%	-1.60%	-2.44%
7	-2.16%	-3.63%	-5.27%
8	-8.56%	-11.73%	-15.28%

Table 4.37 Objective function variations – year 2046

Scenario	Objective Function Variation		
	Minimum	Average	Maximum
1	–	–	–
2	-0.03%	-0.06%	-0.10%
3	-0.03%	-0.07%	-0.12%
4	-0.22%	-0.49%	-0.80%
5	-0.47%	-0.95%	-1.47%
6	-0.77%	-1.44%	-2.19%
7	-2.07%	-3.42%	-4.91%
8	-8.41%	-11.40%	-14.73%

Table 4.38 Objective function variations – year 2056

Numerical results lead to a common conclusion: the doubling of the line is imperative in order to fully exploit the benefits provided by the innovative signalling system. As already shown, generally, the main drawbacks of infrastructure interventions are the necessity of finding major funding and large areas to utilise. In particular, by means of a parameter estimation on the basis of line features (see Cascetta and Pagliara, 2015), it can be stated that the doubling of the line has a cost, approximately, of 800,000,000 €. However, results indicate that such an investment could be recouped in only one year. Indeed, tables show a difference between scenario 1 and scenario 8 of about 3,000,000 €/day, which implies a regain of approximately 900,000,000 €/year; a value even higher than the estimated cost for the infrastructural intervention.

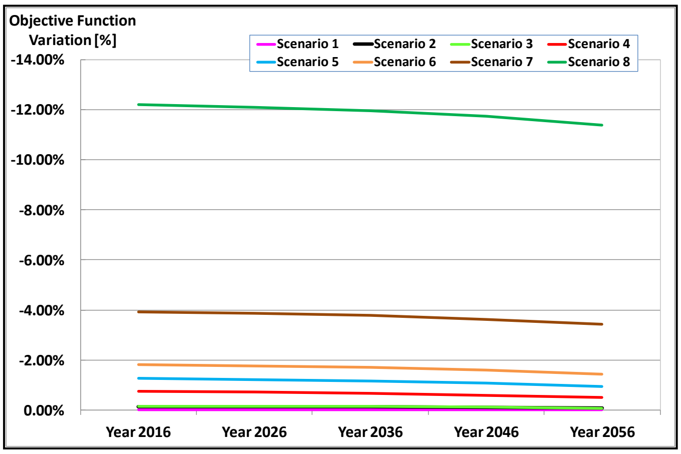


Figure 4.33 Variation of objective function value in average conditions during the analysed time period (2016-2056)

Figure 4.33 depicts the trend of objective function value variations, during the test period, in the case of an average rate of demographic variation. The graph confirms that, without the infrastructural upgrade, fully exploiting advantages from the signalling system improvements would be unfeasible.

The effects of each analysed scenario, in terms of headway between two successive trains, are illustrated in figure 4.34. In particular, as can be seen, by means of the timetable optimisation, the headway can move from 29 to 12 minutes, with a reduction of more than 50%; while, by doubling the line, it is possible to regain only about 7 minutes. However, this infrastructural intervention is essential for reducing headways between two successive convoys to as low as 2 minutes.

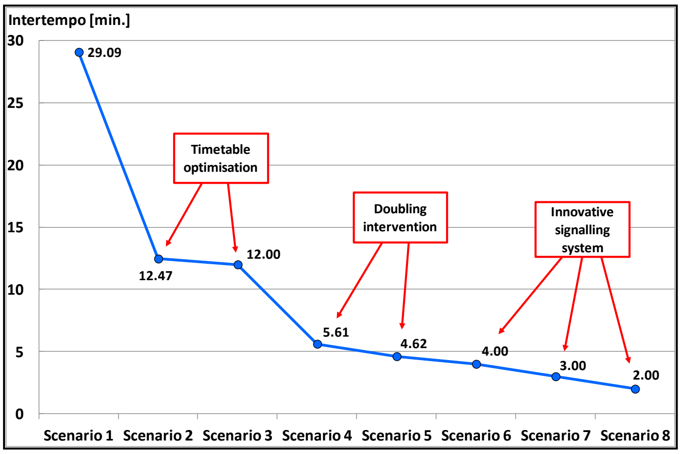


Figure 4.34 Simulation results in terms of headway for each scenario analysed

Scenario	Convoys required	Additional convoys
1	10	0
2	19	9
3	20	10
4	40	30
5	49	39
6	56	46
7	74	64
8	110	100

Table 4.39 Number of convoys required

To be able to guarantee such low headways, the availability of a certain number of convoys is required. Clearly, the lower the headway, the higher the number of trains to be operated (see table 4.39). Therefore, in a cost-benefit perspective, also resources for acquiring additional convoys have to be considered.

Finally, strictly speaking, also installation and maintenance costs related to the new signalling system would have to be taken into account. However, since in the proposed application the implemented signalling system is characterised from an operational perspective, rather than as an assembly of physical devices with specific technological features, this aspect has been neglected.

In conclusion, for any transportation system, the evaluation of the technological feasibility of alternative projects, involving long realisation times, requires the implementation of a cost-benefit approach where the long-term estimation of travel demand is a major requirement; therefore, the proposed framework can represent an effective support tool for such an analysis. In particular, the described method is based on the use of Italian data sources and, therefore, as research prospects, it would be appropriate to apply it in other contexts, such as other Italian railways (in order to verify the reliability of the adopted data sets in different network configurations) as well as other non-Italian railways (in order to test the methodology in the case of different data sources).

CHAPTER 5: CONCLUSION

5.1 Resume of the main achievements

The presented work provides a decision support system for planning and management phases of rail operations in a passenger-oriented perspective. The developed methodology consists in a simulation-optimisation integrated approach and, therefore, suitable optimisation and simulation techniques have been put in place to accurately model rail service, its interaction with travel demand and the related energy consumption issues.

Specifically, a bi-level multi-dimensional constraint optimisation problem is implemented, where it is necessary to minimise an objective function which expresses the user generalised cost and, if properly enhanced, the trade-off between passengers satisfaction and mass-transit agencies costs.

A basic and an extended structure can be distinguished, instead, in the proposed simulation framework. In particular, its backbone is provided by four models (i.e. Service Simulation Model, Travel Demand Model, Supply Model, and Failure Model) whose interactions allow to reproduce, both in ordinary and perturbed situations, the reciprocal influence between supply and demand features, which typically characterises any kind of evaluation concerning transport systems. This basic architecture is, then, improved by means of specific methodological frameworks which allow to:

- take into account the stochastic nature of involved operational factors, such as train performance and delays;
- perform a sensitivity analysis on the degree of reliability offered by the solution obtained by means of a deterministic approach;
- compute, analytically, the operational times within the timetable (e.g. inversion times, reserve times, layover times etc.) so as to be able to put in place energy saving strategies by preserving the service quality offered;

- model the snowball effect generated by the dynamic interaction between rail service and passenger flows so as to perform a reliable estimation of dwell times, thus, adequately, supporting the scheduling phase with the aim of carrying out a stable and robust timetable;
- simulate the effects, on rail service and user satisfaction, of different passenger behavioural patterns during the boarding process;
- extend passenger counts by means of properly calibrated functions which are able to capture the space-time variations of travel demand, thus permitting a reduction in surveyed data flow to be acquired, without prejudging the accuracy of the provided results;
- properly estimate long-term travel demand for supporting cost-benefit evaluations aimed at investigating the feasibility of design solutions on rail systems.

Each one of these specific methodological frameworks has been tested in a real network context, in order to evaluate its effectiveness and suitability. In particular, most of the proposed applications concerned metro systems which, generally, operate in overcrowded conditions and, therefore, imply even a greater necessity of properly taking into account the above cited issues. However, in order to show the potential of the proposed methodology, a rescheduling framework and forecasting techniques for long-term travel demand estimation have been implemented in the case of a regional rail line. Specifically, the differences between the analysed contexts, which have mostly affected the application of the presented approach, are a different spatial characterisation of travel demand and a different timetable structure, which give rise to the necessity of a non-equivalent modelling of passenger behavioural choices.

Numerical results appear promising and confirm the relevance of reconsidering dispatching and rescheduling tasks in a passenger-oriented perspective, as well as the inadequacy of evaluating rail service as astand-alone system, without considering related energy consumption issues and space-time variability of the involved user flows, in order to perform an accurate analysis.

The transfer of the proposed methodology from a research sphere to a practical use has been conceived as consisting in the generation of a dynamic database which, once at dispatchers' disposal, provides them with a support tool for managing rail operations, both in the planning and the management phase. The latter includes both ordinary and disruption operating conditions. In particular, the information content of such a database consists in the identification and quantification, for each possible intervention strategy, related or not to a specific failure event, of relevant impacts on each part of the analysed system. The targets considered in this work are related to user generalised costs as well as operational costs and energy consumption; however, it is understood that further contents can be made available by properly enriching the simulation architecture with suitable modelling frameworks. In this way, dispatchers might be fully conscious of the implications of each possible intervention and have all information to react properly to any contingency, with response times comparable with real-time rescheduling approaches, but without the computational effort they require.

5.2 Research prospects

The main aspect to be addressed, for improving the entire methodology and allowing an operational use of the described database, consists in the development of proper feature learning techniques. The goal is to confer a dynamic structure to such a tool, which enables it to progressively upgrade its information content on the basis of events occurring during the service, so as to minimise the probability that specific conditions to be faced are not included in the database yet. This, jointly with dispatchers' experience, as well as additional information provided by means of the sensitivity analysis which can be performed for each obtained solution, could lead to an efficient management of rail systems and, therefore, to an effective valorisation of such a transport mode with all related benefits.

However, in order to increase the accuracy of the provided outcome, also methodological frameworks, proposed for extending the basic simulation structure, can be individually improved.

Firstly, within rescheduling applications, it would be appropriate to test more articulated metaheuristic techniques so as to verify if a greater complexity actually implies an improvement in terms of how good performed solutions are.

On the other hand, as regards planning tasks, the simulation-based framework developed for estimating dwell times as flow-dependent factors could be enhanced by introducing rail crowding models, allowing to replicate conditions in which passengers could decide not to take the first arriving train, but wait for the following one (hoping it will be less crowded), although this would mean an increase in their waiting time, so as to make the simulation more realistic. Additionally, in this way, an assessment of en-route passenger discomfort could be performed, thus strengthening the passenger-oriented nature of the proposed approach. Dwell times can play a role also in the definition of supplement time rates to be used for implementing eco-driving strategies and, therefore, the two methodological approaches could be combined in an energy saving perspective. Further improvements consist in enabling the proposed framework to compute the statistical distribution associated with resulting dwell times, rather than only their average values, as well as testing other resolution procedures for solving the fixed-point problem generated by the interaction between rail service and travel demand.

Concerning the handling of passenger counts for the aggregate estimation of O-D matrix, as already shown, it calls for several improvements: the adoption of different spatial reference systems, the introduction of additional predictors and the implementation of conversion coefficients to be properly calibrated in order to capture eventual correlations existing among travel demand patterns in different time periods. Moreover, the promising results obtained by applying this methodology to a real metro line need to be further validated by analysing more complex metro contexts and adopting different sampling rates. The goal is to investigate how the layout of the considered line, as well as the ratio between the initial surveyed rate and the simulated one, have affected the quality of the provided outcome.

Finally, it is worth analysing the transferability of the proposed technique for long-term demand estimation by further testing its performance in other contexts and in the case of different data sources.

In the light of the above, it is clear that the discussed topics are still open to debate and present a considerable potential which is worth investigating in greater depth in forthcoming works.

DISSEMINATION OF THE MAIN RESEARCH ACHIEVEMENTS

In the course of these past three years of research activities, within my Ph.D., part of what is here presented has been published in the following contributions:

1. L. D'Acierno, **M. Botte** and B. Montella (2018). *Assumptions and simulation of passenger behaviour on rail platforms.* International Journal of Transport Development and Integration 2(2), pp. 123-135, ISSN: 2058-8305.

2. L. D'Acierno, **M. Botte**, A. Placido, C. Caropreso and B. Montella (2017). *Methodology for determining dwell times consistent with passenger flows in the case of metro services*. Urban Rail Transit 3(2), pp. 73-89, ISSN: 2199-6687.

3. L. D'Acierno, **M. Botte** and B. Montella (2017). *An analytical approach for determining reserve times on metro systems*. Proceedings of the 17th IEEE International Conference on Environment and Electrical Engineering (IEEE EEEIC 2017) and 1st Industrial and Commercial Power Systems Europe (I&CPS 2017), Milano, June 2017, art. no. 7977519, pp. 722-727, ISBN: 978-1-5386-3916-0.

4. **M. Botte**, D. Puca, B. Montella and L. D'Acierno (2017). *An innovative methodology for managing service disruptions on regional rail lines.* Proceedings of the 10th International Conference on Environmental Engineering, Vilnius Gediminas Technical University (VGTU) Press, Vilnius, Lithuania, April 2017.

5. R. Di Mauro, **M. Botte** and L. D'Acierno (2017). *An analytical methodology for extending passenger counts in a metro system*. International Journal of Transport Development and Integration 1(3), pp. 589-600, ISSN: 2858-8305.

6. C. Caropreso, C. Di Salvo, **M. Botte** and L. D'Acierno (2017). *A long-term analysis of passenger flows on a regional rail line.* International Journal of Transport Development and Integration 1(3), pp. 329-338, ISSN: 2858-8305.

7. **M. Botte**, C. Di Salvo, A. Placido, B. Montella and L. D'Acierno (2017). *A Neighbourhood Search Algorithm for determining optimal intervention strategies in the case of metro system failures*. International Journal of Transport Development and Integration 1(1), pp. 63-73, ISSN: 2058-8305.

8. L. D’Acierno, A. Placido, **M. Botte**, M. Gallo and B. Montella (2016). *Defining robust recovery solutions for preserving service quality during rail/metro systems failure*. International Journal of Supply and Operations Management 3(3), pp. 1351-1372, ISSN: 2383-1359.

9. L. D’Acierno, **M. Botte**, C. Di Salvo, C. Caropreso and B. Montella (2016). *A methodology for long-term analysis of innovative signalling systems on regional rail lines*. Transactions on Environment and Electrical Engineering 1(3), pp. 77-85, ISSN: 2450-5730.

10. **M. Botte**, C. Di Salvo, C. Caropreso, B. Montella and L. D’Acierno (2016). *Defining economic and environmental feasibility thresholds in the case of rail signalling systems based on satellite technology*. Proceedings of the 16th IEEE International Conference on Environment and Electrical Engineering (IEEE EEEIC 2016), Florence, June 2016, art. no. 7555878, pp. 251-255, ISBN: 978-1-5090-2319-6.

11. L. D’Acierno, A. Placido, **M. Botte** and B. Montella (2016). *A methodological approach for managing rail disruptions with different perspectives*. International Journal of Mathematical Models and Methods in Applied Sciences 10, pp. 80-86, ISSN: 1998-0140.

12. **M. Botte**, L. D’Acierno, B. Montella and A. Placido (2015). *A stochastic approach for assessing intervention strategies in the case of metro system failures*. Proceedings of the 2015 AEIT International Annual Conference, Naples, October 2015, art. no. 7415258, ISBN: 978-8-8872-3728-3.

13. L. D’Acierno, A. Placido, **M. Botte** and B. Montella (2015). *Preliminary results on different perspectives in managing rail disruptions*. Proceedings of the 6th International Conference on Automotive and Transportation Systems, Salerno, June 2015, pp. 32-37, ISSN: 2227-4588, ISBN: 978-1-61804-316-0.

14. A. Placido, L. D’Acierno, **M. Botte** and B. Montella (2015). *Effects of stochasticity on recovery solutions in the case of high-density rail/metro networks*. Proceedings of the 6th International Conference on Railway Operations Modelling and Analysis, Tokyo, Japan, March 2015.

15. A. Placido, L. D’Acierno, **M. Botte**, M. Gallo and B. Montella (2015). *A sensitivity analysis of recovery solutions in the case of rail disruption management*. Proceedings of the 94th Annual Meeting of the Transportation Research Board, Washington (D.C.), USA, January 2015.

REFERENCES

Acikbas, S., & Soylemez, M.T., (2008). Coasting point optimisation for mass rail transit lines using artificial neural networks and genetic algorithms. IET Electric Power Applications, 2(3), pp. 172-182.

Acuna-Agost, R., Michelon, P., Feillet, D., & Gueye, S. (2011). A MIP-based local search method for the railway rescheduling problem. Networks, 57 (1), pp. 69-86.

Adenso-Díaz, B.; Oliva González, M.; & González-Torre, P. (1999). On-line timetable re-scheduling in regional train services, Transportation Research Part B, 33(6), pp. 387-398.

Albrecht, T., & Oettich S., (2002). A new integrated approach to dynamic schedule synchronization and energy-saving train control. WIT Transactions on The Built Environment, 61, pp. 847-856.

Albrecht, T. (2009).Automated timetable design for demand-oriented service on suburban railways. Public Transport, 1, pp. 5-20.

Albrecht, A., Howlett, P., Pudney, P., & Vu, X. (2013). Energy-efficient train control: from local convexity to global optimization and uniqueness. Automatica, 49(10), pp. 3072-3078.

Antoniou, C., Azevedo, C.L., Lu, L., Pereira, F., & Ben-Akiva, M. (2015). W-SPSA in practice: Approximation of weight matrices and calibration of traffic simulation models. Transportation Research Part C, 59, pp. 129-146.

Appiah, J., & Rilett, L. (2010). Joint estimation of dynamic origin-destination matrices and calibration of microsimulation models using aggregate intersection turn count data, Proceedings of the 89th Annual Meeting of the Transportation Research Board, Washington DC, USA, January 2010.

Ashok, K., & Ben-Akiva, M. (1993). Dynamic origin-destination matrix estimation and prediction for real time-traffic management systems. Proceedings of the 12th International Symposium on the Theory of Traffic Flow and Transportation, Berkeley (CA), USA, July 1993, pp. 465-484.

Baee, S., Eshghi, F., Hashemi, SM., & Moienfar, R. (2012) Passenger boarding/alighting management in urban rail transportation. Proceedings of the Joint Rail Conference – JRC 2012, Philadelphia (PA), USA, April 2012, pp. 823-829.

Balakrishna, R., Ben-Akiva, M., & Koutsopoulos, H.N. (2008). Time-dependent origin-destination estimation without assignment matrices. In: Chung, E., Dumont, A.-G. (Eds.), Transport Simulation: Beyond Traditional Approaches. EPFL Press, pp. 201-213.

Banach, S. (1922) Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. Fundamenta Mathematicae, 3, pp. 133-181.

Bandini, S., Mondini, M., & Vizzari, G. (2014) Modelling negative interaction among pedestrians in high density situations. Transportation Research Part C, 40, pp. 251-270

Bao, X., Li, H., Qin, L., Xu, D., Ran, B., & Rong, J. (2016). Sensor location problem optimization for traffic network with different spatial distributions of traffic information. Sensors, 16(11), 1790.

Barcelò, J., & Montero, L. (2015). A robust framework for the estimation of dynamic OD trip matrices for reliable traffic management. Transportation Research Procedia, 10, pp. 134-144.

Barter, W. (2004). Forecasting robustness of timetables. WIT Transactions on The Built Environment, 74, pp. 563-572.

Bauer, R., & Schöbel, A. (2014). Rules of thumb: Practical online-strategies for delay management. Public Transport, 6(1), pp. 85-105.

Bazaraa, M. S., Sherali H.D., & Shetty C.M. (1993). Nonlinear programming: Theory and algorithms. Wiley (NY), USA.

Beckman, M., McGuire C.B., & Winsten C.B. (1956). Studies in the economics of transportation. Yale University Press, New Haven (CT), USA.

Bell, M.G.H. (1983). The estimation of origin-destination matrix from traffic counts. Transportation Science, 17(2), pp. 198-217.

Bellman, R. (1958). On a routing problem. Quarterly Applied Mathematics, 16, pp. 87-90.

Ben-Akiva, M., Cyna, M., & De Palma, A. (1984). Dynamic model of peak period congestion. Transportation Research Part B, 18, pp. 339-355.

Ben-Akiva, M., & Lerman S. R. (1985). Discrete choice analysis: theory and application to travel demand. The MIT Press, Cambridge (MA), USA.

Ben-Akiva, M., & Morikawa, T. (1990). Estimation of switching models from revealed preference and stated intention. Transportation Research Part A, 24(6), pp. 485-495.

Bera, S., & Rao, K.V.K. (2011). Estimation of origin-destination matrix from traffic counts: the state of the art. European Transport,49, pp. 3-23.

Berbey, A., Galan, R., Sanz Bobi, JD., & Caballero, R. (2012) A fuzzy logic approach to modelling the passengers' flow and dwelling time. WIT Transactions on The Built Environment, 128, pp. 359-369.

Bešinović, N., Roberti, R., Quaglietta , E., Cacchiani, V., Toth, P., & Goverde, R.M.P. (2015). Micro-macro approach to robust timetabling. Proceedings of the 6th International Conference on Railway Operations Modelling and Analysis – RailTokyo2015, Narashino, Japan, March 2015.

Bešinović, N., & Goverde, R.M.P.(2016). Improving robustness of railway timetables: A new two-stage approach. Proceedings of the 9th Triennial Symposium on Transportation Analysis – TRISTAN IX, June 2016, Aruba.

Bešinović, N., Goverde, R.M.P., Quaglietta, E., & Roberti, R. (2016). An integrated micro–macro approach to robust railway timetabling. Transportation Research Part B, pp. 87, 14-32.

Bierlaire, M., & Crittin, F. (2004). An efficient algorithm for real-time estimation and prediction of dynamic OD tables. Operations Research, 52(1), pp. 116-127.

Bifulco G.N. (1993). A stochastic user equilibrium assignment model for the evaluation of parking policies. European Journal of Operational Research, 71, pp. 269-287.

Binder, S., Maknoon, Y., & Bierlaire, M. (2015). Passenger-oriented railway disposition timetables in case of severe disruptions. Proceedings of the 15th Swiss Transport Research Conference – STRC 2015, Ascona, Swizterland, April 2015.

Binder, S., Maknoon, Y., & Bierlaire, M. (2017). The multi-objective railway timetable rescheduling problem. Transportation Research Part C, 78, pp.78-94.

Bizhan, M., Mohammad, K., (2015). Real-time rescheduling for a subway network using model predictive controller by considering actual constraints. Transportation Engineering, 7(1), pp. 141-166.

Blum, JR. (1954) Multidimensional stochastic approximation methods. Annals of Mathematical Statistics, 25, pp. 737-744.

Boccia, M., Mannino, C., & Vasilyev, I., (2013). The dispatching problem on multitrack territories: heuristic approaches based on mixed integer linear programming. Networks, 62 (4), 315-326.

Bocharnikov, Y. V., Tobias, A. M., Roberts, C., Hillmansen, S., & Goodman, C. J. (2007). Optimal driving strategy for traction energy saving on DC suburban railways. IET Electric Power Applications, 1(5), pp. 675-682.

Botte, M., Puca, D., Montella, B., & D'Acierno, L. (2017) An innovative methodology for managing service disruptions on regional rail lines. Proceedings of the 10th International Conference Environmental Engineering – ENVIRO 2017, Vilnius Gediminas Technical University (VGTU) Press, Vilnius, Lithuania, April 2017.

Bouzaiene-Ayari, B., Gendrau, M., & Nguyen, S. (1998). Passenger assignment in congested transit networks: a historical perspective. In P. Marcotte and S. Nguyen (eds.) 'Equilibrium and advanced transportation modelling', Kluwer, pp. 47-71.

Braker, J.G. (1993). Algorithms and applications in timed discrete event systems. Ph.D. dissertation, Delft University of Technology, The Netherlands.

Brännlund, U., Lindberg, P. O., Nõu, A., & Nilsson, J. E. (1998). Railway timetabling using lagrangian relaxation. Transportation Science, 32, pp. 358-369.

Brog, W., & Ampt, E. State of the art in the collection of travel behaviour data. Travel Behaviour for the 1980's. Special Report 201, National Research Council, Washington DC, USA, 1982.

Brouwer, L.E.J. (1912). Uber abbildung der mannigfaltigkeiten. Mathematische Annalen, 71, pp. 97-115

Bruglieri, M., Maja, R., & Tolentino, S., (2017). Optimizing regular symmetric timetables: a method to reach the best modal split for railway. Optimisation online. http://www.optimization-online.org/DB_HTML/2017/05/5991.html.

Buchmuller, S., Weidmann, U., & Nash, A., (2008) Development of a dwell time calculation model for timetable planning. WIT Transactions on The Built Environment, 103, pp. 525-534.

Büker, T., & Seybold, B., (2012) Stochastic modelling of delay propagation in large networks. Rail Transport Planning & Management, 2, pp. 34-50.

Burdett, RL., & Kozan, E. (2014) Performance profiling for predictive train schedules. Rail Transport Planning & Management, 4, pp. 98-114.

Burkolter, D.M., (2005). Capacity of railways in station areas using petri nets. Ph.D. dissertation, Swiss Federal Institute of Technology, Zurich, Switzerland.

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., & Toth, P. (2012). Nominal and robust train timetabling problems. European Journal of Operational Research, 219 (3), pp. 727-737.

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, 63, pp. 15-37.

Cadarso, L., Marín, Á., & Maróti, G. (2013). Recovery of disruptions in rapid transit networks. Transportation Research Part E, 53, pp.15-33.

Cadarso, L., Maróti, G., & Marín, A., (2015). Smooth and controlled recovery planning of disruptions in rapid transit networks. IEEE Transactions On Intelligent Transportation Systems, 16(4), pp. 2192-2202.

Caggiani, L., Dell'Orco, M., Marinelli, M., & Ottomanelli, M. (2012). A metaheuristic dynamic traffic assignment model for O-D matrix estimation using aggregate data. Procedia – Social and Behavioral Sciences, 54, pp. 685-695.

Canca, D., Zarzo, A., Algaba, E., & Barrena, E. (2011). Confrontation of different objectives in the determination of train scheduling, Procedia – Social and Behavioral Sciences, 20, pp. 302-312.

Canca, D., Barrena, E., Zarzo, A., Ortega, F., & Algaba, E. (2012). Optimal train reallocation strategies under service disruptions". Procedia – Social and Behavioral Sciences, 54, pp. 402-413.

Canca, D. (2017). Analysis of the energy-efficient design of railway rapid transit timetables. Proceedings of workshop on mathematical models of optimization for transportation planning, Seville, Spain, March 2017.

Canca., D., & Zarzo, A. (2017). Design of energy-efficient timetables in two-way railway rapid transit lines. Transportation Research Part B, 102, pp. 142-161.

Canca, D., De-Los-Santos, A., Mesa J.A., & Laporte, G. (2017). The railway network design, line planning and capacity problem: an adaptive large neighbourhood search metaheuristic. In book: Advanced Concepts, Methodologies And Technologies For Transportation And Logistics., Edition: 1th, Publisher: Advances in Intelligent Systems and Computing. Springer.

Cantarella, G.E. (1997). A general fixed-point approach to multi-mode multi-user equilibrium assignment with elastic demand. Transportation Science, 31(2), pp. 107-128.

Cantarella, G.E., & Cascetta E. (1995). Dynamic processes and equilibrium in transportation networks: towards a unifying theory. Transportation Science, 29(4), pp. 305-329.

Cantarella, G.E., Cartenì, A., & De Luca, S. (2015). Stochastic equilibrium assignment with variable demand: theoretical and implementation issues. European Journal of Operational Research, 241(2), pp. 330-347.

Carey, M., & Kwiecinski, A. (1995). Properties of expected costs and performance measures in stochastic models of scheduled transport. European Journal of Operational Research, 83, pp. 182-199.

Carey M., & Carville, S. (2000). Testing schedule performance and reliability for train stations. Journal of the Operational Research Society, 51, pp. 666-682.

Carreno, W.C. (2017). Efficient driving of CBTC ATO operated trains. Ph.D. Dissertation, Universidad Pontificia Comillas, Madrid, Spain.

Cascetta, E. (1984). Estimation of trip matrices from traffic counts and survey data: a generalized least squares estimator. Transportation Research Part B, 18(4-5), pp. 289-299.

Cascetta, E., & Nguyen, S. (1988). A unified framework for estimating or updating origin/destination matrices from traffic counts. Transportation Research Part B, 22(6), pp.437-455.

Cascetta, E., Inaudi, D., & Marquis, G. (1993). Dynamic estimators of origin-destination matrices using traffic counts. Transportation Science, 27(4), pp. 363-373.

Cascetta, E., Nuzzolo, A., Russo, F., & Vitetta, A. (1996). A modified logit route choice model overcoming path overlapping problems, specification and some calibration results for interurban networks. Proceedings of the 13th International Symposium on Transportation and Traffic Theory, Lyon, France, July 1996, pp. 697-711.

Cascetta, E., & Russo, F. (1997). Calibrating aggregate travel demand models with traffic counts: Estimators and statistical performance. Transportation, 24, pp. 271-293.

Cascetta, E., & Postorino, M. N. (2001). Fixed point models for the estimation of O-D matrices using traffic counts on congested networks. Transportation Science, 35(2), pp. 134-147.

Cascetta, E. (2009). Transportation systems analysis: models and applications, Springer, New York (NY), USA.

Cascetta, E., Papola, A., Marzano, V., Simonelli, F. & Vitiello, I. (2013). Quasi-dynamic estimation of o–d flows from traffic counts: Formulation, statistical validation and performance analysis on real data. Transportation Research Part B, 55, pp. 171-187.

Cascetta, E. & Pagliara, F. (2015). Le infrastrutture di trasporto in Italia (in Italian). Aracne, Roma, Italy.

Ceder, A. A., & Hassold, S. (2015). Applied analysis for improving rail-network operations. Rail Transport Planning & Management, 5, pp. 50-63.

Cenelec (1999). Railway applications - specification and demonstration of reliability, availability, maintainability and safety (RAMS). EN50126.

Cepeda, M., Cominetti, R., & Florian M. (2006). A frequency-based assignment model for congested transit networks with strict capacity constraints: characterization and computation of equilibria. Transportation Reserach Part B, 40, pp. 437-459.

Chang, SC., & Chung, YC. (2005). From timetabling to train regulation – a new train operation model. Information and Software Technology, 47, pp. 575-585.

Chevrier, R., Pellegrini, P., & Rodriguez, J. (2013). Energy saving in railway timetabling: A bi-objective evolutionary approach for computing alternative running times. Transportation Research Part C, 37, pp. 20-41.

Chu, W-J., Zhang, X-C., Chen, J-H., & Xu, B. (2015) An ELM-Based approach for estimating train dwell time in urban rail traffic. Mathematical Problems in Engineering, 2015, article ID 473432.

Chuang, HJ., Chen, CS., Lin, CH., Hsieh, CH., & Ho, CY. (2008). Design of optimal coasting speed for saving social cost in mass rapid transit systems. Proceedings of 3rd International conference on deregulation and restructuring and power technologies – DRPT 2008, Nianjing, China, April 2008, pp. 2833-2839.

Chung, I.H. (2001). An optimum sampling framework for estimating trip matrices from day-to-day traffic counts. Ph.D. dissertation, University of Leeds, UK.

Cipriani, E., Florian, M., Mahut, M., & Nigro, M. (2011). A gradient approximation approach for adjusting temporal origin–destination matrices, Transportation Research part C, 19, pp. 270-282.

Cipriani, E., Nigro, M., Fusco, G., & Colombaroni, C. (2014). Effectiveness of link and path information on simultaneous adjustment of dynamic O-D demand matrix, European Transport Research Review, 6(2), pp 139-148.

Ciuffini, F. (2014). A method for evaluating the frequency effect on the overall travel time. Ingegneria Ferroviaria, 69(10), pp. 803-821.

Colorni, A., Dorigo, M., & Maniezzo, V. (1991). Distributed optimization by ant colonies. Proceedings of the First European Conference on Artificial Life – ECAL 1991, Paris, France, December 1991, pp. 134-142.

Colorni, A., Dorigo, M., Maniezzo, V.:(1992) An investigation of some properties of an ant algorithm. Proceedings of the Parallel Problem Solving from Nature – PPSN 1992, Brussels, Belgium, September 1992, pp. 509-520.

Conte, C., & Shöbel, A. (2007). Identifying dependencies among delays. Proceedings of the 2nd International Seminar on Railway Operations Modelling and Analysis – RailHannover 2007, Hannover, Germany, March 2007.

Corapi, G., De Martinis, V., Placido, A., & De Luca, G., (2014). Impacts of energy saving strategies (ESSs) on rail services and related effects on travel demand. WIT Transactions on The Built Environment, 135, pp. 709-720.

Corman, F., D'Ariano, A., Pacciarelli, D. & Pranzo, M. (2009). Evaluation of green wave policy in real-time railway traffic management. Transportation Research Part C, 17(6), pp. 607-616.

Corman, F., D'Ariano, A., & Hansen, I.A. (2010a). Disruption handling in large railway networks. WIT Transactions on The Built Environment, 114, pp. 629-640.

Corman F., D'Ariano A., Pacciarelli D. & Pranzo, M. (2010b). A tabu search algorithm for rerouting trains during rail operations. Transportation Research Part B, 44(1), pp. 175-192.

Corman, F., D'Ariano, A., Hansen, I.A., Pacciarelli, D., & Pranzo, M. (2011a). Dispatching trains during seriously disrupted traffic situations. Proceedings of the IEEE International Conference on Networking, Sensing and Control – IEEE ICNSC 2011, Delft, The Netherlands, April 2011, pp. 323-328.

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), pp. 341-362.

Corman, F., D'Ariano, A. Hansen, I. A., & Pacciarelli, D. (2011c). Optimal multi-class rescheduling of railway traffic. Rail Transport Planning & Management, 1, pp. 14-24.

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), pp. 79-94.

Corman, F., & Meng, L. (2015). A review of online dynamic models and algorithms for railway traffic management. IEEE Transactions On Intelligent Transportation Systems, 16(3), pp. 1274-1284

Corman, F., Pacciarelli, D., D'Ariano, A., & Samà, M. (2015). Rescheduling railway traffic taking into account minimization of passengers' discomfort. Proceedings of International Conference on Computational Logistics – ICCL 2015, Delft, The Netherlands, September 2015, pp 602-616.

Cornic, D. (2010). Efficient recovery of braking energy through a reversible dc substation. Proceedings of Electrical Systems for Aircraft, Railway and Ship Propulsion – ESARS 2010, Bologna, Italy, October 2010.

Cremer, M., & Keller, H. (1981). Dynamic identification of O–D flows from traffic counts at complex intersections. Proceedings of the 8th International Symposium on Transportation and Traffic Theory – ISTTT 1981, Toronto, Canada, June 1981, pp. 121-142.

Cremer, M., & Keller, H. (1984). A systems dynamics approach to the estimation of entry and exit O–D flows. Proceedings of the 9th International Symposium on Transportation and Traffic Theory – ISTTT 1984, Delft, The Netherlands, June 1984, pp. 431-450.

Cremer, M., & Keller, H. (1987). A new class of dynamic methods for the identification of origin–destination flows. Transportation Research Part B, 21, pp. 117-132.

Cucala, A.P., Fernández, A.,Sicre, C., & Domínguez, M. (2012). Fuzzy optimal schedule of high speed train operation to minimize energy consumption with uncertain delays and driver's behavioral response. Engineering Applications of Artificial Intelligence, 25(8), pp. 1548-1557.

Cui, Y., Martin, U., & Zhao, W. (2016) Calibration of disturbance parameters in railway operational simulation based on reinforcement learning. Rail Transport Planning & Management, 6, pp. 1-12.

D'Acierno, L., Montella, B., & Gallo, M. (2002). Multimodal assignment to congested networks: fixed-point models and algorithms. Proceedings of the European Transport Conference – ETC 2002, Cambridge, United Kingdom, September 2002.

D'Acierno, L., Montella, B., & De Lucia, F. (2006). A stochastic traffic assignment algorithm based on Ant Colony Optimisation. In 'Ant Colony Optimization and Swarm Intelligence' (Editors: M. Dorigo, L.M. Gambardella, M. Bittari, A. Martinoli, R. Poli and T. Stützle), Lecture Notes in Computer Science, 4150, Springer-Verlag, Berlin, Germany, pp. 25-36.

D'Acierno, L., Cartenì, A., & Montella, B. (2009). Estimation of urban traffic conditions using an Automatic Vehicle Location(AVL) System. European Journal of Operational Research, 196, pp. 719-736.

D'Acierno, L., Montella, B., & Gallo, M. (2011). A fixed-point model and solution algorithms for simulating urban freight distribution in a multimodal context. Journal of Applied Sciences, 11(4), pp. 647-654.

D'Acierno, L., Gallo, M., Montella, B., & Placido, A. (2012). Analysis of the interaction between travel demand and rail capacity constraints. WIT Transaction of Built Environment, 128, pp. 197-207.

D'Acierno, L., Gallo, M., Montella, B., & Placido, A. (2013a). Evaluation of travel demand impacts in the case of rail system failure. Procedia – Social & Behavioral Sciences, 87, pp. 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 16th International IEEE Conference on Intelligent Transport Systems – IEEE ITSC 2013, The Hague, The Netherlands, October 2013, pp. 1059-1064.

D'Acierno, L., Gallo, M., & Montella, B. (2014). Application of metaheuristics to large-scale transportation problems. In 'Large-scale scientific computing' (Editors: I. Lirkov, S. Margenov and J. Wasniewski), Lecture Notes in Computer Science, 8353, Springer, Berlin, Germany, pp. 215-222.

D'Acierno, L., Placido, A., Botte, M., Gallo, M., & Montella, B. (2016a). Defining robust recovery solutions for preserving service quality during rail/metro systems failure. International Journal of Supply and Operations Management, 3(3), pp. 1351-1372.

D'Acierno, L., Placido, A., Botte, M. & Montella B., (2016b) A methodological approach for managing rail disruptions with different perspectives. International Journal of Mathematical Models and Methods in Applied Sciences, 10, pp. 80-86.

D'Acierno, L., Botte, M., & Montella, B. (2017) An analytical approach for determining reserve times on metro systems. Proceedings of the 17th IEEE International Conference on Environment and Electrical Engineering – IEEE EEEIC 2017 and 1st Industrial and Commercial Power Systems Europe - I&CPS 20), Milan , Italy, June 2017, pp. 722-727.

D'Ariano, A., & Albrecht, T. (2006). Running time re-optimization during real-time timetable perturbations. WIT Transactions on the Built Environment, 88, pp. 531-540.

D'Ariano, A., Pacciarelli, D., & Pranzo, M., (2007). A branch and bound algorithm for scheduling trains in a railway network. European Journal of Operational Research, 183(2), pp. 643-657.

D'Ariano, A. (2008). Improving real-time train dispatching: Models, algorithms and applications. Ph.D. dissertation, Delft University of Technology, The Netherlands.

D'Ariano, A., Corman, F., Pacciarelli, D., & Pranzo, M. (2008.). Reordering and local rerouting strategies to manage train traffic in real-time. Transportation Science 42(4), pp. 405-419.

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), pp. 63-84.

D'Ariano, A., Nucci, D., Pacciarelli, D., & Rosti, M. (2016). Intelligent real-time traffic management system for complex and busy railway networks. Proceedings of the 11th World Congress on Railway Research – WCRR 2016, Milan, Italy, January 2016.

D'Ariano, A., Pacciarelli, D., Samà, M., & Corman, F. (2017). Microscopic delay management: Minimizing train delays and passenger travel times during real-time railway traffic control. Proceedings of the 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems – IEEE MT-ITS 2017, Naples, Italy, June 2017,pp. 310-315.

Daamen, W., Lee, Y., & Wiggenraad, P. (2008). Boarding and alighting experiments: overview of setup and performance and some preliminary results. Transportation Research Record, 2042(1), pp. 71-81

Dafermos, S. C. (1980). Traffic equilibrium and variational inequalities. Transportation Science 14, pp. 42-54.

Daganzo, C. F., & Sheffi Y. (1977). On stochastic models of traffic assignment. Transportation Science, N. 11(3), pp. 253-274.

Daganzo, C. F. (1979). Multinomial probit: the theory and its application to demand forecasting. Academic press, New York (NY), USA.

Daganzo, C. F. (1983). Stochastic network equilibrium with multiple vehicle types and asymmetric, indefinite link cost jacobians. Transportation Science, 17, pp. 282-300.

Daly, A. J., & Zachary, S. (1978), Improved multiple choice models. in D.A. Hensher and M.Q. Dalvi (eds.), Determinants of Travel Choice. Saxon House, Westmead.

Davydov, B., Chebotarev, & V., Kablukova, K. (2017). Stochastic model for the real-time train rescheduling. International Transport Development and Integration, 1, pp. 307-317.

De Cuadra, F., Fernandez, A., de Juan, J., & Herrero, M. A., (1996). Energy-saving automatic optimisation of train speed commands using direct search techniques. WIT Transactions on The Built Environment, 20, pp. 337-346.

De Fabris, S., Longo, G., Medeossi, G., & Pesenti, R. (2014). Automatic generation of railway timetables based on a mesoscopic infrastructure model. Rail Transport Planning & Management, 4, pp.2-13.

De Luca, G., & Gallo, M. (2017). Artificial Neural Networks for forecasting user flows in transportation networks: literature review, limits, potentialities and open challenges. Proceedings of the 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems – IEEE MT-ITS 2017, Naples, Italy, June 2017, pp. 939-943.

De Martinis, V., Weidmann, U. A., & Gallo, M. (2014). Towards a simulation-based framework for evaluating energy-efficient solutions in train operation. WIT Transactions on The Built Environment, 135, pp. 721-732.

De Martinis, V., & Weidmann, U. A, (2015). Definition of energy-efficient speed profiles within rail traffic by means of supply design models. Research in Transportation Economics, 54, pp. 41-50.

De Shutter, B., Van den Boom, T. & Hegyi, A. (2002). Model predictive control approach for recovery from delays in railway systems. Transportation Research Record, 1793, pp. 15-20.

De-Los-Santos, A., Laporte, G., Mesa, J.M., & Perea, F. (2017). The railway line frequency and size setting problem. Public Transport, 9, pp.33-53.

Dewildea, T., Selsa, P., Cattryssea, D., & Vansteenwegena, P. (2014). Improving the robustness in railway station areas. European Journal of Operational Research, 235, pp. 276-286.

Dial, R. B. (1971). A probabilistic multipath traffic assignment model which obviates path enumeration. Transportation Research, 5, pp. 83-111.

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. Numerishe Matematik, 1, pp. 269-271.

Ding, Y., Liu, H., Bai, Y., & Zhou, F. (2011). A two-level optimization model and algorithm for energy-efficient urban train operation. Transportation Systems Engineering and Information Technology, 11, pp. 96-101.

Dollevoet, T., Huisman, D., Schmidt, M., & Schöbel, A. (2012). Delay management with rerouting of passengers, Transportation Science 46(1), pp. 74-89.

Dollevoet, T., Corman, F., D'Ariano, A., & Huisman, D. (2014). An iterative optimization framework for delay management and train scheduling. Flexible Services and Manufacturing, 26(4), pp. 490-515.

Domencich, T. A., & McFadden, D. (1975). Urban travel demand: a behavioural analysis. American Elsevier, New York (NY), USA.

Domínguez, M., Fernández-Cardador, A., Cucala, A. P., & Pecharromán, R. R. (2012). Energy savings in metropolitan railway substations through regenerative energy recovery and optimal design of ATO speed profiles. IEEE Transactions on Automation Science and Engineering, 9, pp. 496-504.

Dorigo, M. (1992). Optimization, learning and natural algorithms (in Italian). Ph.D. dissertation, Polytechnic of Milan, Italy.

Dorigo, M., & Stützle, T.(2004). Ant colony optimization. The MIT Press, Cambridge (MA), USA.

Douglas, N. (2012) Modelling train and passenger capacity. Report to Transport for NSW by DOUGLAS Economics, Sydney, Australia.

Dündar, S., & Şahin, I. (2013). Train re-scheduling with genetic algorithms and artificial neural networks for single-track railways. Transportation Research Part C, 27, pp. 1-15.

Eaton, J., & Yang, S. (2014). Dynamic railway junction rescheduling using population based Ant Colony Optimisation. Proceedings of the 14th UK Workshop on Computational Intelligence, Bradford, UK, September 2014.

EAV (2015a). Operating regulations (original document in Italian: Disposizioni di esercizio). Transportation and Rolling Stock Division, Ente Autonomo Volturno (EAV), Naples, Italy.

EAV (2015b). Emergency management (original document in Italian: Gestione delle emergenze). Transportation and Rolling Stock Division, Ente Autonomo Volturno (EAV), Naples, Italy.

Ehlert, A., Bell, M. G. H., & Grosso, S. (2006). The optimization of traffic count locations in road networks. Transportation Research Part B, 40, pp. 460-479.

Eickmann, C., Kettner, M., & Sewcyk, B. (2003). Integrating microscopic and macroscopic models for railway network evaluation. Proceedings of the European transport conference – ETC 2003, Strasbourg, France, October 2003.

Ercolani, M., Placido, A., D'Acierno, L., & Montella, B. (2014). The use of micro-simulation models for the planning and management of metro systems. WIT Transaction of Built Environment, 135, pp. 509-521.

Fan, B., Roberts, C., & Weston, P. (2012). A comparison of algorithms for minimising delay costs in disturbed railway traffic scenarios. Rail Transport Planning & Management, 2, pp. 23-33

Fei, X., Eisenman, S. M., & Mahmassani, H. S. (2007). Sensor coverage and location for real-time prediction in large-scale networks. Proceedings of the 86th TRB Annual meeting, Washington DC, USA, January 2007, pp. 1-15.

Feng, X., Wang, Q., Liu, Y., Xu, B., Liu, H., & Sun, Q. (2014). Ensuring a reasonable passenger capacity utilization rate of a train for its sustainably efficient transport. Applied Research and Technology, 12(2), pp. 279-288.

Feng, J., Li, X., Liu, H., Gao, X., & Mao, B. (2017). Optimizing the energy-efficient metro train timetable and control strategy in off-peak hour with uncertain passenger demands. Energies, 10(4), 436.

Festa, D. C., & Nuzzolo, A. (1989). Analisi sperimentale delle relazioni velocità-flusso per le strade urbane (in Italian). Le Strade, 1226, pp. 459-464.

Fischetti, M., Salvagnin, D., & Zanette, A. (2009). Fast approaches to improve the robustness of a railway timetable. Transportation Science, 43, pp. 321-335.

Flamini, M., & Pacciarelli, D. (2008). Real time management of a metro rail terminus. European Journal of Operational Research, 189 (3), pp. 746-761.

Fletcher, G., & El-Geneidy, A. (2013). The effects of fare payment types and crowding on dwell time: a fine-grained analysis. Transportation Research Record, 2351, pp. 124-132.

Florian, M., & Spiess, H. (1982). The convergence of diagonalization algorithms for asymmetric network equilibrium problems. Transportation Research Part B, 16, pp. 477-483.

Florian, M., & Chen, Y. (1995). A coordinate descendent method for the bi-level O-D matrix adjustment problem. Transportation Operations Research, 2, pp. 165-179.

Fu, Q., Liu, R., & Hess, S. (2012). A review on transit assignment modelling approaches to congested networks: a new perspective. Procedia – Social and Behavioral Sciences, 54, pp. 1145-1155.

Gallo, M., D’Acierno, L., Montella, B. (2010). A meta-heuristic approach for solving the urban network design problem. European Journal of Operational Research, 201, pp. 144-157.

Gallo, M., D’Acierno, L., & Montella, B. (2011a). A multimodal approach to bus frequency design. WIT Transactions on the Built Environment, 116, pp. 193-204.

Gallo, M., Montella, B., & D'Acierno, L. (2011b). The transit network design problem with elastic demand and internalisation of external costs: An application to rail frequency optimisation. Transportation Research Part C, 19(6), pp. 1276-1305.

Gallo, M., D'Acierno, L., & Montella, B. (2012). A meta-heuristic algorithm for solving the road network design problem in regional contexts. Procedia – Social and Behavioral Sciences, 54, pp. 84-95.

Gao, Y., Kroon, L., Schmidt, M., & Yang, L.(2016). Rescheduling a metro line in an over-crowded situation after disruptions. Transportation Research Part B, 93, pp. 425-449.

Gao, Y., Yang, L., & Gao, Z. (2017). Real-time automatic rescheduling strategy for an urban rail line by integrating the information of fault handling. Transportation Research Part C, 81, pp. 246-267.

Gavriilidou, A., Cats, O., Leffler, D., Corman, F., & Hoogendoorn, S. P. (2017). Real-time transfer synchronization of public transport services 1 using passenger data. Proceedings of the 96th TRB Annual Meeting, Washington DC, USA, January 2017.

Ghaemi, N.; Goverde, R. M. P., & Cats, O. (2016). Railway disruption timetable: Short-turnings in case of complete blockage. Proceedings of IEEE International Conference on Intelligent Rail Transportation – IEEE ICIRT 2016, August 2016, Birmingham, United Kingdom.

Ghavihaa, N., Campilloa, J., Bohlinb, M., & Dahlquista, E. (2017). Review of application of energy storage devices in railway transportation. Energy Procedia, 105, pp. 4561-4568.

Ghoseiri, K., Szidarovszky, F., & Asgharpour, M. J. (2004). A multi-objective train scheduling model and solution. Transportation Research Part B, 38, pp. 927-952.

Glover, F. (1986). Future paths for integer programming and links to artificial intelligence. Computers and Operations Research, 13, pp. 533-549.

Glover, F. (1989). Tabu search, Part I. ORSA Journal on Computing, 1(3), pp. 190-206.

Glover, F. (1990). Tabu search, Part II. ORSA Journal on Computing, 2(1), pp. 4-32.

Glover, F., Laguna, M., & Martì, R. (2003). Scatter search. In Ghosh, A. and Tsutsui, S. (Eds.): Advances in evolutionary computation: Theory and applications, Springer-Verlag, New York (NY), USA, pp. 519-537.

Goldberg, D.E. (1989). Genetic algorithms in search. Optimization and machine learning. Kluwer Academic Publishers, Boston (MA), USA.

Gonzalez-Gil, A., Palacin, R., & Batty, P. (2013). Sustainable urban rail systems: Strategies and technologies for optimal management of regenerative braking energy. Energy Conversion and Management, 75, pp. 374-388.

Goverde, R. M. P. (2005). Punctuality of Railway Operations and Timetable Stability Analysis. Ph.D. dissertation, Delft University of Technology, The Netherlands.

Goverde, R. M. P. (2007). Railway timetable stability analysis using max-plus system theory. Transportation Research Part B, 41 (2), pp. 179-201.

Goverde, R. M. P., Bešinovic, N., Binder, A., Cacchiani, V., Quaglietta, E., Roberti, R., & Toth, P. (2016). A three-level framework for performance-based railway Timetabling. Transportation Research Part C, 67, pp. 62-83.

Guenthner, R.P., & Hamat, K. (1988) Transit dwell time under complex fare structure. Transportation Engineering, 114(3), pp. 367-379.

Guo, X., Sun, H., Wu, J., Jin, J., Zhou, J., & Gao, Z. (2017). Multiperiod-based timetable optimization for metro transit networks. Transportation Research Part B, 96, pp. 46-67.

Hadas, Y., & Ceder, A. (2010). Optimal coordination of public-transit vehicles using operational tactics examined by simulation. Transportation Research Part C, 18, pp. 879-895

Hanafi, R., & Kozan, E., (2014). A hybrid constructive heuristic and simulated annealing for railway crew scheduling. Computers & Industrial Engineering, 70, pp. 11-19.

Hansen, I. A. (2006). State-of-the-art of railway operations research. WIT Transactions on The Built Environment, 88, pp. 565-577.

Hansen, I. A. & Pachl, J. (2008). Railway timetable and traffic: analysis-modelling-simulation, Eurail Press.

Hansen, I. A., Goverde, R. M.P., & Van Der Meer, D. J. (2010). Online Train Delay Recognition and Running Time Prediction. Proceedings of the13th International IEEE Annual Conference on Intelligent Transportation Systems – IEEE ITSC 2010, Funchal, Portugal, September 2010, pp 1783-1788.

Harris, N. G. (2006). Train boarding and alighting rates at high passenger loads. Advanced Transportation, 40(3), pp. 249 - 263.

Harris, N. G., & Anderson, R. J. (2007). An international comparison of urban rail boarding and alighting rates. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 221(4), pp. 521-526.

Hassannayebi, E., Sajedinejad, A., & Mardani, S. (2016). Disruption management in urban rail transit system: a simulation based optimization approach. Handbook of Research on Emerging Innovations in Rail Transportation Engineering, IgI-Global.

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.

Hazelton, M. L. (2001). Inference for origin-destination matrices: estimation, prediction and reconstruction. Transportation Research Part B, 35, pp. 667-676.

Heimburger, D. E., Herzenberg, A. Y., & Wilson, N. H. M. (1999). Using simple simulation models in operational analysis of rail transit lines: Case study of Boston’s Red Line. Transportation Research Record, 1667, pp. 21-30.

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), pp. 8-13.

Ho, T. K., & Yeung, T. H. (2001). Railway junction traffic control by heuristic methods. IEE Proceedings Electric Power Application, 148 (1), pp. 77-84.

Holland, J. H. (1975). Adaptation in natural and artificial systems. The University of Michigan Press, Ann Arbor (MI), USA.

Horowitz, J. L. (1981). Identification and diagnosis of specification errors in the multinominal logit model. Transportation Research Part B, 15(5), pp. 345-360.

Horowitz, J. L., Spermann, J. M., & Daganzo, C. F. (1982). An investigation on the accuracy of the Clark approximation for the multinomial probit model. Transportation Science, 16, pp. 382-401.

Howlett, P., Pudney, P., & Vu, X. (2009). Local energy minimization in optimal train control. Automatica, 45 (11), pp. 2692-2698.

Huang, G.-B., Zhu, Q.-Y., & Siew, C.-K. (2006). Extreme learning machine: theory and applications. Neurocomputing, 70, pp. 489-501.

Huang, S., Sadek, A. W., & Guo, L. (2013). A computational-based approach to estimating travel demand in large-scale microscopic traffic simulation models. Computing in Civil Engineering, 27(1), pp. 78-86.

Huang, Y., Ma, X., Su, S., & Tang, T. (2015). optimization of Train operation in multiple interstations with multi-population genetic algorithm. Energies, 8, pp. 14311-14329.

Huerlimann, D. (2001). Objektorientierte modellierung von infrastrukturelementen und betriebsvorgängen im eisenbahnwesen. Dissertation am Institut für Verkehrsplanung und Transporttechnik (IVT) der ETH Zürich.

Huo, J., Wu, J., Kang, L., & Wang, B. (2016). Railway timetable rescheduling based on priority and train order entropy. Computing in Civil Engineering, 30(5).

Ibaiondo, H., Romo, A. (2010). Kinetic energy recovery on railway systems with feedback to the grid. Proceedings of the 14th International Power Electronics and

Motion Control Conference –EPE-PEMC 2010, Ohrid, Macedonia, September 2010, pp. 94-97.

International Union of Railways (UIC), UIC CODE 451-1 (2000). Timetable recovery margins to guarantee timekeeping – Recovery margins, Paris, France: UIC, 2000.

Istituto Nazionale di Statistica – ISTAT (Italian National Institute of Statistics). Population and housing census, http://www.istat.it/it/censimento-popolazione (last access: December 2016).

Istituto Nazionale di Statistica – ISTAT (Italian National Institute of Statistics). Resident population, http://www.istat.it/it/popolazione (last access: December 2016).

Jacobs, J. (2004). Reducing Delays by Means of Computer-Aided 'on-the-Spot' Rescheduling, WIT Transactions on The Built Environment, 74, pp. 603-612.

Jamili, A., & Pourseyed Aghaee, M. (2015). Robust stop-skipping patterns in urban railway operations under traffic alteration situation. Transportation Research Part C, 61, pp. 63-74.

Jiang, Z., Li, F., Xu, R., & Gao, P. (2012). A simulation model for estimating train and passenger delays in large-scale rail transit networks. Journal of Central South University,19, pp. 3603-3613

Jiang, Z., Xie, C., Ji, T., & Zou, X. (2015). Dwell time modelling and optimized simulations for crowded rail transit lines based on train capacity. Promet, 27(2), pp. 125-135

Jong, J., & Chang, E. (2011). Investigation and estimation of train dwell time for timetable planning. Proceedings of the 9th World Congress on Railway Research – WCRR 2011, Lille, France, May 2011.

Kanai, S., Shiina, K., Harada, S., & Tomii, N. (2011). An optimal delay management algorithm from passengers' viewpoints considering the whole railway network. Rail Transport Planning & Management, 1, pp. 25-37.

Kang, L., Wu, J., & Sun, H. (2012).Using simulated annealing in a bottleneck optimization model at railway stations, Transportation Engineering, 138(11), pp. 1396-1402.

Kang, L., & Zhu, X., (2016). A simulated annealing algorithm for first train transfer problem in urban railway networks. Applied Mathematical Modelling, 40, pp. 419-435.

Kattan, L., & Abdulhai, B. (2006). Non iterative approach to dynamic traffic origin/ destination estimation using parallel evolutionary algorithms. Transportation Research Record, 1964, pp. 201-210.

Kattan, L., & Abdulhai, B. (2011). Comparative analysis of evolutionary, local search, and hybrid approaches to O/D traffic estimation. Transportation Engineering, 137(1), pp. 46-56.

Kattan, L., & Abdulhai, B. (2012). Sensitivity analysis of an evolutionary-based time-dependent origin/destination estimation framework, IEEE Transactions on Intelligent Transportation systems, 13(2), pp. 1442-1453.

Ke, B., Chen, M., Lin, C., (2009). Block-layout design using max–min ant system for saving energy on mass rapid transit systems. IEEE Transactions on Intelligent Transportation Systems, 10(2), pp. 226-235.

Ke, B. R., Lin, C. L., & Lai, C. W. (2011). Optimization of train-speed trajectory and control for mass rapid transit systems. Control Engineering Practice, 19, pp. 675-687.

Kecman, P., Corman, F., D'Ariano, A., & Goverde, R. M. P. (2013). Rescheduling models for railway traffic management in large-scale networks, Public Transport 5(1), pp. 95-123.

Kecman, P., & Goverde, R. M. P. (2015) Predictive modelling of running and dwell times in railway traffic. Public Transport, 7(3), pp. 295-319.

Kecman, P., Corman, F., & Meng, L. (2015a). Train delay evolution as a stochastic process. Proceedings of the 6th International Conference on Railway Operations Modelling and Analysis – RailTokyo2015, Narashino, Japan, March 2015.

Kecman, P., Corman, F., Peterson, A., & Joborn, M. (2015b). Stochastic prediction of train delays in real-time using bayesian networks. Proceedings of Conference on Advanced Systems in Public Transport – CASPT 2015, July 2015, Rotterdam, The Netherlands.

Kepaptsoglou, K., & Karlaftis, M. G. (2010). A model for analyzing metro station platform conditions following a service disruption. Proceedings of the13th International IEEE Annual Conference on Intelligent Transportation Systems – IEEE ITSC 2010, Funchal, Portugal, September 2010, pp. 1789-1794.

Keskin, K., & Karamancioglu, A. (2015). Optimal rail vehicle operation with multiple speed constraints. Proceedings of International Congress on Advanced Railway Engineering, Instanbul, Turkey, March 2015, pp. 381-385.

Kettner, M., & Sewcyk B. (2002). A model for transportation planning and railway network evaluation". Proceedings of the 9th World Congress on Intelligent Transport Systems – ITS 2002, Chicago (IL), USA, October 2002.

Khadilkar, H. (2017). Data-enabled stochastic modelling for evaluating schedule robustness of railway networks. Transportation Science, Articles in Advance, published online: https://doi.org/10.1287/trsc.2016.0703.

Khmelnitsky, E. (2000). On an optimal control problem of train operation. IEEE Transactions on Automatic Control, 45(7), pp. 1257-1266.

Khooban, Z., Farahani, R. Z., Miandoabchi, E., & Szeto, W. Y. (2015). Mixed network design using hybrid scatter search. European Journal of Operational Research, 247, pp. 699-710.

Kim, H., Baek, S., & Lim, Y. (2001). O-D matrices estimation using genetic algorithms from traffic counts, Transportation Research Record, 1771, pp. 156-163.

Kim, K., & Chien, S. I.-J. (2011). Optimal train operation for minimum energy consumption considering track alignment, speed limit, and schedule adherence. Transportation Engineering, 137(9), pp. 665-674.

Kim, K. M., Kim, K.T., Han, MS. (2011). A model and approaches for synchronized energy saving in timetabling. Proceedings of 9th World congress on railway research – WCRR 2011, Lille, France; May 2011.

Kirkpatrick S., Gelatt C. D., & Vecchi M. P. (1983). Optimization by simulated annealing, Science, 220(4598), pp. 671-680.

Kleinrock, L. (1975). Queuing System, vol. I e II, Wiley, New York, USA.

Kroon, L. G., Dekker, R., & Vromans, M. J. C. M. (2007). Cyclic railway timetabling: a stochastic optimization approach. In "Algorithmic Methods for Railway Optimization" (Editors: F. Geraets, L.G. Kroon, A. Shoebel, D. Wagner and C.D. Zaroliagis), Lecture Notes in Computer Science, 4359, Springer, Berlin, Germany, pp. 41-66.

Kroon, L., Huisman, D., Abbink, E., Fioole, P. J., Fischetti, M., Maroti, G., Schrijver, A., Steenbeek, A., & Ybema, R. (2009). The New Dutch Timetable: The OR revolution. Interfaces, 39(1), pp. 6-17.

Kroon, L.G., Maróti, G., & Nielsen, L. K., (2015). Rescheduling of railway rolling stock with dynamic passenger flows. Transportation Science, 49, pp. 165-184.

Kumazawa, K., Hara, K., & Koseki, T. (2010). A novel train rescheduling algorithm for correcting disrupted train operations in a dense urban environment. WIT Transactions on The Built Environment, 103, pp. 565-574.

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), pp. 51-62

Laguna, M. (2002). Scatter search. In Pardalos, P.M. and Resende, M.G.C. (Eds.): Handbook of applied optimization, University Press, Oxford, United Kingdom, 183-193.

Lam, W. H. K., Cheung, C. Y., & Poon YF. (1998). A study of train dwelling time at the Hong Kong mass transit railway system. Advanced Transportation, 32(3), pp. 285-295.

Lam, W. H. K., Cheung, C.-Y., & Lam, C. F. (1999). A study of crowding effects at the Hong Kong light rail transit stations. Transportation Research Part A, 33(5), pp. 401-415.

Lancien, D. & Fontaine M. (1981). Computing train schedules to save energy. Revue General des Chemins de Fer, 100, pp. 679-692.

Landex, A., Kaas, A. H., & Hansen, S. (2006). Railway Operation. Technical Report No. 4, Technical University of Denmark: Centre for Traffic and Transport.

Landex, A., & Jensen, L. W. (2013). Measures for track complexity and robustness of operation at stations. Rail Transport Planning & Management, 3, pp. 22-35.

Langdon, M. G. (1984). Improved algorithms for estimating choice probabilities in the multinomial probit model. Transportation Science, 18, pp. 267-299.

Larsen, R., Pranzo, M., D'Ariano, A., Corman, F., & Pacciarelly, D. (2014). Susceptibility of optimal train schedules to stochastic disturbances of process times. Flexible Services and Manufacturing, 26, pp. 466-489.

LeBlanc, L. J., Morlok, E. K., & Pierskalla, W. P. (1975). An efficient approach to solving the road network equilibrium traffic assignment problem. Transportation Research, 9, pp. 309-318.

Levine, J., & Torng, G. (1994). Dwell-time effects of low-floor bus design. Transportation Engineering, 120(6), pp. 914-929.

Levinson, H. S. (1983). Transit travel time performance. Transportation Research Record, 915, pp. 1-6.

Li, X., & Lo, H. K. (2014). Energy minimization in dynamic train scheduling and control for metro rail operations. Transportation Research Part B, 70, pp. 269-284.

Li, D., Goverde, R. M. P., Daamen, W., & He, H. (2014a). Train dwell time distributions at short stop stations. Proceedings of the 17th International IEEE Conference on Intelligent Transportation Systems – IEEE ITCS 2014, Qingdao, China, October 2014, pp. 2410-2415.

Li, X., Shou, B., & Ralescu, D. (2014b). Train rescheduling with stochastic recovery time: A new track-backup approach. IEEE Transactions on systems, man, and cybernetics: systems, 44(9), pp. 1216-1233.

Li, D., Daamen, W., & Goverde, R. M. P. (2016) Estimation of train dwell time at short stops based on track occupation event data: A study at a Dutch railway station. Advanced Transportation, 50, pp. 877-896.

Liebchen, C., & Möhring, R. H. (2002). A Case Study in Periodic Timetabling. Electronic Notes in Theoretical Computer Science, 66(6), pp. 1-14.

Liebchen, C. (2004). Symmetry for Periodic Railway Timetables. Electronic Notes in Theoretical Computer Science, 92, pp. 34-51.

Liebchen, C. (2008). The first optimized railway timetable in practice. Transportation Science, 42(4), pp. 420-435.

Liebchen, C., Lübbecke, M., Möhring, R. H. & Stiller, S. (2009). The concept of recoverable robustness, linear programming recovery, and railway applications. In 'Robust and Online Large-Scale Optimization' (Editors: R.K. Ahuja, R.F. Möhring, and C.D. Zaroliagis), Lecture Notes in Computer Science, 5868, Springer, Berlin, Germany, pp. 1-27.

Liu, R., & Golovitcher, I., (2003). Energy-efficient operation of rail vehicles. Transportation Research Part A, 37(10), pp. 917-932.

Lo, H.P., & Chan, C. P., (2003). Simultaneous estimation of an origin–destination matrix and link choice proportions using traffic counts. Transportation Research Part A, 37(9), pp. 771-788.

Longo, G., & Medeossi, G. (2013). An approach for calibrating and validating the simulation of complex rail networks. Proceedings of the 92nd TRB Annual Meeting, Washington DC, USA, January 2013.

Louwerse, I., & Huisman, D. (2014). Adjusting a railway timetable in case of partial or complete blockades, European Journal of Operational Research, 235(3), pp. 583-593.

Lu, C. C., Zhou, X., & Zhang, K., (2013a). Dynamic origin–destination demand flow estimation under congested traffic conditions. Transportation Research Part C, 34, pp. 16-37.

Lu, S. F., Hillmansen, S., Ho, T. K., & Roberts, C., (2013b). Single-train trajectory optimization. IEEE Transactions on Intelligent Transportation Systems, 14(2), pp. 743-750.

Lu, L., Xu, Y., Antoniou, C., & Ben-Akiva, M. (2015). An enhanced SPSA algorithm for the calibration of Dynamic Traffic Assignment models. Transportation Research Part C, 51, pp. 149-166.

Lukaszewicz, P. (2000). Driving techniques and strategies for freight trains. WIT Transactions on The Built Environment, 50, pp. 1065-1073.

Maher, M. (1983). Inferences on trip matrices from observations on link volumes: a Bayesian statistical approach. Transportation Research Part B, 17(6), pp. 435-447.

Mahmassani, H. S., Peeta, S. (1995). System optimal dynamic assignment for electronic route guidance in a congested traffic network. In: Gartner N.H., Improta G. (eds) Urban Traffic Networks. Transportation Analysis. Springer, Berlin, Heidelberg, pp. 3-37.

Mannino, C., & Mascis, A. (2009). Real-time traffic control in metro stations. Operations Research, 57(4), pp. 1026-1039.

Mannino, C., (2011). Real-time traffic control in railway systems. Proceedings of 11th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems – ATMOS 2011, September 2011, Saarbrücken, Germany.

Manski, C. F., & McFadden, D. (1981). Structural analysis of discrete data with econometric applications. The MIT Press, Cambridge (MA), USA.

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, pp. 516-539.

Martì, R., Laguna, M., & Glover, F. (2003). Principles of scatter search. European Journal of Operational Research, 169, pp. 359-372.

Marzano, V., Papola, A., & Simonelli, F. (2009). Limits and perspectives of effective O-D matrix correction using traffic counts. Transportation Research Part C, 17(2), pp. 120-132.

Mascis, A., & Pacciarelli, D. (2002). Job shop scheduling with blocking and no-wait constraints. European Journal of Operational Research, 143(3), pp. 498-517.

Mathews, J. H., & Fink, K. D. (2004). Numerical Methods Using MATLAB. Pearson Prentice Hall, Upper Saddle River (NJ), USA.

Mazzarello, M., & Ottaviani, E. (2007). A traffic management system for real-time traffic optimization in railways. Transportation Research Part B, 41(2), pp. 246-274.

Meng, X., Jia, L., & Qin, Y. (2010). Train timetable optimizing and rescheduling based on improved particle swarm algorithm. Transportation Research Record, 2197, pp. 71-79.

Meng, L., & Zhou, X. (2014). Simultaneous train rerouting and rescheduling on an N-track network: a model reformulation with network-based cumulative flow variables. Transportation Research Part B, 67, pp. 208-234.

Meng, L., & Zhou, X. (2011). Robust single-track train dispatching model under a dynamic and stochastic environment: a scenario-based rolling horizon approach. Transportation Research Part B, 45 (7), 1080-1102.

Metropolis, N., Rosenbluth, A., Rosenbluth M., Teller, A., & Teller, E. (1953). Equation of state calculations by fast computing machines. Journal of Chemical Physics, 21, pp. 1087-1092.

Meyer, J. F., Movaghar, A., & Sanders, W. H. (1985). Stochastic activity networks: structure, behavior and application. Proceedings of International Workshop on Timed Petri Nets, Torino, Italy, July 1985, pp. 106-115.

Middelkoop, D., & Bouwman, M. (2001). SIMONE: Large scale train network simulations. Proceedings of the 2001 Winter Simulation Conference, Piscataway (NJ), USA, December 2001, pp. 1042-1047.

Middelkoop, A. D. (2010). Headway generation with ROBERTO. WIT Transactions on the built environment, 114, pp. 431-439.

Min, Y. H., Park, M. J., Hong, S. P., & Hong, S. H. (2011). An appraisal of a column generation-based algorithm for centralized train-conflict resolution on a metropolitan railway network. Transportation Research Part B, 45, pp. 409-429.

Miyatake, M., & Matsuda, K. (2009). Energy saving speed and charge/discharge control of a railway vehicle with on-board energy storage by means of an optimization model IEEJ, Transactions on Electrical and Electronic Engineering, 4, pp. 771-778.

Miyatake, M., & Ko, H. (2010). Optimization of train speed profile for minimum energy consumption. IEEJ Transactions on Electrical and Electronic Engineering, 5, pp. 263-269.

Montella, B., Gallo, M., & D'Acierno, L. (2000). Multimodal network design problems. Advances in Transport, 5, pp. 405-414.

Moreira-Matias, L., Gama, J., Ferreira, M., Mendes-Moreira, J., & Damas, L. (2016). Time-evolving O-D matrix estimation using high-speed GPS data streams. Expert Systems With Applications, 44, pp. 275-288.

Movaghar, A., & Meyer, J. F. (1984). Performability modeling with stochastic activity networks. Proceedings of Real-Time Systems Symposium – RTSS 1984, Austin (TX), USA, December 1984, pp. 215-224.

Nachtigall, K. (1996). Periodic network optimization with different arc frequencies. Discrete Applied Mathematics, 69(1), pp. 1-17.

Nagy, V. (2016). Theoretical method for building OD matrix from AFC data. Transportation Research Procedia, 14, pp. 802-1808.

Narayanaswami, S., & Rangaraj, N. (2013). Modelling disruptions and resolving conflicts optimally in a railway schedule. Computers & Industrial Engineering, 64(1), pp. 469-481.

Nash, A., & Huerlimann, D. (2004). Railroad simulation using OpenTrack. WIT Transactions on The Built Environment, 74, pp. 45-54

Nasri, A., Fekri Moghadam, M., & Mokhtari, H. (2010). Timetable optimization for maximum usage of regenerative energy of braking in electrical railway systems. Proceedings of the International symposium on power electronics, electrical drives, automation and motion – SPEEDAM 2010, Pisa, Italy, June 2010, pp. 1218-1221.

Newell, G. F. (1971). Application of queuing theory. Chapman and Hall LTD, London, United Kingdom.

Newell, G. F. (1980). Traffic flows in transportation networks. MIT Press, Cambridge, Mass.

Nguyen, S., & Pallottino, S. (1986). Assegnazione dei passeggeri ad un sistema di linee urbane: determinazione degli ipercammini minimi (in Italian). Ricerca Operativa, 32, pp. 207-230.

Nigro, M. (2009). Adjustment of the O-D matrix under the intra-period dynamic assumption with different data sources (in Italian). Ph.D. dissertation, Roma Tre University, Italy.

Nihan, N. L., & Davis, G. A. (1987). Recursive estimation of origin–destination matrices from input–output counts. Transportation Research Part B, 21, pp. 149-163.

Nihan, N. L., & Davis, G. A. (1989). Application of prediction-error minimization and maximum likelihood to estimate intersection O–D matrices from traffic counts. Transportation Science, 23, pp. 77-90.

Niu, H., & Zhou, X. (2013). Optimizing urban rail timetable under time-dependent demand and oversaturated conditions. Transportation Research Part C, 36, pp. 212-230.

Niu, H., Zhou, X., & Gao, R. (2015). Train scheduling for minimizing passenger waiting time with time-dependent demand and skip-stop patterns: Nonlinear integer programming models with linear constraints. Transportation Research Part B, 76, pp. 117-135.

Novačko, L., Šimunović, Lj., & Krasić, D. (2014). Estimation of origin-destination trip matrices for small cities. Promet – Traffic & Transportation, 26(5), pp. 419-428.

Nuzzolo, A., Russo, F., & Crisalli, U. (2002). Transit Network Modelling. The schedule-based dynamic approach. Franco Angeli, Milan, Italy.

Nuzzolo, A., Crisalli U., & Rosati, L. (2012). A schedule-based assignment model with explicit capacity constraints for congested transit networks. Transportation Research Part C, 20, pp. 16-33.

Okutani, I., & Stephanades, Y. (1984). Dynamic prediction of traffic volume through kalman filtering theory. Transportation Research Part B, 18, pp. 1-11.

Oliveira, E., & Smith, B. M. (2000). A job-shop scheduling model for the single-track railway scheduling problem. Technical Report No. 21, School of Computing, University of Leeds, England.

Oliveira, E. (2001). Solving single-track railway scheduling problem using constraint programming. Ph.D. dissertation, Delft University of Technology, The Netherlands.

Ortuzar, J. de D. (1992). Stated Preference in travel demand modelling. Proceedings of the 6th World Conference on Transportation Research – WCTR 1992, Lyon, France, July 1992.

Ortuzar, J., de D. & Willumsen, L. G. (2011). Modelling transport - 4th edition. John Wiley and Sons, Ldt, Chichester, United Kingdom.

Osservatorio sui comportamenti di mobilità degli italiani – AudiMob (Observatory on the Italian mobility behaviour). Regional mobility statistics, http://www.isfort.it/sito/statistiche/Audimob.htm (last access: December 2016).

Pape, U. (1974). Implementation and efficiency of Moore-algorithms for the shortest route problem. Mathematical programming, 7, pp. 212-222.

Papola, A., Marzano, V., & Tinessa, F. (2017). A new closed-form route choice model with flexible covariance structure: the CoNL route choice model. Presented at the XXII SIDT National Scientific Seminar, Bari, Italy, September 2017.

Park, B., & Zhu, K. (2007). Time-dependent origin-destination estimation: Genetic algorithm-based optimization with updated assignment matrix. KSCE Journal of Civil Engineering, 11(4), pp. 199-207.

Peeters, L. (2003). Cyclic railway timetable optimization. Ph.D. dissertation, Erasmus University Rotterdam, The Netherlands.

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, pp. 23-34.

Placido, A., Cadarso L., & D'Acierno, L. (2014a) Benefits of a combined micro-macro approach for managing rail systems in case of disruptions. Transportation Research Procedia, 3, pp. 195-204, ISSN: 2352-1465.

Placido, A., De Martinis, V., Montella, B., Gallo, M., & D'Acierno, L. (2014b). Effects of travel demand levels on optimal strategies for metro system management in failure contexts. Procedia – Social and Behavioral Sciences, 111, pp. 819-828.

Placido, A. (2015). The definition of a model framework for the planning and the management phases of the rail system in any kind of service condition. Ph.D. dissertation, University of Naples Federico II, Italy.

Placido, A., D'Acierno, L., Botte, M., & Montella, B. (2015) Effects of stochasticity on recovery solutions in the case of high-density rail/metro networks. Proceedings of the 6th International Conference on Railway Operations Modelling and Analysis – RailTokyo2015, Narashino, Japan, March 2015.

Potts, R. B., & Oliver, R. M. (1972). Flows in transportation networks. Academic Press, New York (NY), USA.

Pouryousef, H., & Lautala, P. (2015). Hybrid simulation approach for improving railway capacity and train schedules. Rail Transport Planning & Management, 5, pp. 211-224.

Powell, W. B., & Sheffi, Y. (1982). The convergence of equilibrium algorithms with predetermined step sizes. Transportation Science, 16(1), pp. 45-55.

Puong, A. (2000). Dwell time model and analysis for the MBTA Red Line. Research Memo, Massachusetts Institute of Technology, Cambridge (MA), USA.

Quaglietta, E. (2011). A Microscopic Simulation Model for supporting the design of railway systems: development and applications. Ph.D. dissertation, University of Naples Federico II, Italy.

Quaglietta, E., Punzo, V., Montella, B., Nardone, R., & Mazzocca, N. (2011). Towards a hybrid mesoscopic-microscopic railway simulation model. Proceedings of the 2nd IEEE International Conference on Models and Technologies for Intelligent Transportation Systems – IEEE MT-ITS 2011, Leuven, Belgium, June 2011.

Quaglietta, E., & Punzo, V. (2013). Supporting the design of railway systems by means of a Sobol variance-based sensitivity analysis. Transportation Research Part C, 34, pp. 38-54.

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 16th International IEEE Conference on Intelligent Transportation Systems – IEEE ITSC 2013, October 2016, The Hague, The Netherlands, pp. 1035 - 1040.

Radtke, A. & Bendfeldt, J. (2001). Handling of railway operation problems with RailSys. Proceedings of the 5th World Congress on Rail Research – WCRR 2001, Cologne, Germany, November 2001.

Radtke, A. & Watson, R. (2007). Railway simulation in the United Kingdom, RTR, 1, pp. 24-28.

Ramos, A., Pena, M., Fernndez-Cardador, A., & Cucala, A. P., (2007). Mathematical programming approach to underground timetabling problem for maximizing time synchronization. Proceedings of the International Conference on Industrial Engineering and Industrial Management, Madrid Spain, December 2007, pp. 88-95.

Rausand, M., & Høyland, A. (2004) System reliability theory: models, statistical methods, and applications.Wiley, Hoboken (NJ), USA.

Romo, L., Turner, D., & Ng, L. S. B. (2005). Cutting traction power costs with wayside energy storage systems in rail transit systems. Proceedings of Joint Rail Conference – JRC2005, Pueblo (CO) USA, March 2005, pp. 187-192.

Roy, B., & Sussmann, B. (1964). Les Problèmes d'ordonnancement avec contraintes disjonctives. Note DS n.9 bis, SEMA, Montrouge.

Samà, M., Pellegrini, P., D'Ariano, A., Rodriguez, J., & Pacciarelli, D. (2016). Ant colony optimization for the real-time train routing selection problem. Transportation Research Part B, 85, pp. 89-108.

Samà, M., D'Ariano, A., Corman, F., & Pacciarelli, D., (2017a). A variable neighbourhood search for fast train scheduling and routing during disturbed railway traffic situations. Computers & Operations Research, 78(1), pp. 480-499.

Samà, M., D'Ariano, A., Pacciarelli,D., Pellegrini, P., & Rodriguez, J. (2017b). Ant Colony Optimization for train routing selection: operational vs tactical application. Proceedings of the 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems – IEEE MT-ITS 2017, Naples, Italy, June 2017, pp. 298-303.

Sanders, W. H., & Meyer, J. F. (2001). Stochastic activity networks: formal definitions and concepts. In 'Formal Methods and Performance Analysis' (Editors: E. Brinksma, H. Hermanns, and J.-P. Katoen), Lecture Notes in Computer Science, 2090, Springer, Berlin, Germany, pp. 315-343.

Sato, K., Tamura, K., & Tomii, N. (2013). A MIP-based timetable rescheduling formulation and algorithm minimizing further inconvenience to passengers. Rail Transport Planning & Management, 3, pp. 38-53.

Savrasovs, M., & Pticina, I. (2017). Methodology of OD matrix estimation based on video recordings and traffic counts. Procedia Engineering, 178, pp.289-297.

Sayarshad, H. R., & Ghoseiri, K. (2009). A simulated annealing approach for the multi-periodic rail-car fleet sizing problem. Computers & Operations Research, 36, pp. 1789-1799.

Scheepmaker, G. M., Goverde, R. M. P. (2015). The interplay between energy-efficient train control and scheduled running time supplements. Rail Transport Planning & Management, 5, pp. 225-239.

Schlechte, T., Borndörfer, R., Erol, B., Graffagnino, T., & Swarat, E. (2011). Micro–macro transformation of railway networks. Rail Transport Planning & Management, 1, pp. 38-48.

Schöbel, A. (2007). Integer programming approaches for solving the delay management problem. In: 'Robust and Online Large-Scale Optimization' (Editors: F. Geraets, L. G. Kroon, A. Schöbel, D. Wagner, C. D. Zaroliagis), Lecture Notes in Computer Science, 4359, Springer, Berlin, Germany, pp. 145-170.

Serafini, P., & Ukovich, W. (1989). A mathematical model for periodic event scheduling problems, SIAM Journal on Discrete Mathematics, 2, pp. 550-581.

Seriani, S, & Fernandez, R. (2015). Pedestrian traffic management of boarding and alighting in metro stations. Transportation Research Part C, 53, pp. 76-92.

Sewcyk, B. & Kettner, M. (2001). Network Evaluation Model NEMO. Proceedings of the 5th World Congress on Rail Research – WCRR 2001, Cologne, Germany, November 2001.

Shakibayifar, M., Corman, F., & Hassannayebi, E. (2017). An integrated rescheduling model for minimizing train delays in the case of line blockage. Operational Research, Online First: https://doi.org/10.1007/s12351-017-0316-7.

Shao, M., Sun, L., & Shao, X. (2016). Sensor location problem for network traffic flow derivation based on turning ratios at intersection. Mathematical Problems in Engineering, 2016, article ID 9012724.

Sheffi, Y., & Powell, W. B. (1981). A comparison of stochastic and deterministic traffic assignment over congested networks. Transportation Research Part B, 15, pp. 53-64

Sheffi, Y., & Powell, W. B. (1982). An algorithm for the equilibrium assignment problem with random link times. Networks, 12, pp. 191-207.

Sheu, J. W., Lin, W. S. (2011). Automatic train regulation with energy saving using dual heuristic programming. IET Electrical Systems in Transportation, 1(2), pp. 80-89.

Sicre, C., Cucala, P., Fernández, A., Jiménez, J. A.; Ribera I. & Serrano A. (2010). A method to optimise train energy consumption combining manual energy efficient driving and scheduling. WIT Transactions on The Built Environment, 114, pp. 549-560.

Sicre, C., Cucala, A. P., Fernandez, A., & Lukaszewicz, P. (2012). Modeling and optimizing energy-efficient manual driving on high-speed lines. IEEJ Transactions on Electrical and Electronic Engineering, 7, pp. 633-640.

Sicre, C., Cucala, A. P., & Fernández-Cardador, A. (2014). Real time regulation of efficient driving of high speed trains based on a genetic algorithm and a fuzzy model of manual driving. Engineering Applications of Artificial Intelligence, 29, pp. 79-92.

Siegler, L. E. (1987). Leonardo Pisano Fibonacci, The Book of Squares, An Annotated Translation Into Modern English. Academic Press Incorporated, Orlando (FL), USA.

Silvestrin, P. V., & Ritt, M. (2017). An iterated tabu search for the multi-compartment vehicle routing problem, Computers and Operations Research, 81, pp. 192-202.

Simonelli, F., Marzano, V., Papola A., & Vitiello, I. (2012). A network sensor location procedure accounting for o–d matrix estimate variability. Transportation Research Part B, 46, pp.1624-1638.

Smith, M. J. (1979). The existence, uniqueness and stability of traffic equilibria. Transportation Research Part B, 13(4), pp. 295-304.

Spall, J.C. (1992). Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. IEEE Transactions on Automatic Control, 37(3), pp. 332-341.

Spall, J.C. (1998). Implementation of the simultaneous perturbation algorithm for stochastic optimization, IEEE Transactions on Aerospace and Electronic Systems, 34, pp. 817-823.

Sparing, D. , & Goverde, R. M. P. (2013). An optimization model for periodic timetable generation with dynamic frequencies. Proceedings of the 16th International IEEE Conference on Intelligent Transportation Systems – IEEE ITSC 2013, October 2013, The Hague, The Netherlands, pp. 785-790

Sparing, D., & Goverde, R. M. P. (2017). A cycle time optimization model for generating stable periodic railway timetables. Transportation Research Part B, 98, pp. 198-223

Spiess, H., Florian, M. (1989). Optimal strategies: a new assignment model for transit networks. Transportation Research Part B, 32(2), pp. 82-102.

Srinivas, M., & Patnaik, L. M. (1994). Genetic algorithms: A survey. Computer, 27(6), pp. 17-26

Stathopoulos, A., & Tsekeris, T. (2004). Hybrid meta-heuristic algorithm for the simultaneous optimization of the O-D trip matrix estimation. Computer-Aided Civil and Infrastructure Engineering, 19(6). pp. 421-435.

Steiner, R., Klohr, M., & Pagiela, S. (2007). Energy storage system with UltraCaps on Board of Railway Vehicles. Proceeding of European Conference on Power Electronics and Applications, Aalborg, Denmark, September 2007.

Su, S., Li, X., Tang, T., & Gao, Z. (2013). A subway train timetable optimization approach based on energy-efficient operation strategy. IEEE Transactions on Intelligent Transportation Systems, 14(2), pp. 883-893.

Subiono (2000). On classes of min-max-plus systems and their applications. Ph.D. dissertation, Delft University of Technology, The Netherlands.

Sun, Y., Cao, C., & Wu, C. (2014). Multi-objective optimization of train routing problem combined with train scheduling on a high-speed railway network. Transportation Research Part C, 44, pp. 1-20

Tanaka, S., Kumazawa, K., & Koseki, T. (2009). Passenger flow analysis for train rescheduling and its evaluation. Proceedings of International Symposium on Speed-up, Safety and Service Technology for Railway and Maglev Systems 2009 – STECH'09, Niigata, Japan, June 2009.

Tavassoli, A., Mesbah, M., & Hickman, M. (2017).Quantifying error in transit assignment using smart card data in a large-scale multimodal transit network. Proceedings of Transportation Research Board 96th Annual Meeting, Washington DC (USA). January 2017.

Teymourfar, R., Asaei, B., Iman-Eini, H., & Nejati fard, R. (2012). Stationary super-capacitor energy storage system to save regenerative braking energy in a metro line. Energy Conversion and Management, 56, pp. 206-214.

Tirachini, A., Hensher, D. A., & Rose, J. M. (2013). Crowding in public transport systems: Effects on users, operation and implications for the estimation of demand. Transportation Research Part A, 53, pp. 36-52.

Toledo, T., Koutsopoulos, H. N., Davol, A., Ben-Akiva, M., Burghout, W., Andreasson, I., Johansson, T. & Lundin, C. (2003). Calibration and validation of microscopic traffic simulation tools: Stockholm case study. Transportation Research Record, 1831, pp. 65-75.

Toledo, T., & Kolechkina, T., (2013). Estimation of dynamic origin-destination matrices using linear assignment matrix approximations, IEEE Transactions on Intelligent Transportation Systems, 14(2), pp. 618-626.

Toletti, A., & Weidman, U. (2016). Modelling customer inconvenience in train rescheduling. Proceedings of the 16th Swiss Transport Research Conference – STRC 2016, Ascona, Switzerland, May 2016.

Toletti, A., De Martinis, V., & Weidmann, U. (2016). Energy savings in mixed rail traffic rescheduling: an RCG approach. Proceedings of the 19th IEEE International Conference on Intelligent Transportation Systems – IEEE ITSC 2016, Rio de Janeiro (Brazil), November 2016, pp. 2430-2435.

Tolouei, R., Psarras, S., & Prince, R. (2016). Origin-destination trip matrix development: conventional methods vs. use of mobile phone data. Proceedings of the European transport conference - ETC 2016, Barcelona, Spain, October 2016.

Tonosaki, Y., Koizumi, Y., Tajima, M., Miyoshi, M., Takeba, T., & Miyatake, M. (2016). Punctual train operation with energy-saving driving advisory system in dense traffic railway. Proceedings of the IEEE International Conference on Intelligent Rail Transportation – IEEE ICIRT 2016, Birmingham, United Kingdom, August 2016.

Törnquist, J., & Persson, J. A. (2005). Train traffic deviation handling using Tabu Search and Simulated Annealings. Proceedings of the 38th Hawaii International Conference on System Sciences – HICSS 2005, Big Island (HI), USA, January 2005.

Törnquist, J., & Persson, J. A. (2007). N-tracked railway traffic re-scheduling during disturbances. Transportation Research Part, B 41(3), pp. 342-362.

Transportation Research Board (2000). Highway capacity manual. National Research Council. Washington D.C.

Tsekeris, T., Dimitriou, L., & Stathopoulos, A. (2007). Simultaneous origin-destination matrix estimation in dynamic traffic networks with Evolutionary Computing. In 'Applications of Evolutionary Computing' (Editor: M. Giacobini), Lecture Notes in Computer Science, 4448, Springer, Berlin, Germany, pp 668-677.

Tympakianaki, A., Koutsopoulos, H. N. & Jenelius, E. (2015). c-SPSA: Cluster-wise simultaneous perturbation stochastic approximation algorithm and its application to dynamic origin–destination matrix estimation. Transportation Research Part C, 55, pp. 231-245.

Umiliacchi, S., Nicholson, G., Zhao, N., Schmid, F., & Roberts, C. (2016). Delay management and energy consumption minimisation on a single-track railway, IET Intelligent Transport Systems 10(1), pp. 50-57.

Veelenturf, L. P., Kidd, M. P., Cacchiani, V., Kroon, L. G. & Toth, P. (2016). A railway timetable rescheduling approach for handling large scale disruptions. Transportation Science, 50(3), pp. 841-862.

Veelenturf, L. P., Kroon, L. G., & Maróti, G. (2017). Passenger oriented railway disruption management by adapting timetables and rolling stock schedules. Transportation Research Part C, 80, pp.133-147.

Vuchic, V. R. (2005). Urban transit operations, planning and economics. John Wiley & Sons, Incorporated, Hoboken (NJ), USA.

Walpena, J., Macinelli, E. M., & Lotito, P. A. (2015). A heuristic for the OD matrix adjustment problem in a congested transport network. European Journal of Operational Research, 242, pp. 807-819.

Wang, L., Qin, Y., Xu, J., & Jial, L. (2012). A fuzzy optimization model for high-speed railway timetable rescheduling. Discrete Dynamics in Nature and Society, 2012, article ID 827073.

Wang, Y., Ma, X., Liu, Y., Gong, K., Henrickson, K. C., Xu, M., & Wang, Y. (2016). A two-stage algorithm for origin-destination matrices estimation considering dynamic dispersion parameter for route choice. PLOS ONE 11(1), pp. 1-24.

Webster, F. W. (1958). Traffic signal settings. Technical Paper 39, Road Research Laboratory, HMSO, London, United Kingdom.

Webster, F. W., Cobbe, B. M. (1966). Traffic signals. Technical Paper 56, Road Research Laboratory, HMSO, London, United Kingdom.

Wendler, E. (2001). Quality management in the operation planning process by means of harmonized modelling. Proceedings of the 5th World Congress on Rail Research – WCRR 2001, Cologne, Germany, November 2001.

Weston, J. G. (1989). Train service model – technical guide. London Underground operational research note 89/18.

Wiggenraad PBL (2001) Alighting and boarding times of passengers at Dutch railway stations analysis of data collected at 7 stations in October 2000. Ph. D. dissertation, Delft University of Technology, The Netherlands.

Wille, A., & Bühlmann, P. (2004). Tri-graph: a novel graphical model with application to genetic regulatory network, Technical report, ETH Zürich, Switzerland.

Wille, A., & Bühlmann, P. (2006). Low-order conditional independence graphs for inferring genetic networks. Statistical Applications in Genetics & Molecular Biology, 5(1), pp.1-34.

Williams, H. C. W. L. (1977). On the formation of travel demand models and economic evaluation measures of user benefit. Environment and Planning A, 9(3), pp. 285-344.

Wirasinghe, S. C., Szplett, D. (1984). An investigation of passenger interchange and train scheduling time at LRT stations: (ii) estimation of standing time. Advanced Transportation, 18(1), pp. 13-24.

Wong, K. K., & Ho, T. K. (2004). Coast control for mass rapid transit railways with searching methods. IEE Proceedings – Electric Power Applications,151(3), pp. 365-376.

Wong, R. C. W., & Leung, J. M. Y. (2004). Timetable synchronization for mass transit railway. Proceedings of the 9th International Conference on Computer-Aided Scheduling of Public Transport – CASPT 2004, San Diego, (CA), USA, August 2004.

Wong, K. K., & Ho, T. K. (2007). Dwell-time and run-time control for DC mass rapid transit railways. IET Electric Power Applications, 1(6), pp. 956-966.

Wong, R. C. W., Yuen, T. W. Y., Fung, K. W., & Leung, J. M. Y. (2008). Optimizing timetable synchronization for rail mass transit. Transportation Science, 42(1), pp. 57-69.

Wu, J. H., Chen, Y., & Florian, M. (1995). The continuous dynamic network loading problem: a mathematical formulation and solution method. Transportation Research part B, 32, pp.173-187.

Wu, J., Kang, L., Sun, H., & Jia, X. (2013). Track allocation optimization in railway station: mean-variance model and case study. Transportation Engineering, 139(5), pp. 540-547.

Xiaoming, X., Keping, L., Xiang, L. (2016). A multi-objective subway timetable optimization approach with minimum passenger time and energy consumption. Advanced Transportation, 50, pp. 69-95.

Xu, X., Liu, J., Li, H., & Zhou, Y. (2013). Probabilistic model for remain passenger queues at subway station platform. Journal of Central South University, 20, pp. 837-844.

Xu, X., Li, K., & Yang, L. (2016a). Rescheduling subway trains by a discrete event model considering service balance performance. Applied Mathematical Modelling, 40 pp.1446-1466.

Xu, X., Li, K., Li, X. (2016b). A multi-objective subway timetable optimization approach with minimum passenger time and energy consumption. Advanced Transportation, 50, pp. 69-95.

Xu, P., Corman, F., Peng, Q., & Luan, X. (2017a). A Timetable rescheduling approach and transition phases for high speed railway traffic during disruptions. Proceedings of Transportation Research Board 96th Annual Meeting, Washington DC (USA). January 2017.

Xu, P., Corman, F., Peng, Q., & Luan, X. (2017b). A train rescheduling model integrating speed management during disruptions of high-speed traffic under a quasi-moving block system. Transportation Research Part B, 104, pp. 638-666.

Xu, W., Zhao, P., & Ning, L. (2017c). A passenger-oriented model for train rescheduling on an urban rail transit line considering train capacity constraint. Mathematical Problems in Engineering, 2017, article ID 1010745.

Yamamura, A., Koresawa, M., Aadchi, S., Inagi, & T., Tomii, N. (2013). Dwell time analysis in Railway Lines using Multi Agent Simulation. Proceedings of the 13th World Conference on Transportation Research – WCTR 2013, Rio de Janeiro, Brazil, July 2013.

Yan, X., Cai, B., Ning, B., & ShangGuan, W. (2016). Online distributed cooperative model predictive control of energy-saving trajectory planning for multiple high-speed train movements. Transportation Research Part C, 69, pp. 60-78.

Yan, F., & Goverde, R. M. P. (2017). Railway timetable optimization considering robustness and overtakings. Proceedings of the 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems – IEEE MT-ITS 2017, Naples, Italy, June 2017, pp. 292-297.

Yang, H., Meng, Q., & Bell, M. G. H. (2001). Simultaneous estimation of the origin destination matrices and travel-cost coefficient for congested networks in a stochastic user equilibrium. Transportation Science, 35(2), pp. 107-123.

Yang, L., Li, K., Gao, Z., & Li, X., (2012). Optimizing trains movement on a railway network, Omega, 40, pp. 619-633.

Yang, X., Li, X., Gao, Z., Wang, H., & Tang, T., (2013). A cooperative scheduling model for timetable optimization in subway systems. IEEE Transactions on Intelligent Transportation Systems, 14(1), pp. 438-447.

Yang, H., & Yagar, S. (1995). Traffic assignment and signal control in saturated road networks. Transportation Research Part A, 29(2), pp. 125-139.

Yang, H., & Zhou, J. (1998). Optimal traffic counting locations for origin-destination matrix estimation. Transportation Research Part B, 32(2), pp. 109-126.

Yin, J., Tang, T., Yang, L., Gao, Z., & Ran, B., (2016). Energy-efficient metro train rescheduling with uncertain time-variant passenger demands: An approximate dynamic programming approach. Transportation Research Part B, 91, pp. 178-210.

Yin, J., Yang, L., Tang, T., Gao, Z., & Ran, B. (2017). Dynamic passenger demand oriented metro train scheduling with energy-efficiency and waiting time minimization: Mixed-integer linear programming approaches. Transportation Research Part B, 97, pp. 182-213.

Yu, V. F., & Lin, S. W. (2014). Multi-start simulated annealing heuristic for the location routing problem with simultaneous pickup and delivery. Applied Soft Computing, 24, pp. 284-290.

Yuan, J. (2006). Stochastic modelling of train delays and delay propagation in stations. Ph.D. dissertation, Delft University of Technology, The Netherlands.

Zhan, S., Kroon, L. G., Veelenturf, L. P. & Wagenaar, J. C. (2015). Real-time high-speed train rescheduling in case of a complete blockage. Transportation Research Part B, 78, pp. 182-201.

Zhan, S., Zhao, J. & Peng, Q. (2016). Real-time train rescheduling on high-speed railway under partial segment blockages. Journal of the China Railway Society, 38(10), pp. 1-13.

Zhang, T., Chaovalitwongse , W. A., & Zhang, Y. (2012). Scatter search for the stochastic travel-time vehicle routing problem with simultaneous pick-ups and deliveries. Computers & Operations Research, 39, pp. 2277-2290.

Zhang, Q., Han, B., & Li, D. (2008) Modeling and simulation of passenger alighting and boarding movement in Beijing metro stations. Transportation Research Part, C 16(5), pp. 635-649

Zhao, F., & Zeng, X. G. (2006). Simulated annealing genetic algorithm for transit network optimization, Computing in Civil Engineering, 20(1), pp. 57-68.

Zhao, J., & Rahbee, A. (2007). Estimating a rail passenger trip origin-destination matrix using automatic data collection systems. Computer-Aided Civil and Infrastructure Engineering, 22, pp. 376-387.

Zhao, N., Roberts, C., Hillmansen, S., & Nicholson, G. (2015). A multiple train trajectory optimization to minimize energy consumption and delay. IEEE Transactions on Intelligent Transportation Systems, 16 (5), pp. 2363-2372.

Zhu, Y., & Goverde R. M. P. (2017). Dynamic passenger assignment during Disruptions in Railway Systems. Proceedings of the 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems – IEEE MT-ITS 2017, Naples, Italy, June 2017, pp. 146-151.