



UNIVERSITÀ DEGLI STUDI DI TRIESTE

Sede Amministrativa del Dottorato di Ricerca

**XXII CICLO DELLA
SCUOLA DI DOTTORATO DI RICERCA IN
INGEGNERIA CIVILE ED AMBIENTALE**

CAPACITY AND RELIABILITY ON RAILWAY NETWORKS: A SIMULATIVE APPROACH

Settore scientifico-disciplinare ICAR/05

DOTTORANDO
dott. ing. GIORGIO MEDEOSSI

RESPONSABILE DOTTORATO di RICERCA
prof. ing. AURELIO MARCHIONNA
Università di Trieste

RELATORE e TUTORE
prof. ing. GIOVANNI LONGO
Università di Trieste

ANNO ACCADEMICO 2008/2009

Ben fastu e ben varastu...

Acknowledgements

When I was in the primary school, I was really interested in all transportation systems, but especially in agriculture tractors and trains. This hobby grew up with me, travelling to Switzerland by train with my family, reading magazines for truck drivers and watching tractors on the hills. I decided for trains when I was sixteen, and I started reading train timetables of Italy, with its international connections....

First of all, I wish to thank Prof. Giovanni Longo for the confidence placed in me and for giving me the opportunity to take part to an incredible number of very interesting projects that made this PhD a unique experience. I thank also prof Roberto Camus and prof Fabio Santorini for supporting me during these three years.

This dissertation would not have been possible, without a fruitful cooperation with RFI (Rete Ferroviaria Italiana). Special thanks to all RFI technicians and engineers with whom we cooperated and developed the applicative parts of the work.

I am indebted to Stefano de Fabris, who developed the data-analysis tool and improved it with any function I needed. I really appreciated the support of Stefano Strami, especially for reading and correcting this thesis. I would also thank Ing. Giuliano Stabon and Boris Sasic, for any enjoyable pause we spent together, Matteo Visintin for sharing my passion for Mac computers, Fabio Lamanna and Lorenzo Tomasella for tolerating my music and my way of working.

I am also grateful to Daniel Hürlimann, for solving any question concerning OpenTrack and endowing it with the interfaces I requested, and to all professors, researchers and PhD students I met in these three years and especially prof. Ulrich Weidmann, dr. Marco Lüthi and dr. Gabrio Caimi: the experience gained in Zurich was fundamental for this thesis.

Finally, let me thank my family in mother tongue:

O ringrazi di cûr i miei gjenitôrs, che àn simpri crodût in me, ancje cuant che no lu meritavi, e lis mes sûrs Teresa, simpri atenta e diponibila, e Anna, che mi a simpri sburtât a lâ un toc plui indevant. Grazie ai miei cusins, a lis nonis, agnis, al barba e ai amîs che mi permetin di no pensâ dome ai trenos.

Abstract

An efficient train operation is a primary success factor for all infrastructure managers, since it allows operating a higher number of trains without significant infrastructure investments.

As known, a trade-off exists between capacity and punctuality, forcing planners to find an equilibrium allowing the highest number of slots to be operated with satisfying punctuality indicators. This is particularly challenging in nodes, where the combination of different stochastic parameters on various lines and for different trains dramatically increases modelling tasks.

In the last years, railway simulators have become a very powerful instrument to support the different steps of the planning process: from the layout design to capacity investigations and offer model validations. More recently, the possibility of an automatic import of infrastructure layouts and timetables widened the application spectrum of micro-simulators to large nodes and to more detailed stochastic stability evaluations.

Stochastic micro-simulators can reproduce most processes involved in rail traffic and comprehend not only its deterministic aspects, but also human factors. This is particularly relevant in order to simulate traffic under realistic conditions, considering variability at border, various driving styles and stop times. All these parameters have to be calibrated using real-world collected data for single trains or train families, considering their different behaviour in the network and at its border.

Since a perfect representation of all stochastic and deterministic parameters involved in rail traffic is not possible, a calibrated model must be validated to evaluate its precision before using it in practice. Calibration has been tested on the Palermo - Punta Raisi single-track line, on the Trieste - Venice double-track line and in the node of Turin.

The model is first used to forecast reliability of the operations after infrastructure and timetable changes. Results have been compared ex-post with real traffic data, showing remarkable reliability. An approach is then presented, in which stochastic micro-simulation is used to represent the relationship between robustness, capacity and a number of other important factors, such

as traffic variability or running time supplements. The approach can be used to estimate the buffer times, and the running time supplements to obtain a given reliability level.

First, micro simulation with its advantages and weaknesses is presented; then, after a presentation of the most common reliability measures, the a new indicator is explained. Third, calibration, validation and application of the case studies is described; in the last part, an approach to evaluate the trade-off between different parameters is presented.

Sommario

La domanda di trasporto collettivo, in particolare ferroviario, è fortemente cresciuta nelle grandi aree urbane anche per effetto di specifici strumenti pianificatori volti a favorire l'utilizzo di alternative sostenibili sotto il profilo ambientale, nel rispetto della configurazione del territorio. La conseguente saturazione dei nodi e delle linee e la contemporanea necessità di aumentare la regolarità del servizio offerto impongono ai gestori un deciso aumento nella precisione della pianificazione dell'esercizio.

Come è noto, esiste un trade-off tra capacità e regolarità dei servizi ferroviari, che obbliga i pianificatori a trovare un equilibrio, massimizzando il numero di treni e garantendo nel contempo un soddisfacente livello di regolarità. Il mantenimento di tale equilibrio risulta particolarmente complesso nei nodi, dove la combinazione di fenomeni stocastici in numerose linee e per diversi servizi aumenta notevolmente le difficoltà di modellizzazione della circolazione reale.

Negli ultimi anni, gli strumenti di microsimulazione della circolazione ferroviaria sono divenuti uno strumento potente a supporto delle diverse fasi di pianificazione, dalla progettazione funzionale dell'infrastruttura alle stime di capacità e alla validazione degli orari. Più recentemente, la possibilità di ottenere l'importazione automatica del modello infrastrutturale e degli orari ha ampliato lo spettro di applicazione della microsimulazione ai nodi complessi ed alle valutazioni ex-ante della stabilità degli orari.

La microsimulazione consente di riprodurre la gran parte dei processi coinvolti nella circolazione ferroviaria, comprendendo non solo i suoi aspetti deterministici, ma anche il fattore umano. Ciò risulta particolarmente importante al fine di simulare la circolazione in condizioni reali, considerando i ritardi in ingresso nell'area di simulazione, i diversi stili di guida e la variabilità dei tempi di fermata nelle stazioni. Tali parametri necessitano tuttavia di un'adeguata calibrazione mediante dati reali per singoli servizi o famiglie di treni, considerando le diversità di comportamento all'interno della rete simulata ed al cordone.

Dato che non è possibile una rappresentazione perfetta di tutti i parametri stocastici e deterministici che definiscono la circolazione

ferroviaria, un modello calibrato deve essere validato per valutarne la precisione, prima di venire utilizzato in pratica. La calibrazione è stata testata sulla linea a semplice binario Palermo-Punta Raisi, sulla Venezia - Trieste, a doppio binario e nel nodo di Torino.

In primis, il modello è stato utilizzato per la stima ex-ante della regolarità in seguito a modifiche infrastrutturali e di orario. Dopo l'attivazione dei provvedimenti simulati, i dati della circolazione reale sono stati confrontati con quelli simulati, dimostrando la notevole attendibilità della stima.

E' stata quindi sviluppata una metodologia, in cui la microsimulazione viene utilizzata per rappresentare la relazione tra capacità, regolarità dei servizi ed una serie di altri fattori, quali ad esempio gli allungamenti sul tempo di percorrenza e i fenomeni stocastici. L'approccio può essere utilizzato per la stima delle riserve e degli allungamenti da inserire nell'orario per ottenere una data regolarità.

La tesi muove dalla presentazione della microsimulazione, delle sue potenzialità e dei suoi limiti; quindi, dopo la descrizione delle principali misure di regolarità, viene introdotto un nuovo indicatore. Seguono la presentazione delle metodologie di calibrazione e validazione di un modello di microsimulazione e la loro applicazione ai casi di studio. Viene infine presentata la metodologia per la stima coordinata di regolarità e capacità, le cui potenzialità sono illustrate nei casi di studio.

Contents

Introduction.....	15
1.1 Research objectives.....	17
1.2 Thesis outline.....	17
Timetabling and Delays.....	19
2.1 Basic principles for railway capacity	20
2.1.1 Minimum Headway time	20
2.1.2 Buffer Time	22
2.1.3 Running time supplements.....	23
2.1.5 Capacity, Stability and Robustness.....	27
2.2 Capacity, Robustness and Reliability: A Literature Review.....	28
2.2.1 UIC leaflet 406	28
2.2.2 Deterministic Models	30
2.2.3 Stochastic models.....	32
2.2.4 Real traffic analysis.....	36
2.2.5 Simulation.....	37
2.2.6 Conclusions.....	38
2.3 Reliability	39
A new reliability indicator	43
3.1 Case study.....	48
3.2 Reliability measures in the case studies.....	51
3.2.1 Reliability measure at the end of a trip.....	51
3.2.2 Global evaluation for a single-track line.....	52

3.2.3 Global evaluation for a node with mixed traffic	53
3.3 Conclusions	57
Calibration and validation of a micro-simulation model	59
4.1 Simulation Models.....	60
3.1.1 Macro-simulation models.....	61
4.1.2 Synchronous microscopic simulation models.....	62
4.1.3 Asynchronous simulation models.....	64
4.1.4 Asynchronous vs Synchronous simulation.....	65
4.2 Building a synchronous simulation model.....	66
4.2.1 Stochastic phenomena and secondary delays.....	67
4.3 Calibration of motion equation.....	70
4.3.1 Acceleration.....	71
4.3.2 Full-speed behaviour.....	72
4.3.3 Braking.....	72
4.3.4 "Global" calibration	73
4.4 Issues of synchronous simulation.....	76
4.4.1 Dispatching rules	76
4.4.2 Complex dispatching on lines.....	77
4.5 Calibration: Case Studies.....	81
4.5.1 Venice - Trieste	81
4.5.2 Palermo - Punta Raisi	84
The trade-off between capacity and reliability.....	93
5.1 Estimation of the trade-off between capacity and robustness.....	94
5.1.1 Creating a dense timetable.....	94
5.1.2 Buffer times	97
5.1.3 Distributed running time supplements.....	98
5.1.4 Concentrated running time supplements.....	100
5.1.5 Stop time supplements	101
5.1.6 Stochastic phenomena.....	105
5.2 Model application.....	110
5.3 Case Studies	111
5.3.1 Turin Node	112
5.3.2 Rome - Formia	112
5.3.3 Venice - Trieste	113
5.4 Main Results.....	114

5.4.1 Buffer time - Timetable Robustness	115
5.4.2 Running time supplements - Reliability	118
5.4.3 Stochastic delays - Reliability.....	121
5.4.4 Evaluation of Infrastructure improvements	122
5.5 Comparison with stochastic differential equations.....	124
Conclusions.....	127
6.1 Main Results.....	127
6.2 Recommendations and Future Research.....	129
Bibliography	131

*Noi canteremo [...] le stazioni ingorde,
divoratrici di serpi che fumano; [...] le
locomotive dall'ampio petto, che scalpitano
sulle rotaie, come enormi cavalli d'acciaio
imbrigliati di tubi [...]*

(F.T. Marinetti, Manifesto del Futurismo, 1909)



Introduction

Rail travel demand has grown strongly in large urban areas and growth in rail service is expected to continue in the coming years, since the mobility demand would generally increase and rail services may represent a very good alternative to the private car or bus transit in large towns. However, especially in densely inhabited areas, laying new tracks to meet the growing demand requires long planning and construction times at very high costs. Moreover, the economic crisis of 2008 has forced operators to reduce costs, which are substantially paid by states and regions.

As a result the investments and the operational costs have to be reduced, but also the quality levels should be increased. The goal for the Infrastructure Managers is therefore to maximise the use of the existing infrastructures, eventually improving them with the most advanced IT technologies. But technological improvements are non the less still insufficient to increase usable capacity significantly. An efficient train operation is a primary success factor for all infrastructure managers, since it allows operating a higher number of trains without significant infrastructure investments.

A key success factor is an efficient train operation, since it does not represent a cost, but allows operating a higher number of trains. As known, a trade-off exists between capacity and punctuality, forcing planners to find an equilibrium allowing the highest number of slots to be operated with satisfying punctuality indicators.

This is particularly challenging in nodes, where the combination of different stochastic parameters on various lines and for different trains dramatically increases modelling tasks.

In the last years, a number of tools have been introduced to support timetable planning, offering a precise representation of the occupation of each train on a line and eventually the possibility to simulate traffic in high detail. On the other hand, various studies have investigated the possibility to create an optimal timetable using operations research algorithms; some of these have also been developed to complete software tools and are currently used to support timetable planning. Much more based on the experience of planners is the definition of the measures, which have to be inserted in a timetable to allow trains to recover delays and to prevent delay propagation. These are based on “rules of thumb” of each Infrastructure Manager, and therefore normally different in each country.

The first element to evaluate such delay prevention and compensation measures is a deep knowledge of the real traffic, with standardised, intuitive and simple methods. The conventional performance measures present a number of weaknesses, which either lead to time-consuming analysis or do not offer a precise representation of real phenomena.

The results of a detailed traffic analysis can be used as input for an accurate model of real operations. Among other approaches, stochastic micro-simulation represents one of the most precise ways to model train operations on a network, obtaining knock-on delay and punctuality estimations and allowing users to evaluate various rolling stock, infrastructure layouts and timetable.

Although already widely used since some years, especially for infrastructure planning tasks, very little literature can be found regarding the use of stochastic micro-simulation to estimate the relationship between capacity and reliability. Another blank spot in the existing literature is the calibration and validation of such models, necessary since an exact representation of all parameters and of the dispatchers’ behaviour is not provided at the moment. This residual error must be estimated to enable a precision evaluation of the models.

1.1 Research objectives

The thesis aims mainly at modelling train traffic on large networks to support the improvement of capacity utilisation and timetable design at a desired reliability level. Aiming at this final goal, a number of objectives are derived:

1. Getting more insight into the stochastic characteristics of train movements in Italy, on their relationship with different infrastructures and timetables.
2. Defining an innovative way to evaluate service reliability, focussing on the point of view of the Infrastructure Managers.
3. Developing a calibration procedure of a synchronous micro-simulation model, which enables detailed representation of very different situations.
4. Estimating the precision of a calibrated micro-simulation models in a number of case studies.
5. Developing an approach to the representation of the relationship between network capacity and reliability, including the impact of a number of factors, and testing it in the case studies.

1.2 Thesis outline

The thesis is structured in four parts, illustrated in Figure 1.1. First, an introduction to the problem is described in this chapter 1.

Chapter 2 gives first an overview of the railway timetabling issues concerning the possibility to compensate disturbances. The chapter continues with a literature review covering the most important approaches to railway capacity, timetabling, and to the ex-ante evaluation of timetable reliability. In the last part, the most common indicators to measure service reliability are explained and discussed.

In Chapter 3, a new service reliability indicator which allows a synthetic and simple but precise service quality estimation is introduced. The same parameter can be used to evaluate the robustness of simulated train operations and compare results of different scenarios. Then the indicator is tested on a single-track suburban line in Palermo and in the node of Turin.

Chapter IV begins with the description of the different railway simulation methods and tools; the most common approaches to micro-simulation are then compared. Some issues concerning the construction of a synchronous micro-simulation model are then presented, together with the weaknesses of the approach. A simple method to calibrate and validate a micro-simulation model is then explained and tested in three case studies. In Chapter III and IV, the inputs for the chapter V are defined. Therefore in Figure 1.1 they are parallel and their combination leads to Chapter V

In Chapter V, an approach is described, which allows to use the calibrated model used to measure the variation of service reliability under different conditions. In particular, the impact of an increasing number of services and of growing running time supplements is tested, to obtain a representation of the trade off between capacity and reliability. The approach is tested in two different case studies.

Finally, Conclusions are drawn in chapter 6, together with some recommendations for further development.

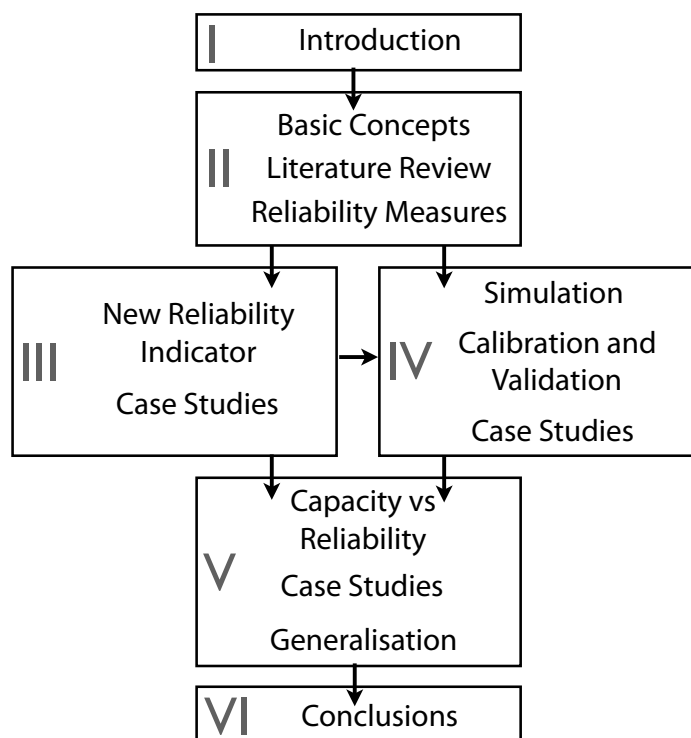
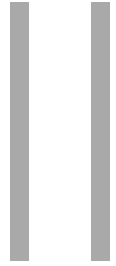


Figure 1.1: Thesis structure.

*Un bello e orribile mostro si sferra,
corre gli oceani, corre la terra:
Corusco e fumido come i vulcani,
i monti supera, divora i piani
Sorvola i baratri; poi si nasconde
per antri incogniti, per vie profonde;
Ed esce; e indomito di lido in lido
come di turbine manda il suo grido [...]
Passa benefico di loco in loco
Su l'infrenabile carro del foco. [...]*

(G.Carducci, Inno a Satana, 1863)



Timetabling and Delays

The very high cost of the railway infrastructure, and its relatively low capacity compared to other transportation systems require a careful planning for both investments and offer, to allow the maximum infrastructure utilisation and high reliability standards. In particular, on a given infrastructure, it is necessary to consider and combine all deterministic and stochastic aspects of train movements in order to obtain a realistic representation of reality, which is a precondition for an efficient timetable planning.

This chapter starts with an introduction to capacity planning and timetabling process, with a focus on practice in Italy. A brief literature review describes the most recent achievements in capacity and robustness estimation and in the analysis of real traffic data. Finally, the most used reliability measures are presented.

This brief introduction to timetabling and traffic analysis is intended to give the reader some key elements useful to better understand the next parts and not as a complete description of the problem and its issues.

2.1 Basic principles for railway capacity

Once a railway line has been built and the interlocking system has been installed, its capacity is nearly fixed until expensive technological or civil works improve its performance. As a result, the very first issue in timetable planning is a detailed capacity modelling. Once the minimum headway time between trains has been calculated, it has to be increased to feasible values by the introduction of margins. Margins, running time and headway time are the basic elements to plan a timetable.

2.1.1 Minimum Headway time

The *maximum capacity* of a line section is the maximum number of trains that can be carried under ideal (deterministic) conditions. In this case, the scheduled headway between two trains is the *minimum headway*. The minimum headway depends on the *blocking times* (Pachl, 2002) of any scheduled pair of trains. The blocking time is the time interval in which a section of a track (usually a block section) is exclusively allocated to a train at its scheduled speed and blocked for any other train. Thus, the blocking time lasts from issuing a movement authority (i.e. by clearing a block signal) until the possibility of issuing a movement authority to another train to enter the same section. The blocking time of a signal block is usually much longer than the time the train occupies the block physically. In a classic line with two-aspect signalling, the blocking time of a block section consists of the following time intervals (Figure 2.1):

- time for clearing the signal
- a certain time for the driver to view the clear aspect of the signal that gives the approach indication indication to the signal at the entrance of the block section (this can be the preceding block signal or a separate approach signal),
- the running time between the block signals,
- the clearing time to clear the block section and, if requires, the overlap with the full length of the train,
- the release time to “unlock” the blocking system.

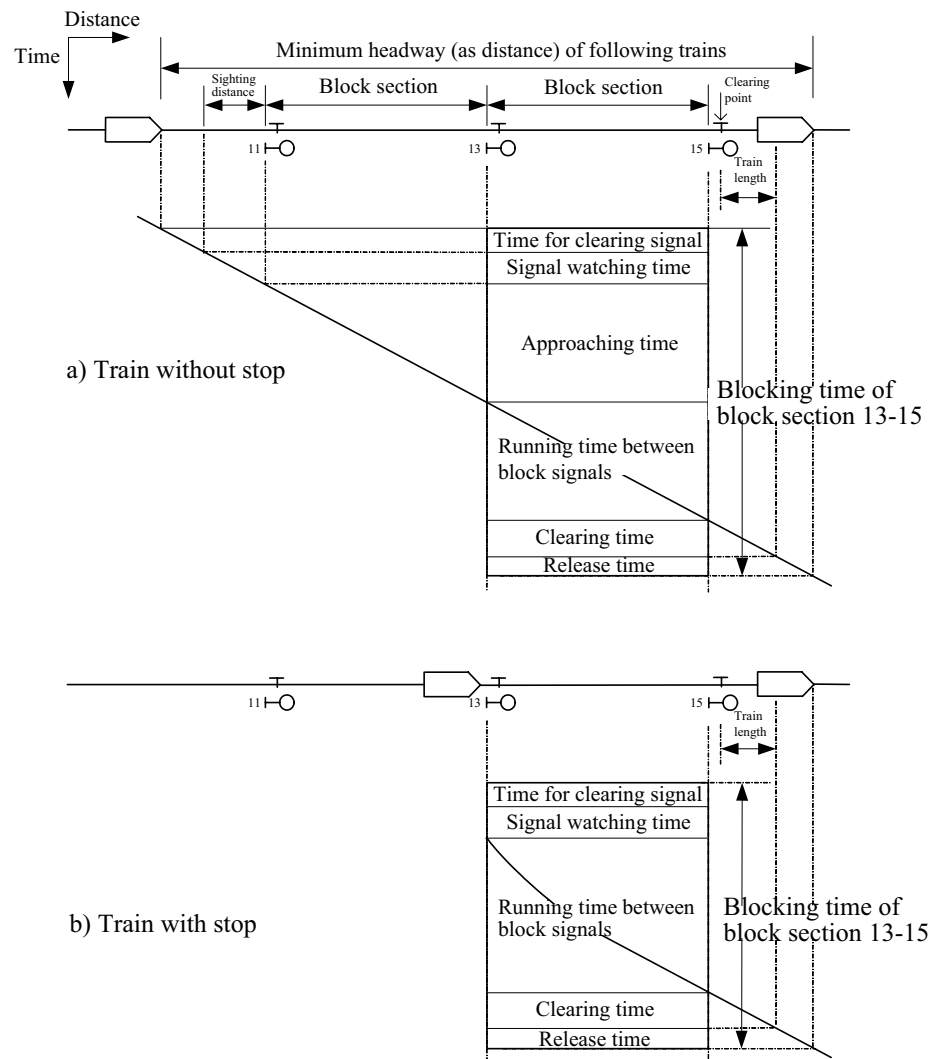


Figure 2.1: Blocking time of a block section (Source: Pacht, 2002)

Drawing the blocking times of all block sections that a train passes into a time-distance diagram yields the so-called *blocking time stairway* (Figure 2.2). The blocking times directly establish the signal headway as the time interval between two successive trains in each block section. The minimum line headway is the headway between two trains not only considering one block section but the whole blocking time stairways of the line. In this case the blocking time stairways of two trains just touch each other in at least one block section, which is the critical block section. In this case any delay of the first train in this section leads to propagation to the second train, since no free time is provided to compensate variation.

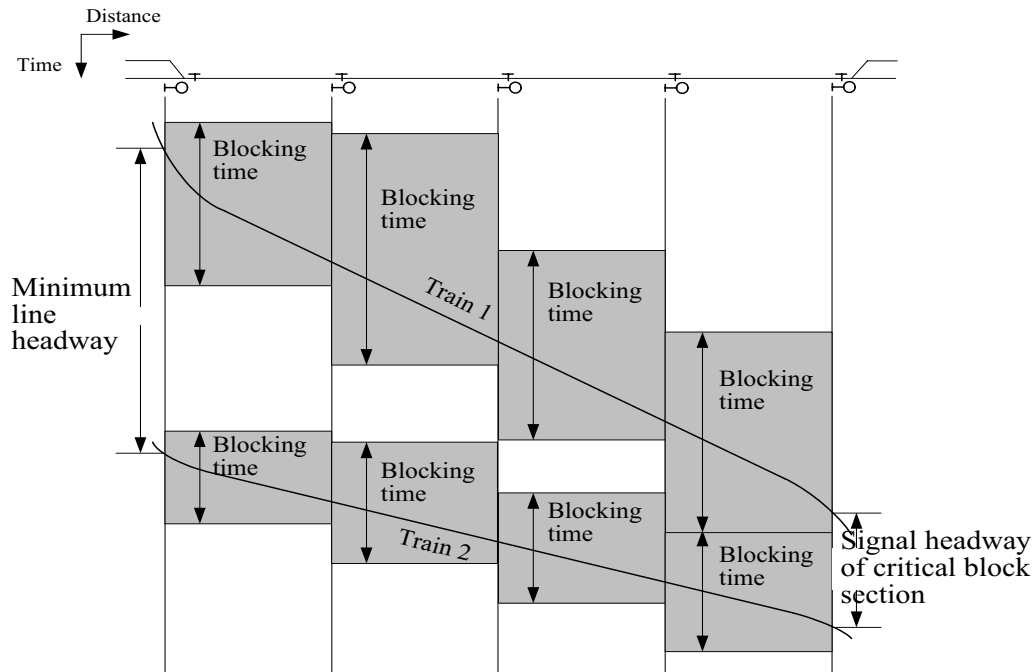


Figure 2.2: Blocking time stairways, signal headways and minimum line headway (Source: Pacht, 2002)

2.1.2 Buffer Time

To reduce the possibility of delay propagation between trains, buffer times are added to minimum headway times, increasing them and therefore reducing available capacity. As a result, while buffer times are required to obtain high traffic reliability, they must be as short as possible to keep high capacity levels. The importance of buffer times and the relationship with traffic variability is clearly shown by (Hansen, 2005) in Figure 2.3. The distance between the two dashed lines represents the buffer time, while the curve represent the real release variability distribution of the 1st train (left) and the real start of blocking time of the same section for the 2nd train. The intersection area of these curves represents the probability for the second train to see a yellow signal and therefore the probability of a conflict, which leads to higher block section occupation and delay. When a short buffer time is used, the probability of conflict is higher, while it becomes lower increasing it. Since the area of the curve is a function of the trains running variability, also the buffer time to obtain a low conflict probability must be evaluated considering the expected distributions.

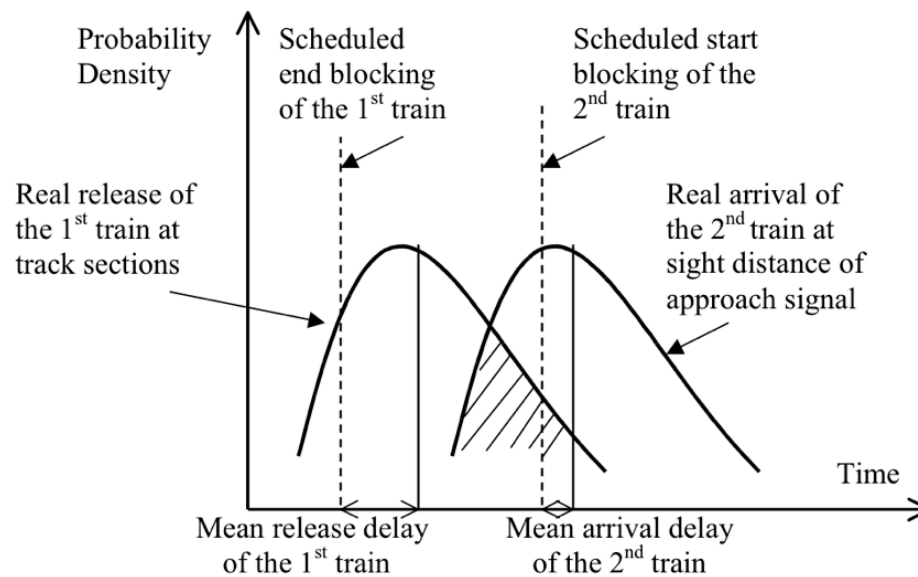


Figure 2.3: Relationship between buffer time and conflicts probability (Source: Hansen, 2005)

Despite the key role played by buffer times in timetable planning, a number of rules of thumb are still used in practice by most operators, probably due to the high reliability and traffic density already obtained in many countries or to the dependency of any model from the expected traffic variability, which may be different when the timetable is operating.

2.1.3 Running time supplements

To obtain a high reliability of train services, it is necessary to consider not only the minimum running time possible with a given trainset, but also to compensate the normal stochastic phenomena that normally influence train running without generating delays. This means that scheduled running times are the sum of the minimal running times and *running time supplements*.

Higher running times generally lead to a better reliability of train services. However, higher supplements also lead to higher planned running times, reducing the attractiveness of train for customers. Moreover, trains running slower increase the infrastructure occupation, reducing available capacity; on the other side trains running at maximum performance despite supplements can enter earlier than planned in critical sections (i.e. nodes), being forced to brake and re-accelerate and therefore increasing congestion in such sections.

The International Union of Railways (UIC, 2000) has published Leaflet 451-1 on the size of running time supplements. In their recommendations, the supplements are the sum of a distance-dependent supplement and a percentage of the technically minimal running time. The distance-dependent supplement is 1.5 min/100 km for locomotive-hauled passenger trains and 1 min/100 km for multiple-unit passenger trains. The running-time-dependent supplements vary between 3% for stopping trains and 7% for fast trains. For locomotive hauled trains, the percentage also depends on the total weight of the train. Supplements for cargo trains are generally higher. Furthermore, the running-time-dependent supplement can be replaced by a second distance-dependent supplement in some cases.

While the mentioned recommendations are given as example for a generic railway operator, slightly different strategies and rules have been developed in each country about dimensioning and allocating supplements. In most cases these are based on the experience acquired by planners depending on the specific conditions, and are not supported by theoretical models. In a generic way, we can consider:

a) *Distributed supplements*, as previously described regarding the UIC Leaflet 451-1, these supplements are constant and distributed through along the train route, as distance- or running time-dependent percentage. Such supplements can compensate the different behaviour of train drivers, small departure delays, adherence reduction and other variability causes.

b) *Concentrated supplements*, which allow trains to enter punctual into nodes or critical stations. Concentrated supplements allow relatively high delays to be generated in the open track compensating them before or at critical points, obtaining higher punctuality rates, especially if it is only measured in such points. Concentrated supplements are very efficient to reduce arrival delays.

c) *Stop-time supplements* compensate both arrival delay and dwell time variability: they allow punctual departure but cause higher capacity consumption at platforms. If placed just before critical points in nodes, they reduce the probability of secondary delays and therefore a reduction of buffer time in such sections.

2.1.4 “Total Headway Time”

The timetabling software ROMAN currently used by RFI in Italy is endowed with a running time calculator, which calculates the running time of a specific trainset for a given service. The planner adds to this minimum running time the distributed,

concentrated and stop-time supplements, respecting the recommendations contained in an internal leaflet called "*Scenario Tecnico*" ("Technical Scenario") which contains the performances for each track section of the network.

This leaflet contains not only the recommended supplements, but also the headway times, comprehending the technical minimal headway times and the buffer times evaluated by the experience of engineers. This "total" headway times are called "*Specifiche Tecniche di Circolazione*" (Technical Traffic Specifics). The leaflet is published yearly, and contains headway times for following trains, crossing itineraries in stations and at junctions, in event of take-overs and regarding the use of two-way working.

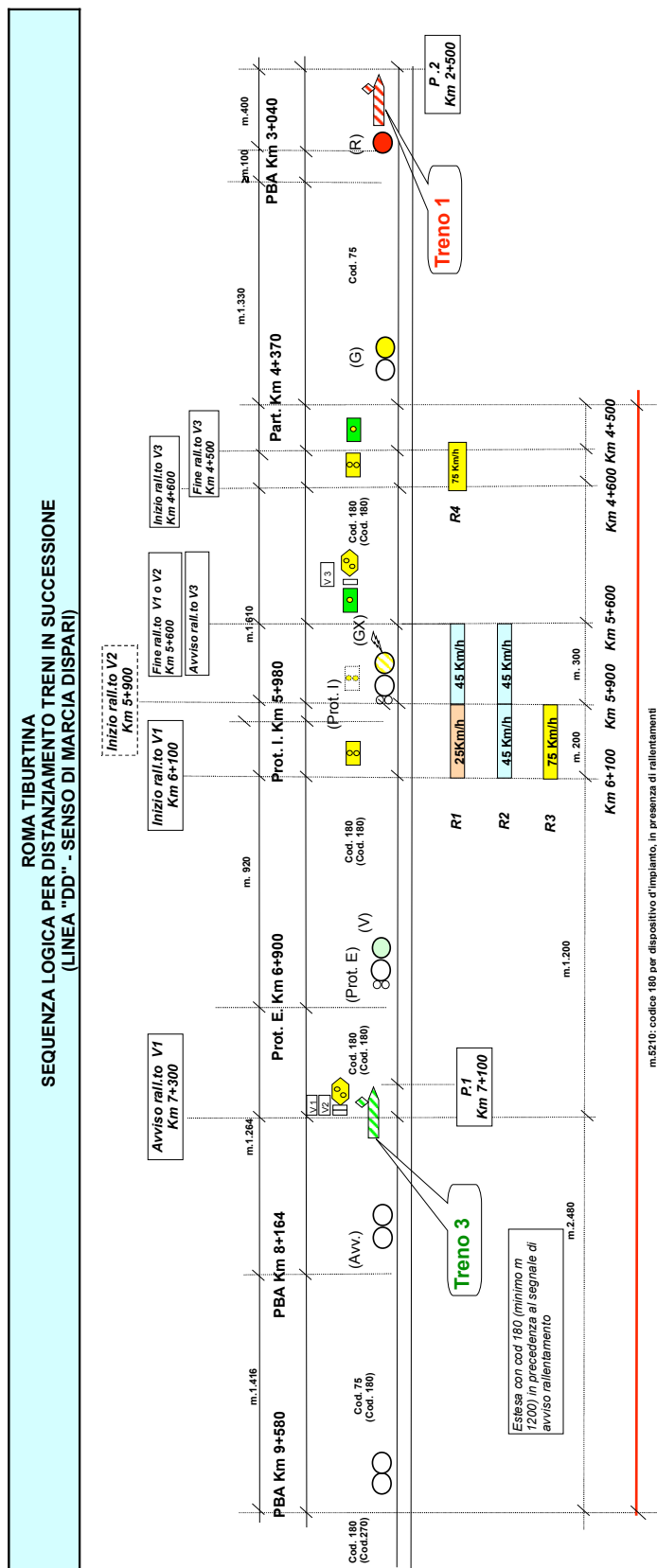
For example, on the Turin - Milan conventional line, with about 1600-meter-long block sections, and a line speed around 160 km/h and continuous ATP, for each section and station (example translated by the Author):

Headway: Normal: 6'; 5' only for two following trains. If this minimum is used, for a third train the "normal" has to be increased by the difference between the "normal" and the "minimal"

Headway for crossing itineraries in a station: 5' for trains ending/starting in the station, 4' elsewhere.

Headway in event of overtakings: 4' before the takeover, 3' after the train has passed

The leaflet also contains indications regarding the impact of works on the infrastructure, regarding both lines (speed restrictions, temporary single-track,...) and stations (unavailable platforms, tracks, restrictions...). An example of such impact calculation is presented in Figure 2.4. It regards the impact evaluation of major improvement works near Roma Tiburtina station, in presence of the continuous analogic and multi-aspect ATP system ("RSC"); the title of the picture is "Headway sequence of trains - High Speed Line, Odd trains". The signal aspects between two following trains are shown, which help headway time calculation with the combination of multi-aspect signalling with fixed speeds and line speed reductions.



Nota 1) Il valore di codice indicato tra parentesi (180) rappresenta il massimo codice di linea disponibile al treno per dispositivo di impianto, in presenza di rallentamento.

In presenza di un abbattimento codice (da "270" a "180") correlato ai rallentamenti per la realizzazione della NCI, abbattimento esteso dal km 9+580 al km 4+370, sono stati determinati i peritempo di un treno con $V=105$ km/h nella tratta con abbattimento codice, in presenza dei seguenti rallentamenti:

- 25 km/h dal km 6+100 al 5+900 e 45 km/h dal 5+900 al 5+600 (R1), in presenza, o meno, di un rallentamento a 75 km/h dal km 4+600 al km 4+500 (R4) – caso utilizzo fasci di rotale senza deroga;
- 45 km/h dal km 6+100 al km 5+600 (R2), in presenza, o meno, di un rallentamento a 75 km/h dal km 4+600 al km 4+500 (R4) – caso utilizzo fasci di rotale con deroga;
- 75 km/h dal km 6+100 al km 5+900 (R3), in presenza, o meno, di un rallentamento a 75 km/h dal km 4+600 al km 4+500 (R4) – caso utilizzo sistema Essen.

Inoltre, è stata determinata la percorrenza di un treno, in presenza delle combinazioni di rallentamenti e di abbattimento codice, tra le progressive 7+100 e 2+500, senza limitazioni circa la disposizione a via libera dei segnali (segnali ai km 6+900, 5+980, 4+370, 3+040 superati a via libera senza condizionamenti).

Figure 2.4: Headway time calculation in presence of speed restrictions. (Source: RFI)

2.1.5 Capacity, Stability and Robustness

The capacity of a railway network does not exist as such: it depends on the way the infrastructure is utilised. Therefore, a reliable capacity evaluation has to be based on a pre-defined timetable for train operation on the given infrastructure. Depending on the buffer times between slots with a given timetable different punctuality rates can be obtained, which are used to measure the quality indicators for Infrastructure Managers.

As presented in (UIC, 2004), on a given infrastructure, capacity is based on the interdependencies existing between:

- the **number of trains**. When train intensity increases, less capacity is left for buffers and margins.
- the **average speed**. Faster services require braking distances, and therefore headways, which grow proportionally more than the average speed.
- the **stability**. As presented in 2.1.2/3 margins and buffers have to be added to the running time of trains and between train paths to ensure that minor delays are suppressed instead of amplifying and so causing (longer) delays to other trains. *Stability* is defined as the ability of a system to compensate for delays and return to its initial state. The word **robustness** is often used instead of stability: *Robustness* is the ability of a system to withstand model errors, parameters variations or changes in the operational conditions.
- the **heterogeneity**. The capacity consumption of the same number of trains will increase proportionately to the average speed difference between slow and fast services.

The relation between these parameters is clearly shown in the "capacity balance", as illustrated in Fig. 2.5. In this qualitative model, an axis for each parameter is drawn from a unique origin. A chord links the points on the axes, corresponding to the value of each parameter. The length of the chord represents the capacity. Capacity utilisation is defined by the positions of the chord on the four axes. Increasing capacity means increasing the length of the chord.

Since the average speed on a line is given and the heterogeneity of services depends on demand and cannot be decided by Infrastructure Managers, these two elements can be considered as nearly fixed. Capacity allocation and timetable

planning can be viewed as a search for an equilibrium in the trade-off between capacity (number of trains) and stability (traffic quality).

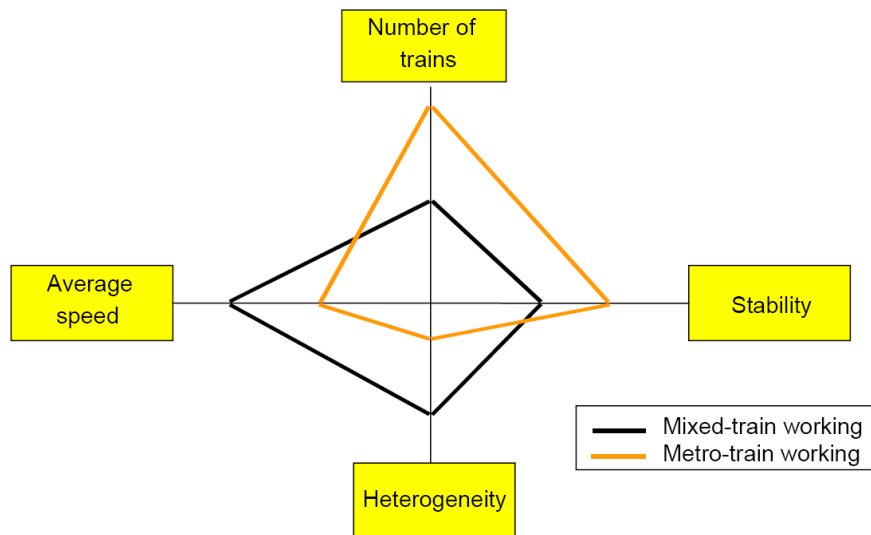


Figure 2.5: Capacity balance. (Source: UIC, 2004)

2.2 Capacity, Robustness and Reliability: A Literature Review

In the past decades, a range of researchers from different fields has studied railway capacity, timetabling and real traffic obtaining relevant results. Capacity studies aim at evaluating the maximum number of trains that can be carried on a network but, since effective capacity is strictly related to traffic variability (2.1.2) and to a given timetable structure, most approaches focus on a combination of such problems. As commonly presented in literature, we will distinguish timetable optimisation, timetable quality assessment and simulation models. A very complete description of the most important works is presented by Hansen and Pachl (2008).

2.2.1 UIC leaflet 406

The UIC 406 leaflet (UIC, 2004) describes a simple method to be used to evaluate the capacity utilisation on lines and nodes with a given timetable. It is performed in simple steps:

1) First a timetable for a large network, comprehending the line section to be analysed, has to be defined. As a result, the timetable in the analysis area depends on the infrastructure and timetable outside the analysis area (the so-called network effects).

2) Timetable graphs on the focus section are compressed, without considering possible conflicts on the other parts of the network and therefore deleting the possible network effects. Compression means that no buffer time is left between the occupation steps on the timetable graph, without changing the train order, running times and scheduled overtakings.

3) The value of the infrastructure occupation [% of time-window] is used to calculate capacity consumption (Fig 2.6)

4) New slots are added, if possible, until no more trains can be incorporated into the timetable, or the infrastructure occupation reaches congestion. The new slots form the *usable capacity*, the leftover capacity is the *lost capacity*.

If the *buffer times* already incorporated in the timetable (in light blue) are sufficient for timetable stability, the capacity consumption can theoretically reach 100% of the time-window considered. If it is greater, the infrastructure is congested.

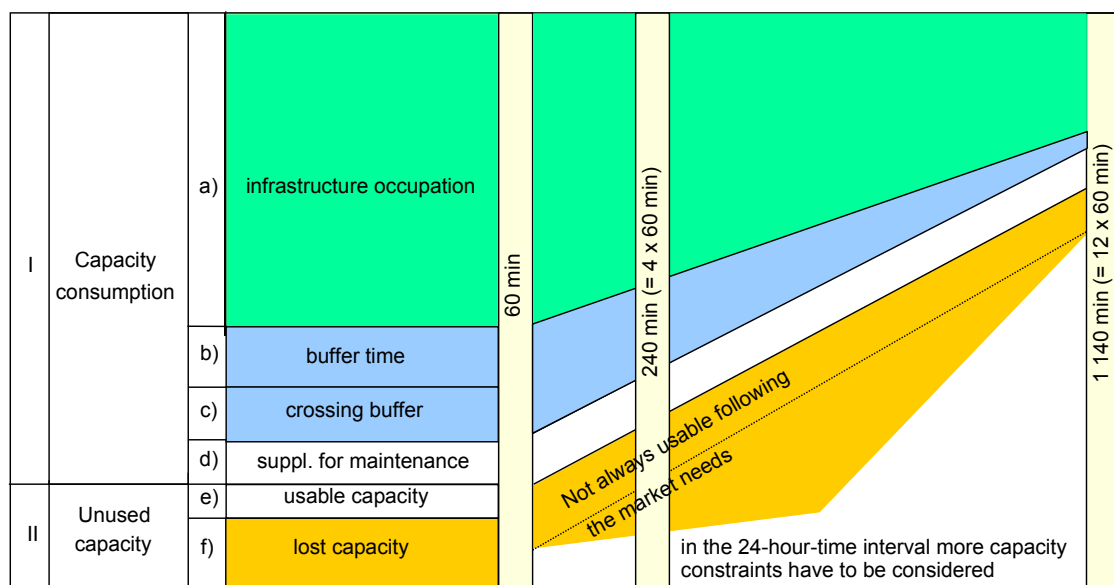


Figure 2.6: Determination of capacity consumption. (Source: UIC, 2004)

2.2.2 Deterministic Models

Timetable Optimization

Deterministic running, dwell and headway times are used all over the world for conventional timetabling, often leading to high-quality results thank to experience. Operations research algorithms, using such deterministic inputs can be used to solve a very wide range of problems, from timetabling to capacity assessment and to routing.

Serafini and Ukovich (1989) developed a mathematical model for the PESP (Periodic Event Scheduled Problem). Using this model in presence of timetabling restrictions, a feasible regular-interval timetable can be found. This model was first applied by Schrijver and Steenbeck (1994) obtaining the semi-automatic timetabling system CADANS, which presents a feasible solution respecting the given restrictions. The model has been integrated in the Dutch railway timetable development system DONS (Design Of Network Schedules). Since CADANS considers only a macroscopic infrastructure model, the real track and signal layout is not considered, requiring a more detailed feasibility check. This gap was filled by Zwaneveld (2001) developing an algorithm for solving the routing problem in a station to optimality using a branch-and-cut algorithm. The algorithm has been implemented in the module STATIONS as a part of DONS.

Liebchen (2003) extends the PESP model with symmetry constraints. Although this often leads to suboptimal solutions, these constraints can speed up the process of finding a good solution considerably.

Caimi et al. (2009) have solved the timetabling problem also for partially periodic timetables, which used on mixed-traffic lines in most countries. Introducing the Periodic Service Intention, a framework where the customer-relevant information about train services can be described, the problem becomes the search for a timetable that fulfils the requirements specified in the Periodic Service Intention. Results for a test case in the central part of Switzerland show that solutions are found with slightly longer computation times compared to fully periodic timetables, but obtaining a timetable that is usable in more real contexts.

Several other models exist to solve the timetabling problem, which are not mentioned here; a complete overview can be found in (Hansen & Pacht, 2008).

Routing and scheduling

To consider also the real topology of each station, a number of integrated scheduling and routing models have been developed. Carey (1994a) considers a rail corridor, assuming unlimited capacity in the stations; more recently, Carey and Carville (2003) use a heuristic approach which focuses for scheduling and routing in a single complex station; the model has been extended to a line or network by Carey and Crawford (2005). This approach can be extended considering a range of objectives, which are more understandable to planners.

Using very detailed topology, Herrmann (2005) proposed a two-level approach to saturate a station starting from a given line plan and maintaining high robustness. On the first step the draft timetable is generated using an aggregated topology, considering the smallest possible periodicity respecting safety restrictions. A detailed topology is used in the second step to decide feasibility of the previously generated timetables and analyse the derived schedules. Stability measures as properties of the used graph are used as input for optimisation problem solved with heuristics. Results on the Bern station region show that the tighter the timetable becomes the more effective is the improvement of the available capacity.

Caimi et al. (2005) consider the problem of generating robust train routings through a station, given a timetable and a layout of the station. Robustness is defined as the length of the time-slot which can be allocated to each train, which corresponds to the variability each train can have without causing secondary delays. The model has been tested on the Bern station region, where initial solutions are computed within minutes.

Timetable quality assessment

A slightly different approach to capacity assessment is used in the deterministic analytical model CAPRES (CAPacité de Reseaux ferroviaries) (Curchod and Lucchini, 2001). Capacity of a basic periodic timetable is analysed by saturating the network with additional trains. Running and dwell times are specified by a minimum and maximum value. The user defines the train services that are inserted in the timetable. The tool is useful in assessing unused capacity and detecting bottlenecks especially to plan infrastructure improvements.

A number of research studies in The Netherlands has led to the development of the software tool PETER (Performance Evaluation of Timed Events in Railways) (Goverde, 2005) based on Max-Plus-Algebra which performs efficient and reliable

timetable robustness analysis also on large-scale networks. The models are suitable for evaluating the overall stability of timetables of interconnected lines, but cannot be used to estimate the distributions of knock-on delays and the punctuality level of the scheduled trains, as they are still based on a deterministic modelling approach.

As a part of the RECIFE project, which aims at developing a software for capacity evaluations at stations, Delorme et al (2007) proposed a model to evaluate delay propagation using a shortest path problem resolution. The software suite also comprehends a capacity assessment tool based on a multi-objective combinatorial set packing problem (SPP) which saturates the timetable of the study area respecting infrastructure and user-defined constraints. The model uses a microscopic infrastructure model and headway times imported by the SISYFE simulator. The algorithm has been tested at a railway junction and on the station of Lille Flandres in France.

Routing and scheduling of trains on lines and stations have been intensively investigated, and a high number of different models for the assessment of capacity have been developed. Some timetabling models are also already in use or have been tested under real conditions. The main weakness of deterministic models is represented by their unsuitability to cope with stochastic process times and the need for pre-calculated running times, buffer and running time supplements.

2.2.3 Stochastic models

Timetables are conventionally developed on the basis of deterministic running and dwell times. However, the real traffic is conditioned by a number of stochastic phenomena, both in stations and in the open track. Since optimisation models do not consider this variability, stochastic approaches have been developed to evaluate buffer times and timetable slack to be used as input in timetable planning. The analytic stochastic models can be divided into queueing and delay propagation models.

Queueing models

Queueing models are used to estimate *waiting times* on networks. *Scheduled waiting time* is the time lost in the timetable due to infrastructure restrictions. Because of conflicts running time, dwell time or transfer time may be forced to be longer than the minimum process time. This additional time is called *scheduled waiting time*. The amount of waiting gives a prediction of bottlenecks and of buffer times which have to be inserted in timetable planning with corresponding traffic

intensity. Queueing models are mainly applied in strategic capacity assessment studies since they require no timetable.

Queueing models have been developed at the RWTH Aachen since 1974. Schwanhauser (1974, 1994) first developed models to assess the capacity in terms of train mix, frequency of different train types and mean inter-arrival times. Then he obtained analytical solutions to evaluate the required buffer time on lines maintaining acceptable total delay levels. The approach was first implemented in the mid 1980s in the software tool STRLE (Streckenleistungsfähigkeit).

The approach was extended by Wakob (1985) considering also stations. The model is based on the decomposition of station topology in route sections, called TFK (TeilFahrstrassenKnoten) which are used by only one train at one time as a part of an entire route. A series of TFK forms a route, while each TFK can be part of many different routes, depending on the topology. Each TFK is modelled as independent, single-server queue with a service time corresponding to the minimum headway time. Trains have gamma-distributed inter-arrival times, random entry order and different priorities. The approach has been implemented in the software tool ALFA (Analytische Leistungsfähigkeitermittlung der FAhrstrassenknoten) and more recently in the tool ANKE (Analytische Kapazitätsermittlung) (Vakhtel 2002), which can use the standard infrastructure model and database SPURPLAN (Bruenger, 1995, see also Paragraph 2.3). The model has been refined by Wendler (1999) considering conflicts between train triples and an improved algorithm. Since these methods assume random train orders, they are not valid for delay propagation in periodic timetables; moreover, these models are timetable free, i.e. only train types and frequencies are given as input. Therefore these models are very useful for capacity or buffer time estimation but cannot predict punctuality for a given timetable.

Schwanhauser's approach is used by Meng (1991) to calculate the required buffer time to preserve train connections at stations and evaluate delay propagation. Buffer time is calculated for all stations in the network with respect to all future timetables and combinations of train connections.

A slightly different approach is presented by Huisman et al. (2002) using a queueing network model suitable for long-term large network capacity investigations with a simple infrastructure model. The model divides the network into stations, junctions and (only double-track) line sections. By a careful definition of those components, the network is transformed in a product form queueing network, which allows a detailed analysis of these separate components. Close form expressions for

average waiting times are obtained and different network designs, traffic scenarios and capacity expansions may be evaluated. The simple infrastructure modelling is based on some assumptions: first, every train may use any platform in stations; second, trains arrive according to a Poisson process, while the occupation times and minimum headway times are exponentially distributed. Third, artificial queues are inserted between the components of the network with unlimited capacities, which is not accurate for track sections with limited capacity.

Delay propagation models

Delay propagation models have been investigated at a larger number of universities with different approaches. Using exponential delay distributions characterised by the mean delay and the delay probability Weigand (1981) developed a delay propagation model particularly suitable to analyse delay propagation due to connections at stations. Weigand also demonstrated that on cyclic timetables traffic remains stable if the average buffer time exceeds the average primary delay on each circuit in the network. This model was generalised by Muehlhans using general probability distributions for the primary delays. Cumulative Distribution Functions are derived for the evolution of delays. Both approaches consider only secondary delays due to connections at stations, so they cannot be used to explicitly represent delay propagation due to traffic conflicts.

Carey and Kwiecinski (1994) developed a stochastic model to simulate the knock-on delay occurring on a single track due to traffic conflicts, also considering speed variations. Carey (1999) presented also ex-ante stability measures for public transport that can be used for estimating the effects of perturbations and compare different timetables. The proposed heuristics can be used to recursively generate more robust timetables, using a stochastic delay propagation model to distribute time margins optimally.

Considering real traffic data, Kaminsky developed an empirical approach to assess the optimal headway between two trains. He fixes a threshold, the 80% of trains, which does not propagate delay to a following train. Kaminsky determines the blocking times for each block in the line finding the critical block considering minimum running times. The difference between the margins calculated using the 80% threshold and the planned ones gives the buffer time between two trains. The model considers a detailed infrastructure model, but not the delay and running time variability of the second train. The timetable planner has to adapt iteratively the

timetable to the calculated buffer times respecting the infrastructure constraints and timetable structure.

Vromans (2005) presents a stochastic model to improve the robustness of timetables by optimising the distributions of running and dwell time supplements to minimise delays under given stochastic disturbances. The total running and dwell times are fixed in advance using conventional rules and realistic disturbances are simulated. The model is based on deterministic headway times, thus neglecting the impact of traffic conflicts on train speed and track occupations. The model is quite powerful and has been used to test and improve 2008 timetable in The Netherlands.

Yuan (2006) developed a delay propagation model based on blocking time theory to estimate the knock-on delays caused by route conflicts and late transfer connections including the impact on the punctuality of trains. The model reflects the constraints of the signalling and interlocking system, the train protection rules and the impact of rescheduling on real operations. Starting from real distributions, the model enables accurate predictions of the knock-on delays of trains suffered at critical track sections and of the resulting punctuality of train arrivals and departures at the station.

Using stochastic differential equations, Stok (2008) presented a model for line capacity assessment. The model uses MonteCarlo simulations in which stochastic differential equations representing train motion are solved, obtaining a conflict probability between two following trains. Running time variability is given by a Brownian Motion component and a stochastic optimal control function is integrated to model drivers behaviour, which tries to minimise power consumption when running on time and to minimise delays when they occur. Although it can be used also for complex networks, the model has been tested only on simple lines.

Conte (2008) presented a stochastic approach to identify delay propagation using a graphical model, called tri-graph, that is a representation of probability distributions. The method has the advantage of being able to identify even complex dependencies without “a priori” knowledge of the track topology and of the operational rules of the interlocking system. Until now the model has only been tested as delay management support in a part of the German networks.

On the basis of extensive real world delay data, Flier et al. (2009) propose to predict the risk (conflict probability) of planned trains paths using linear regression models. In a shortest path model, regression models to evaluate the risk of entire paths through the corridor are combined. A Pareto frontier of solutions is computed

to consider the trade-off between risk and travel time. The approach is very efficient to add slots to existing timetables of complex networks; since only a macroscopic infrastructure model is used, a more detailed planning of the obtained paths including track and platform routing has to be performed.

2.2.4 Real traffic analysis

Since introduction of automatic train describers, operation data are automatically collected and stored by infrastructure operators in large databases initially used only for quality surveys. More recently, specific tools have been developed, which help planners to improve timetables with customisable statistics and diagrams.

In Switzerland Ullius developed OpenTimeTable (2004), an easy-to-use timetable analysis tool based on SBB data collected at stations and timetabling points. The software visualises delays in graphs comparing them to the planned timetable or in various distributions. The tool is still improved within the SBB and currently used by timetable planners.

In the Netherlands, a consistent works has been carried out (Goverde, 2005) to develop a powerful tool to analyse train describer (TNV) logfiles. Compared to data collected only at timetabling points, TNV data offer a more accurate infrastructure and train movement description, comprehending all block sections, and therefore blocking times. Starting from these trajectories at section-level, accurate speed estimates are computed obtaining reliable arrival and departure time estimates.

Some authors focused on the analysis of the distributions of train delays. Schwanhauser (1974) proposed the use of the negative-exponential distribution to fit non-negative arrival delays at stations. More recent studies (Herrmann, 1996, Goverde et al. 2001, Wendler and Naehrig, 2004 among others) confirm this assumption. Many delay propagation models use the negative exponential distribution to represent the initial delays.

During the last years, different distributions have been used to fit departure, arrival and running times distributions of trains. Yuan (2006) presented a method for fine-tuning the parameters of distribution models. Several commonly applied distributions for train event and process times have been compared using the Kolmogorov-Smirnov goodness-of-fit test. Results for a major station in the Netherlands show that a location-shifted log-normal distribution can be considered the best for the arrival delays both at the station platform and at the approach signal

of the station home signal. The Weibull distribution is proposed for non-negative arrival delays, departure delays and free dwell times of trains.

An innovative approach to distributions modelling has been recently presented by Bueker (2009). To bypass the need of fitting to a given distribution, and to obtain a function which can be adapted to all phenomena, the author proposes a piecewise continuous distribution function, with each continuous segment represented by an act of distribution involving several parallel phases. To guarantee the efficiency of the computational algorithm, a complexity-reduction algorithm has also been developed. The approach is suitable for both delay propagation models and for enhanced estimated-time-of-arrival (ETA) models.

2.2.5 Simulation

Simulation has been used by some researchers to analyse the influence of delays scenarios on train traffic. Various tools have been developed mainly at universities and are commonly used by infrastructure managers and railway consultants especially to support infrastructure planning process or validate timetables. A description of the most used simulation models will be presented in Chapter III; this paragraph presents a brief review of the scientific works based on simulation and not simply describing the tools.

Carey and Carville (2000) developed a discrete simulation algorithm called AATPS (Automatic Train Timetabling and Platforming Systems) and used it to evaluate the effect of parameter and policy changes. The authors demonstrated that on-the-day platform changes and allowing late trains to depart after less than their usual dwell time have dramatic effects in improving punctuality. They also used the model to evaluate and validate the heuristic measures of reliability most commonly used in Britain. More recently the same authors (2003) used the algorithm to evaluate improvements to timetables in large stations.

Middelkoop and Bouwmann (2000, 2001) used the macrosimulation tool SIMONE to evaluate several traffic scenarios in the Netherlands. Some other studies have been performed using the same tool to evaluate the influence of planning norms on punctuality.

Rudolph and Demnitz (2003) describe the improvements for the regular-interval timetable for North Rhine Westphalia using synchronous micro-simulation. The tool has been used for blocking-stairways based timetable planning and stochastic simulation.

Rudolph (2004) developed a strategy to improve punctuality through a better allocation of running time margins. Given the total amount of margins calculated using the UIC Leaflet, after a theoretical analysis of processes and operations, Rudolph demonstrates that the most effective supplements are allocated just before major stops or nodes, because this maximises the probability of a punctual departure from that stations. Rudolph proposes to move the supplements to the dwell time, creating two co-exixting timetables: an operator's schedule and a published timetable. Rudolph uses stochastic microsimulation to verify the effect of this strategy on punctuality.

In a broad-spectrum PhD thesis, Watson (2008) describes the use of microsimulation in Britain, comparing different tools, describing some issues with the use of them and describing a case study. The author uses stochastic microsimulation to evaluate the impact of the ERTMS Level 2 ATP system on timetable robustness (delay propagation) on the congested commuter line between London Victoria and East Croydon.

Description of model validation and its results can be found in some consultancy reports, such as Kohls (2007). In the mentioned study, which follows the specific requirements of Network Rail, the impact of a new station on a line is evaluated. Detailed delay scenario description and calibration results are presented with an analysis of the deviation from real data in each section and for each train family over the entire line.

Micro-simulation is used in many universities for Master and Bachelor Thesis; however, as clearly shown in this review, a detailed study on simulation accuracy in representing railway operations and on parameter calibration has not been produced yet.

2.2.6 Conclusions

Summarising, the trade-off between increasing the utilisation of railway networks and improving the reliability of train operations has been widely studied in the last years.

Queuing models estimate the waiting time of trains: they can be used in strategic planning to evaluate impact of infrastructure improvements or higher number of trains. These models are very efficient, but timetable-independent, so they cannot be used to forecast punctuality. Max-Plus-Algebra Models can evaluate timetable slack and delay propagation on large networks, but they do not estimate

punctuality for the simulated trains. Some stochastic delay propagation models consider real process distributions, but they do not consider the real behaviour of drivers with different interlocking and ATC systems. More advanced stochastic propagation models partially overcome these weaknesses, but do not offer planners intuitive operations representation.

Some authors purpose the use of micro simulation to estimate timetable robustness, but the use of micro simulation requires some specific attention as regards in particular model calibration and validation.

A number of issues and examples concerning this model calibration and validation are discussed in Chapter V. A method for the evaluation of the trade-off between capacity and timetable robustness using calibrated micro-simulation model is then presented and applied to different case studies.

2.3 Reliability

As stated in (Kittelson et al., 2004), reliability of transit services is influenced by some factors, that fall within the competence of the transit company and also by other elements which may not be completely controlled by the operators. as for instance the traffic conditions or demand levels. Also in the railway system train running times, dwell times usually vary a lot because of operational elements or external aspects, so that the planned timetable sometimes is not respected. In both cases reliability could be an important issue and it should be measured.

There could be a number of measures for reliability:

- on-time performance;
- regularity of headway between successive vehicles;
- missed trips;
- distance travelled between mechanical breakdowns.

Usually the on-time performance is considered when the headway is equal or higher than 10 minutes so that passengers consult published timetables. On the contrary, if the headways are lower than 10 minutes, the headway adherence is mainly used (Kittelson et al., 2004). On heavy-rail transit systems the lower frequency of services forces passengers to consult schedules to minimise waiting times at stops:

as a result, the regularity of headways is not considered, and punctuality is normally the only quality measure used in practice.

According to (Kittelson et al., 2004), the on-time performance should be measured, at specific locations of particular interest for passengers, over a series of days. Of course on-time is referred to the planned timetable. These measures should refer to arrivals or departures according to the number of passengers alighting or boarding respectively. Moreover the manual define "on-time" "as being 0 to 5 minutes late" (a departure from a timepoint 0 to 5 minutes late) while some further considerations are drawn for early departures. The LOS ranges are defined in terms of on-time percentage. The thresholds refer to 5 round trips per week when discussing the meaning of the thresholds. Different thresholds mean different perceptions in quality of service.

Also in practise punctuality or on-time performance (P) is usually defined as the percentage of trains arriving within a certain number of minutes form the scheduled arrival time. The threshold values used to measure the punctuality of train operations depends on the country and the operator: in most countries a threshold equal to 5 minutes is used, however, some operators already use 3 minutes or 1 minutes. In some cases they are planning to introduce even more restrictive parameters. Punctuality of operations can be measured at terminal stations, at major stations on the network or at all timetabling points.

In Italy, punctuality is only measured at terminal station of each train, using different threshold for long-distance and commuter traffic. Long distance services, comprehending conventional Intercity, High Speed and overnight services are considered punctual if they arrive with delays less then 15 minutes; a more restrictive 5-minute allowance is used for commuter trains. This has a great impact on traffic planning and control, producing different strategies to improve punctuality in timetabling and by choosing priorities in route assignment as a function of this indicator. For example it may happen that a punctual long distance train could be delayed by 10 minutes in order to reduce by 1 minute the delay of a regional train running with a delay of 6 minutes and obtain two punctual trains for the statistics. In the same way, running time supplements at last timetabling points improve punctuality: even trains affected by significant delays on most of their itinerary can have good performances in reports.

Punctuality is very widely used since it is simple to be calculated and easy to be understood by passengers; however, it gives very limited information about real

operations, since it does not consider the large amount of small delays and stochastic variability that can have a considerable impact on the quality of train operations. In other words, punctuality can be used as very synthetic reliability measure, but is not precise enough to allow traffic modelling in particular on complex networks. Other performance indicators can be found in literature, which are used for more detailed statistical analysis of train traffic and not to define reliability levels:

- Average train delay (A) has the advantage of considering real delay distributions, but it is strongly influenced by few very large delays. Moreover, a large amount of early running trains can lead to very low values which do not represent real delay variability.
- Average non-negative train delay (A_n) could overcome this last weakness, but it underestimates the good behaviour of a system where most trains are early running.
- Camus et al., 2005, proposed a “weighted delay index”, which allows to consider not only the percentage of trips which are on time or not, but also the amount of delay they have. This performance measure takes into consideration the product of two factors: the amount of delay and its probability. It could be used for transit services with headway higher than 10 minutes but not too high, so that passengers could wait for the following trip and delay could be comparable to headway. This does not happen exactly for rail services; in this case headway are usually much higher so that passengers choose the course and usually do not wait for the next trip. Moreover in railway operation high delays may create serious disruptions which are not perceived by users but influence the global reliability of the system. Nevertheless, this approach is a first step towards the consideration of delay distribution within reliability indexes.

As stated before, punctuality (P) is the reliability measure normally used on heavy rail systems. This has the great advantage to be easy to calculate and understood also by customers, who are normally most interested in on-time arriving and normally have no perception of small variations in schedule keeping. But this indicator presents a number of weaknesses regarding in particular its impossibility to represent real traffic distributions:

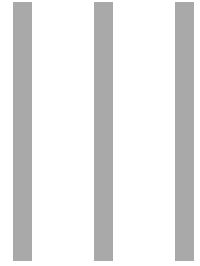
- it indicates the percentage of trains delayed more than a given threshold, but not by how many minutes the services have been delayed (this problem has been already pointed out in (Camus et al, 2005));

- it does not consider trains running out of schedule, but within the allowed delay;
- significantly early running trains are not considered, although they can cause conflicts with other trains in nodes.

In rail transit systems, early running has normally no negative impact on traveler's quality of service perception, since trains must wait the scheduled departure, minimising the risk of train loss. But the presence of more restrictive capacity constraints, in particular at important nodes and stations suggests that early running trains have also to be considered, because they could produce traffic conflicts with other trains, causing secondary delays.

[...]Si era evidentemente appoggiata alla sbarra per godersi la vista del nostro treno, superdirettissimo, espresso del nord, simbolo per quelle popolazioni incolte, di miliardi, vita facile, avventurieri, splendide valige di cuoio, celebrità, dive cinematografiche, una volta al giorno questo meraviglioso spettacolo, e assolutamente gratuito per giunta. [...]

(D.Buzzati, Qualcosa era successo, 1958)



A new reliability indicator

As clearly explained in chapter 2.3, conventional reliability measures like punctuality or mean delay are easy to understand and widely accepted, but do not represent delay distributions in a reliable way.

An alternative parameter R can be obtained by a weighted sum of mean delay and its standard deviation, allowing a better representation of distributions, in particular regarding the impact of higher variability on train traffic:

$$R = k \times \sigma + A \quad (3.1)$$

where σ is the standard deviation of the considered delay distribution, k is the coefficient to weight the importance of standard deviation on traffic quality (some tests on Italian cases show that values about 0.5 could guarantee a good representation, without exceeding in sensitivity) and A is the average train delay. Like the Average Delay (A), this indicator has the disadvantage of being very sensitive to very large delays and to be not easy linkable to a delay perception or to a Level of

Service. To overcome these weaknesses, a new indicator could be used, which considers the real delay distribution together with a punctuality threshold.

This new indicator is the “*Delay Frequency Index*” F . It allows to consider both early and delayed arrival/departure. The indicator includes the different importance of early and delayed events and filters very high delays due to disruptions. This last aspect may avoid quality underestimation due to unusual high disruptions. The index F is defined as follows:

$$F = \sum_{i=1}^n \left(\frac{N_i}{N} \times \frac{D_i}{P} \times f \right) \quad (3.2)$$

where N is the total number of trains, N_i is the number of trains arriving in the delay interval i , D_i is the delay in interval i , P is the selected on-time bound and f is a weight coefficient (in this study $f = 1$ for delays (f_p) and $f = -0,5$ for early running trains (f_n) have been used). A sort of upper bound D_{max} is used for D_i to separate normal variability from large delays and therefore to limit the weight of these system failures on the indicator. In particular if

$$D_i > D_{max} \Rightarrow D_i = D_{max} \quad (3.3)$$

and D_{max} equal to 20 minutes has been used in this study. Figure 1 represents graphically the basic principles of F .

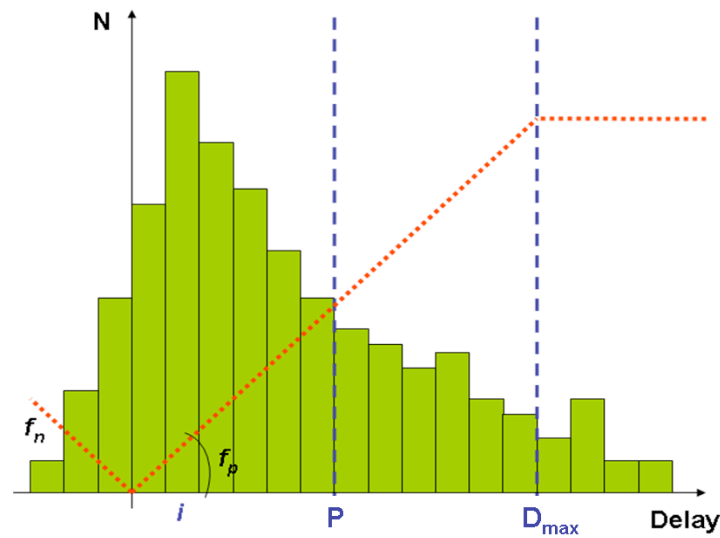


Figure 3.1 Parameters involved in F.

The resulting index F is a synthetic percentage indicator, which is smaller as the traffic quality increases and shows values higher than 100 for non-acceptable high variability and delays. F indicates both running time deviation and punctuality: therefore it can be used as comprehensive parameter for quality of traffic and service measurement.

Four real examples are represented in Figure 2, where very different distributions are compared. The distributions 1 – 2 and 3 – 4 have respectively identical P -levels. The tables show that the average train delay A , the average non-negative train delay A_n and the weighted sum of mean delay and its standard deviation R values have a high variability that complicates their use as absolute performance indicators. On the contrary F is more stable, but it must be connected to an absolute scale in order to increase its comprehensibility.

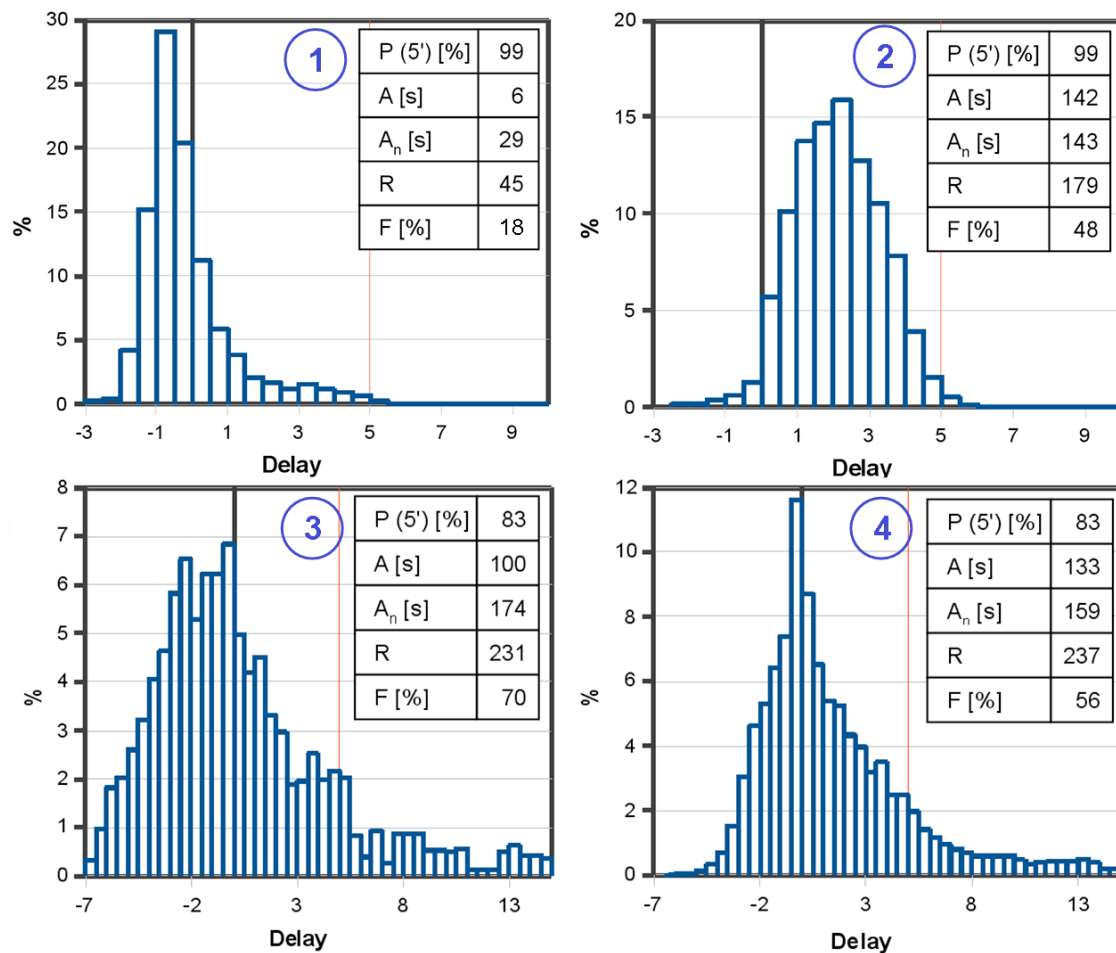


Figure 3.2 Comparison between performance indicators for different distributions with the same on-time percentage (P).

Since it is expressed in percentage, a number of intervals can be defined, which allows to associate F index to Level of Service (LOS) thresholds. After an extensive analysis of the Italian railway nodes and lines, a very simple LOS Scale is proposed, which can be qualitatively compared to the actual reliability LOS (P-LOS) based on (TCQSM). A precise and concise comparative analysis between the two scales is not possible, since F describes the distributions and P the number of delayed transits. A simplified comparison is presented in Table 3, where typical Italian delay distributions have been used to determine F.

F-LOS have in general larger intervals, but with a stricter A range: as a result F normally gives higher LOS than the corresponding P in the D and E intervals. F results only in very wide and delayed distributions, whereas in P-LOS an F is also assigned to distributions where the worst 20% of records have 6' delays. On the other side, only

very tight distributions obtain an A in F-LOS, also if the 96% of trains have less than 5' delay.

LOS	F	P
A	$F < 20$	$P > 95$
B	$20 < F < 40$	$95 < P < 90$
C	$40 < F < 60$	$90 < P < 85$
D	$60 < F < 80$	$85 < P < 80$
E	$80 < F < 100$	$80 < P < 75$
F	$F > 100$	$P < 75$

Table 3.1: LOS intervals for the F index compared to transit On-Time Performance LOS (P)

As first and illustrative applications, LOS have been calculated for the four distributions represented in Figure 3.2. In Table 3.2 each row refers to correspondent distribution, the percentage of trains with selected and growing delays thresholds are shown in 5 separate columns while P and F with their respective LOS are included in the last four columns

	D<-1' [%]	D<1' [%]	D<3' [%]	D<5' [%]	D<10' [%]	P	P -LOS	F	F -LOS
1	0	91	94	99	100	99	A	18	A
2	0	8	63	99	100	99	A	48	C
3	36	60	75	83	91	83	D	70	D
4	15	49	70	83	94	83	D	56	C

Table 3.2 Comparison of cumulative delay percentages with the P and F indicators and LOS for the distributions depicted in Figure 2.

The distributions 1 and 2 are quite different as regards reliability. The first one represents an ideal situation with a very good adherence between real and planned timetable and some early arrivals. The second distribution refers to a worse situation where a large amount of trains has a delay even if low. They have the same P-LOS "A". On the contrary the application of the new index leads to different F-LOS (respectively A and C) and this result seems to be more adherent to the system behaviour. Similar considerations may be exposed for distributions 3 and 4. They have the same P-LOS even if the last one is better as regards reliability because it is

quite closer to the planned value. The proposed approach allows to distinguish between them leading to F-LOS D and C respectively.

3.1 Case study

P and F reliability indexes have been tested in real life cases study in different Italian contexts. In particular the results of two significant cases are presented in the following: the Palermo – Punta Raisi International Airport line and the Turin node.

The Palermo – Punta Raisi International Airport is a single track line, whose length is 35 km. Eight intermediate stations are presents where train operations may take place. The figure 3.3 shows a simplified schema of this line.

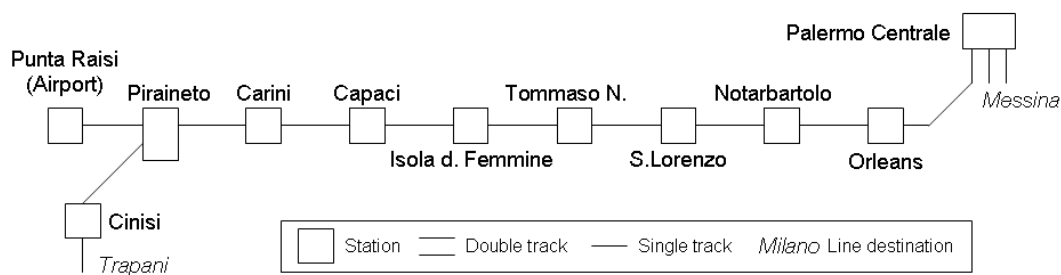


Figure 3.3 Simple layout of the Palermo-Punta Raisi Airport link.

On this infrastructure a regular-interval timetable is offered. During the peak hours the number of trains per direction rises up to 4 trains/hour. The service is mainly dedicated to commuters, even if it performs the connection between the city of Palermo and its international airport, which is really important for tourist trips (it is the most important airport of Sicily).

Over this line, most of trains depart/end in Palermo and Punta Raisi International Airport and so there are no additional delays due to traffic phenomena which take place outside the system. Nevertheless, delays are usually present due to overlong stop time at urban stops and numerous level crossings with surface streets. The figure 3.4 shows the plotting of real traffic behaviour over two months compared to the planned timetable (black lines).

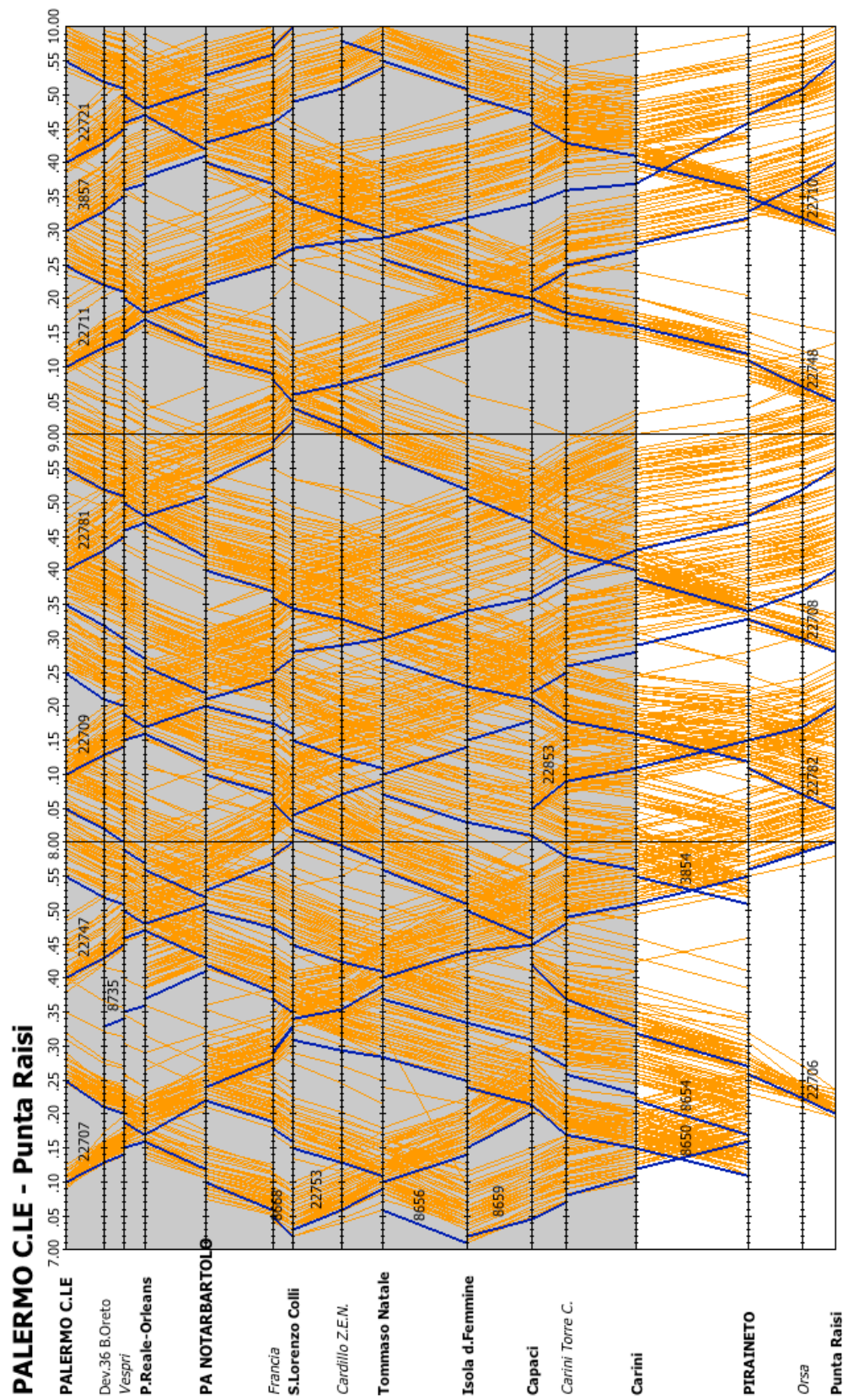


Figure 3.4 Railway traffic diagram for the Palermo-Punta Raisi International Airport line.

Nevertheless, the measured punctuality of this line is relatively high at terminal stations due to running time margins placed at the end of each trip. In this case it is really evident that this reliability measure does not represent the real behaviour of the system.

The second case study is the Turin railway node. Turin is the second largest city in northern Italy with 4 million inhabitants in its region. The Turin railway node is quite complex. Ten lines converge to the node (5 double track and 5 single track lines) as shown in the figure 3.5. Some of them are dedicated to short distance commuter traffic, while other lines play an important role at national and European level. For example Turin belongs to the Fifth Pan-European Corridor (Lyon – Turin – Milan – Venice – Trieste – Ljubljana – Kiev).

Within the node many junctions and 14 stations are placed, one of them is dedicated to freight traffic. Turin main station (Turin Porta Nuova) is a terminal station with 20 tracks.

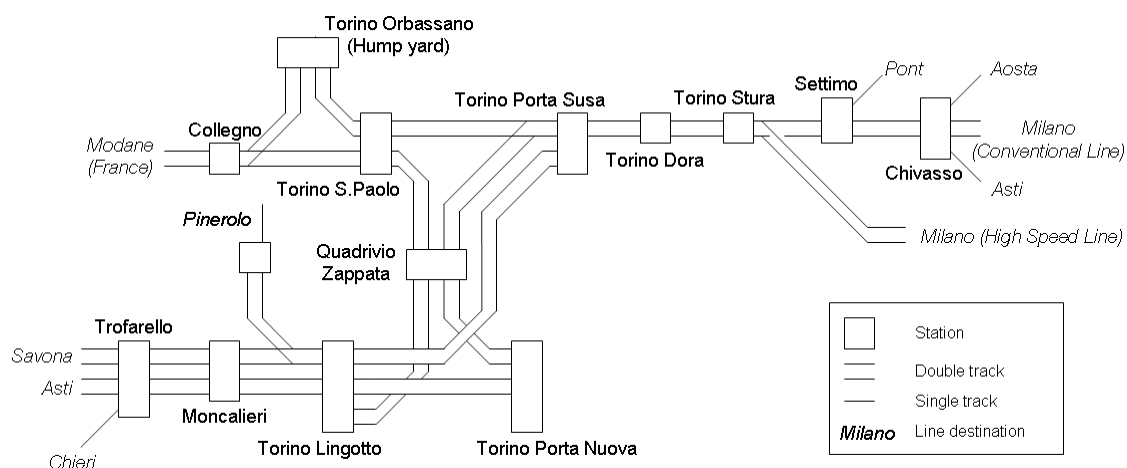


Figure 3.5 Simple layout of Turin railway node.

The node is interested by high and mixed traffic volumes. Traffic is mixed because of the presence of both freight and passengers services and, concerning passengers, both commuters and long distance services are present. Regardless of its complexity and of the fact that the lines normally link the node with other cities located outside the node itself, reliability in Turin is actually really high.

3.2 Reliability measures in the case studies

Reliability indicators have been tested on Palermo-Punta Raisi International Airport line and Turin node. The indexes have been calculated first considering the arrival distributions at the end of the train trip and then globally over the whole line and node respectively. In fact different ways of measuring the indexes may lead to reliability results really different even if referred to the same railway behaviour.

3.2.1 Reliability measure at the end of a trip

The first case refers to the reliability indexes measured at the end of train trips. This is the method currently used in Italy. The results are shown in table 3, which includes a comparison between P, F and arrival rates at given delay thresholds for some commuter lines. In this table each row refers to a specific trip type. The first column indicates the case study (Sicily or Turin), the second and third columns identify the train direction and so the corresponding ending station. In the case of Turin node, only stations within the node have been considered as starting and ending points. The following five columns show the percentage of Delay (D) lower than given thresholds and they depict the observed arrival distributions. Finally the last four columns show the percentage of arrivals with a delay lower than 5 minutes (P), the corresponding P-LOS (according to (Kittelson et al., 2004)), the proposed F index and the corresponding F-LOS.

It should be noticed that F-LOS are normally equal or better than P-LOS. This happens because the observed real-life distributions often have a relevant part (round 10%) within a delay between 5 and 7 minutes. These records are simply considered as delays in P, according to the methodology proposed in (TCQSM,2004), while in F their real values are computed. This result is coherent with the methodology proposed in (Camus et al., 2005) and it represents a sort of extension to railway system of that approach. The proposed F index allows considering the real amount of delay and not only the number of delayed trains.

Of course, regardless of the applied method, the reliability measure at the ending station does not allow to consider the punctuality at intermediate stops, which could be important, on the contrary, both for passengers and for infrastructure manager.

Case Study	Arrivals at	From	D<-1'	D<1'	D<3'	D<5'	D<10'	P	P - LOS	F	F - LOS
Sicily	Punta Raisi	PA Centrale	35	56	72	83	97	83	D	61	D
Sicily	PA Centrale	Punta Raisi	6	28	56	75	94	75	E	72	D
Turin (TO)	Chivasso	TO Porta Nuova	15	49	70	83	94	83	D	56	C
Turin (TO)	TO Porta Nuova	Chivasso	36	60	75	83	91	83	D	70	D
Turin (TO)	Collegno	TO Porta Nuova	3	33	61	79	93	79	E	64	D
Turin (TO)	TO Porta Nuova	Collegno	27	58	77	87	95	87	C	53	C
Turin (TO)	Trofarello	TO Porta Nuova	9	39	66	81	94	81	D	60	C
Turin (TO)	TO Porta Nuova	Trofarello	11	48	79	90	96	90	C	44	C

Table 3.3: Comparison of cumulative delay percentages with the P and F indicators and LOS

3.2.2 Global evaluation for a single-track line

In this section the whole railway line between Palermo and Punta Raisi International Airport is considered. In particular the reliability is determined with reference to all the existing stations and not only in the ending one. The aim is to perform a sort of comprehensive reliability estimation.

In the following, departure delay distributions are considered for the first station while arrival delay distributions are used at all the intermediate and ending stations.

Both P and F indexes have been calculated for each delay distribution. Their average values have been used as global performance indicator for the whole line. This procedure leads to a more realistic reliability index, since it considers the effective reliability on the line, which of course affects the capacity on the line and interests the boarding and alighting passengers at the intermediate stops.

In this case the observed higher delays in the central part of the line cause lower perceived quality in real life, which would not appear if only the reliability at the ending stations was considered.

In Figure 3.6, delay distributions at all stations for Palermo – Punta Raisi International Airport services are depicted. In particular, darker parts refer to on-time rates while lighter grey represent growing delays or early running. As a result, the darker is the histogram, the more punctual are the services. On the other side, significant light intervals indicate worse delay distributions. On the same figure the

line with squares represents the estimated P rates while the line with rhombs represents F levels. Average P and F values are indicated in the small box.

P seems sensitive to very high delay rates at few intermediate stations. This leads to a global LOS E for the whole line, despite of LOS D obtained if only the delay at the trip end was considered. F index is more stable, since distributions' shape from station to station is not as variable as it would appear by considering the high variability of the P-values. In some stations, 30% of trains are concentrated between 3-minute and 7-minute delay, causing higher P variations. As a result, a LOS D for the proposed F index allows to better represent the real perception of quality from both traffic and passengers point of view.

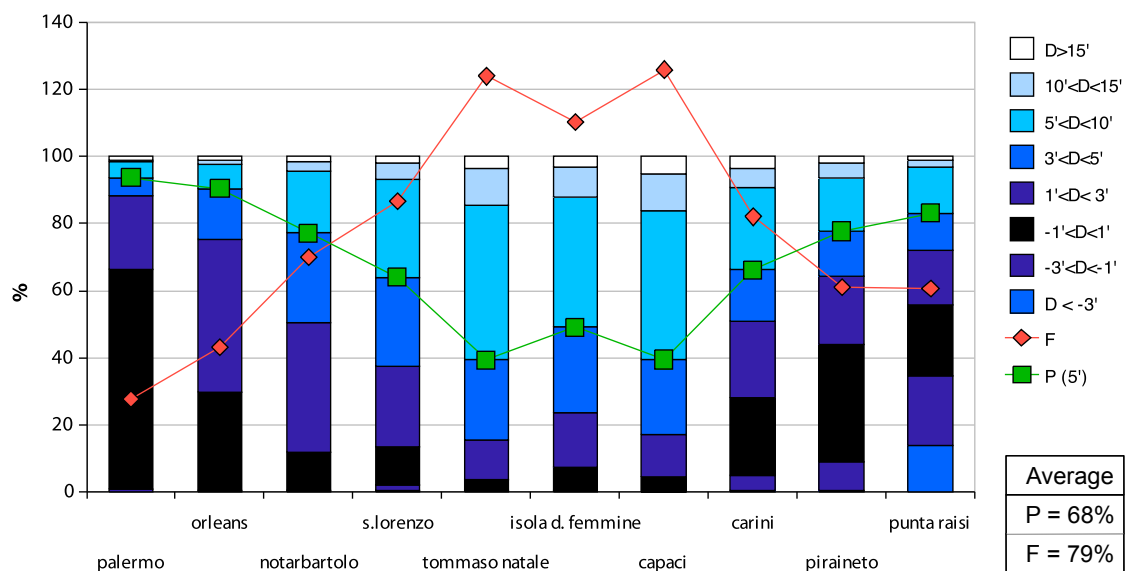


Figure 3.6: Percentage arrival distribution at all stations on the Palermo – Punta Raisi Airport Link, compared with P and F.

3.2.3 Global evaluation for a node with mixed traffic

It is quite more difficult to estimate a global reliability index for a whole complex node, where many line converge and different services are present. Nevertheless, in this section P and F indexes have been estimated as global performance indicator for Turin node.

Global LOS have been calculated on the basis of average P and F values for each line and train category, weighted on the respective number of trains. In this study the

same relative importance has been given to different train categories, although in some countries high-speed or commuter trains may have higher priority.

The figure 3.7 uses the same graphical language as figure 3.6. In particular delay distributions, P and F values for selected stations are represented, while P and F LOS are finally calculated.

Different behaviours for train categories can be noticed. High-speed services are often characterised by early arrivals, commuter trains are quite punctual while long-distance Intercity trains present a variable behaviour. This is due to the fact that in Italy High-speed services have significant running-time supplements scheduled before their entrance into the node so that their delays decrease while approaching Turin Porta Nuova station. These very significant early running causes poor F-LOS (E), because the often registered 5- or even more-minute early arrivals, while better P-LOS estimations. Lower reliability estimations seem more adherent to reality as longer track occupations within the node may lead to higher problems for the infrastructure manager. For commuter services the differences between P and F LOS mainly depend on the shape of the delay distribution, while in both cases intercity trains reliability is very bad.

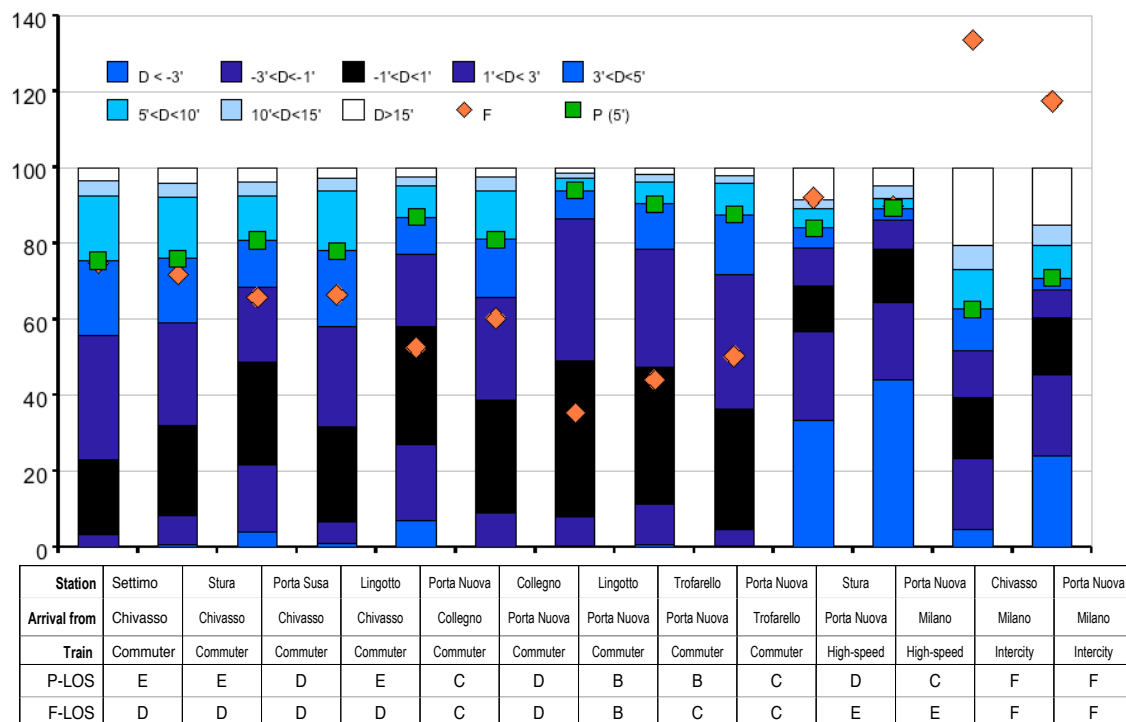


Figure 3.6: Percentage arrival distribution at selected stations and for various train types in Turin node; comparison between P and F indexes.

The table 3.4 shows first the P and F values and relative LOS for different lines and kind of services. Then the same values have been determined for trains categories and the last row refer to the whole node. Also the number of analysed trains is reported for each row.

The figure 3.8 shows the same results of table 4 in a graphical way. It allows to understand immediately the differences between the methods. These differences are more evident for high-speed trains and for some commuter services according to their delay distributions. As a result the reliability of the whole node is different and the new approach leads to a better LOS estimation which is more consistent with passengers' perception and infrastructure manager point of view.

Trains	Line	Number of records	P	P-LOS	F	F-LOS
High speed	Turin – Milan	675	84	D	81	E
High speed	Milan – Turin	675	86	C	93	E
Intercity	Turin – Chivasso	832	94	B	24	B
Intercity	Chivasso – Turin	832	68	F	123	F
Commuter	Turin – Chivasso	4438	87	C	46	C
Commuter	Chivasso – Turin	4438	79	D	67	D
Intercity	Turin – Trofarello	1812	82	D	62	D
Intercity	Trofarello – Torno	1812	56	F	142	F
Commuter	Turin – Trofarello	5156	91	B	42	C
Commuter	Trofarello – Torno	5156	92	B	37	B
Commuter	Turin – Collegno	2626	82	D	58	C
Commuter	Collegno – Turin	2626	79	E	64	D
Commuter	Turin – Moncalieri Sangone	1924	86	C	56	C
Commuter	Moncalieri Sangone – Turin	1924	81	D	62	D
Commuter	Settimo – Trofarello	1962	78	E	67	D
Commuter	Trofarello – Settimo	1962	86	C	49	B
High speed	all lines	1350	85	C	87	E
Intercity	all lines	5288	73	F	93	E
Commuter	all lines	32212	85	C	52	C
all trains	all lines	38850	84	D	59	C

Table 3.4: Global performance indicators for all lines and train categories.

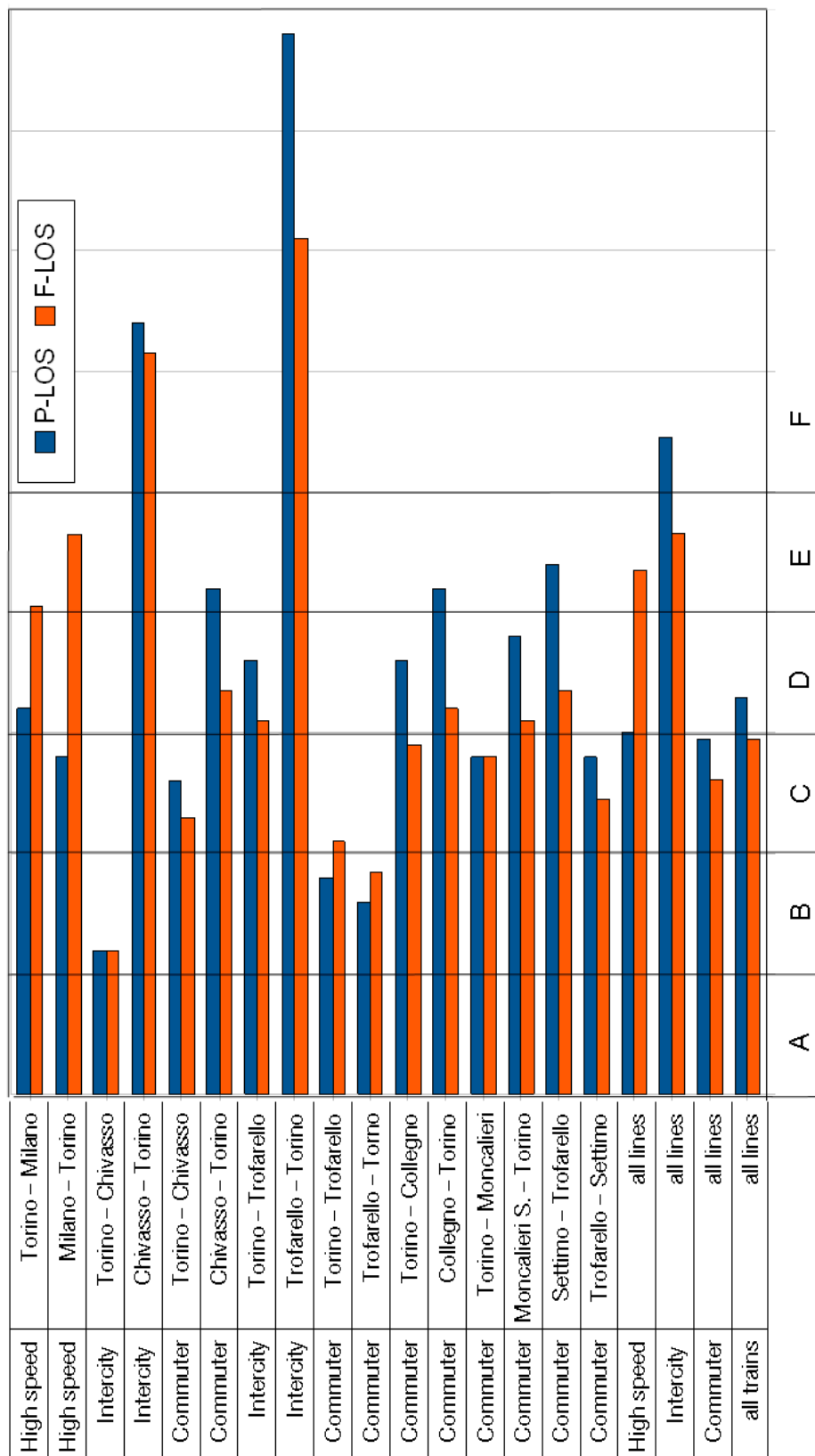


Figure 3.8 Comparison between P-LOS and F-LOS for all lines and train categories.

3.3 Conclusions

The “Frequency of delay Index” (F), proposed in this paper, may represent a new performance measure for the reliability of railway services. It allows taking into consideration both values and shape of the delay distribution of the railway services and it leads to results which better represent the real effects of punctuality on railway traffic, if compared to the ones of the existing approaches. It is coherent with passengers’ perspective but may be useful also for the infrastructure manager who should control train operations.

The applications of this method to two important case studies in Italy shows that it may be used for real-life problems. Moreover, the tests underline the importance of using global reliability measures (referred to a line, to a node, to different train categories and so on) together with punctual estimations in order to obtain a more complete evaluation of the system behaviour.

*Stanotte, signor Cavaliere. Ha fischiato. L'ho sentito
fischiare... Il treno?
Sissignore. E se sapesse dove sono arrivato! In
Siberia... oppure oppure... nelle foreste del Congo... Si
fa in un attimo, signor Cavaliere!*

L. Pirandello, “il Treno ha fischiato” 1914)

IV

Calibration and validation of a micro- simulation model

Simulation of railway networks has a long tradition, starting many decades ago in railway laboratories, where models in the scale 1:76 were used to reproduce networks and control them using realistic interlocking systems. The growth in the computational power of computers and the creation of graphical interfaces on one side and the relative simplicity of the basic rules of train traffic on the other have led to the development of simulation tools.

Simulation tools were first able to simulate relatively small networks, considering all trains in a deterministic way. These tools were mainly used to support infrastructure planning and especially regarding capacity estimations. More recently, the further increase in the performance of computers and the possibility of an automatic import of infrastructure layouts and timetables widened the application spectrum of micro-simulators to large nodes and to more detailed stochastic stability evaluations.

Compared to deterministic simulation, the stochastic one presents the great advantage of considering also variability in process times, increasing the precision of the outputs. But the possibility of obtaining realistic results is strictly related with the quality of the input data used to model stochastic phenomena, and which have to be accurately calibrated.

Calibration starts with the analysis of real data, which can be collected at track circuits or on-board, using digital train event recorders or GPS, in order to obtain higher precision. Since a perfect representation of all stochastic and deterministic parameters involved in rail traffic is not possible, a calibrated model must be validated to evaluate its precision before using it in practice. Calibration has been tested on the Venice - Trieste and Palermo - Punta Raisi lines and in the node of Turin, where the obtained model has also been used to forecast reliability of the operations.

First, different approaches to simulation are presented, then synchronous micro-simulation with its advantages and weaknesses is analysed in detail. Calibration, of motion equation using GPS and station-passing data, its validation and applications to three case studies are described in the second and third part of the chapter.

4.1 Simulation Models

Simulation has been defined by Robinson (1994) as “a model that mimics reality” and by Gamerman and Lopez (2006) as “treatment of a real problem through reproduction in an environment controlled by the experimenter”.

Simulation is *deterministic* if all parameters are defined by the user and do not contain any random components. Deterministic models are used to represent real systems which are too complex to be evaluated analytically. In *stochastic* simulations random components are introduced to better represent one or more phenomena. Since stochastic simulations are used to evaluate the behaviour of a system with some random factors, results of a single simulation run have no statistic relevance; as a consequence, a number of *multiple simulations* must be performed. Deterministic simulations support timetable planning or the design of new infrastructures, while stochastic models allow timetable robustness or stability analysis.

Simulation is *static* if time does not play a role. A dynamic simulation model shows how a system evolves over time. Railway simulation models are dynamic, since they are explicitly built to study traffic evolution in a given time interval.

In *continuous* simulation models, the value of state variables change continuously in time, therefore it is calculated with analytic continuous resolution of state equations. In *discrete* simulation models, state variables are calculated only at fixed time intervals (fixed-increment time advance models) or when an event occurs, independently from the time-span between two successive events (*next-event time*). Railway simulators are continuous when they solve the motion equation with a continuous algorithm, while more simplified models use pre-defined process and event times in a discrete approach.

Macro-simulation models use a simplified infrastructure model to reduce computational time and therefore allow simulation of larger networks. *Micro-simulation* models offer a description of infrastructure which reproduces the functionality of interlocking, safety and block systems.

3.1.1 Macro-simulation models

Macro-simulation models have been developed mainly to simulate timetables at network level.

FASTA (FAhrplan STAbilitaet, Nordeen, 1996) is a discrete-event simulation system developed at the EPFL Lausanne. The network is modelled as a directed graph where nodes correspond to stations and edges to the line sections; infrastructure constraints are represented by minimum headway times. Running times are calculated during the simulation, also considering random parameters in stochastic simulations.

SIMONE (Bouwman et al., 2004) is a discrete-event simulation tool used by ProRail, the Dutch Infrastructure Manager since 1999. Receiving automatic input from the timetabling system DONS, SIMONE perform stochastic simulation to assess punctuality with different timetable or infrastructure scenarios. The infrastructure model is made by timetable points and line sections; in each timetable point a train is assigned to a platform group. A train must wait outside the station area if all possible platforms are already occupied. Stochastic disturbances to dwell and running time to define primary delays can be inserted to perform multiple simulations, producing a variety of statistics at different aggregation level and considering the most commonly used measures.

A stochastic macro-simulation model has been developed at the RWTH Aachen (Weidner, 2007) to be integrated in the strategic planning tools of DB. The model represents the impact of real process times distributions on a given timetable producing statistics about punctuality and the number of missed connections. The low computation load of the model enabled studied on a very large scale, such as on the entire SBB network (Akermann, 2008).

MERIT (Modelling the Reliability of Infrastructure and Timetable) (AEAT, 2002) is a discrete-event simulation tool proposed by DeltaRail for simulations on large networks. The tool is based on a significantly simplified model in which margins and technical times are approximated from the planned values and instead of distributions of primary delays, it uses randomly generated incidents to assess timetable robustness. More accurate micro-simulation models have replaced MERIT by Network Rail since 2005

The deterministic macro-simulator NEMO (Netz-EvaluationsMOdell) (IVE, 2005) is a strategic planning tool for the evaluation of infrastructure scenarios. NEMO allows the automatic generation of future capacity requirements considering traffic forecasts. The tool considers freight and passenger traffic in a separate way. Freight traffic is given as an O/D matrix that distinguishes different kinds of goods; NEMO calculates the corresponding number of loaded freight wagons per segment and builds up block trains for higher traffic amounts and single wagon traffic based on a production network for the resting parts. Passenger traffic is based on given line concepts, among which the expected number of passenger trains and therefore have no direct link from demand forecasts. Freight and passenger traffic are assigned to the network, leading to network loads.

4.1.2 Synchronous microscopic simulation models

A number of synchronous micro-simulation models have been presented in the last years. Since all models use the same approach, solving the motion equation of trains which are moving together on a microscopic network with respect to interlocking and safety system functionality, differences between models are mainly in the flexibility to represent different technologies, in the capability to receive automatic inputs or generate specific outputs.

VISION (Visualisation and Interactive Simulation of Railway Networks) is owned and supported by DeltaRail. It has been very widely used in Britain, with some licences also in Spain and Italy (RFI). The model focuses mainly on the infrastructure

design for small network sections. This is reflected in the simple interface for building models and in the outputs. The lack of scalability, the slow simulation run times, the limited number of outputs seriously limit the potentials of the tool for a wider use.

The simulation tool SIMU VII (Klahn, 1992 and 1994), has been developed at the Institute for Transport, Railway Construction and Operation (IVE), University of Hanover for the simulation of railway nodes. The infrastructure is represented by a directed graph, while complex structures formed by more arcs are used to reproduce the technological equipment such as the functions of train protection systems. The model has been refined to allow simulation of networks and integrated with a tool, called SIMUPLAN, to support timetable construction with a simple conflict detection and an interactive graphical interface. Incidents and initial delays can be inserted to evaluate timetable behaviour under perturbed conditions with multiple simulations. To go beyond the weaknesses of SIMU VII in representing specific technological features and enable simulation of even larger nodes the completely new object-oriented SIMU ++ was developed, in 1999 renamed RailSys.

RailSys (RMCon, 2008) is a comprehensive signal-berth simulation package. The software database structure allows simple storing of very large models, which is reflected by the success at many infrastructure operators. The model features an improved timetable construction system and an automatic slot search algorithm (Hauptmann, 2000). Refined simulation technique to prevent deadlocks, a number of input and output capabilities and the presence of variable process and event times made this tool suitable for large-network stochastic simulations already in 2002 (Rudolph and Demnitz, 2003).

OpenTrack (Huerlimann, 2001) is a user-friendly simulation tool developed at the Institute for Transport Planning and Systems (IVT) of the ETH Zurich and now supplied and refined by OpenTrack Railway Technology Ltd. Together with RailSys, OpenTrack is the most used simulation software. Initially developed some years later than RailSys, but with a more intuitive interface and high quality and variety of outputs especially for deterministic simulations, OpenTrack is more used among consultancies and at universities. On the other hand, the file-based infrastructure models, and the lack of an internal stochastic simulation analysis tool have limited its use for stochastic simulations of large networks. More recently the improvement of input and output capabilities, of the representation of stochastic process-times and of the dispatching rules have made this tool suitable for reliable stochastic analysis.

4.1.3 Asynchronous simulation models

While synchronous models have been developed in different countries and universities, asynchronous simulation tools have been mainly studied at the Institute of Transport Engineering (VIA) of the RWTH Aachen. In a standard asynchronous simulation, trains with highest priority are simulated first, and conflicts among them solved with a first-come - first-served strategy; the resulting infrastructure occupations are stored. Then the process is repeated for each priority group, more and more saturating the time-windows which are still unused. Therefore no conflict among trains with different priority is possible: high-priority services are never forced to brake or stop by conflicts with low-priority trains.

The simulation model STRESI (STReckenSimulation) has been presented in 1985 and can therefore be considered the first simulation which has been widely used. Since no graphical user interface was available at that time, the only infrastructure to be considered were lines, whose characteristics were inserted in tables, and standard stations were only defined by the number of tracks, the corresponding speed and some other parameters. Stochastic phenomena were limited to incidents within the network (and therefore no initial delays), but multiple simulations were already possible.

At the same Institute, the powerful infrastructure model called SPURPLAN (Bruenger, 1995) was developed, which can be stored in a database and contains a detailed infrastructure description. SPURPLAN is a graph, which contains not only the infrastructure and interlocking systems, but also route priorities at each station and for each train category or family. This allows automatic routing of trains and platform assignment at stations. The model was first developed for the RUT timetabling system based on blocking time theory currently in use in Germany.

On the basis of the SPURPLAN graph, a software family comprehending the already mentioned capacity assessment algorithm based on queueing theory ANKE (2.3.2, Vakhtel, 2002), the asynchronous simulation tool BABSI (BetriebABwicklungSimulation, Groeger, 2002) and the dispatching-support tool ASDIS (ASynchrone DISposition, Jacobs, 2002).

BABSI uses the asynchronous scheduling approach segment-wise, where the ends of a segment are determined for each train by a scheduled stop with overtaking possibility. When all occupations for trains of a given priority level are calculated and stored and shown in the graph, conflicts are detected and solved by scheduling all

trains segment by segment in chronological order. If after resolution of a series of conflicts, the timetable for a train is not acceptable, the segments for the train are merged to improve its schedule, reducing possibilities of a take-over. As a result the tool can work as a semi-automatic timetabling tool, which uses a timetable draft as input and solves all conflicts producing a feasible timetable. The behaviour of the conflict resolution algorithms depends on the priority assigned to the possible dispatching measures (rerouting, running time extensions, etc...). The tool has been integrated in the timetabling system to support planners, suggesting conflict resolutions or inserting new slots into existing timetable structures. Further functions are the (deterministic) simulation of a given timetable with pre-defined buffer times to detect conflicts and the stochastic simulation to evaluate timetable robustness.

4.1.4 Asynchronous vs Synchronous simulation

Asynchronous and synchronous models show different strengths and weaknesses, which make them complementary tools to support railway planning. Synchronous simulation allows a more realistic representation of train traffic, with all trains simultaneously moving on lines and therefore interacting between each other, with secondary delays transferred from high-priority to low-priority trains and vice-versa. On synchronous models it is also simple to implement a wide variety of incidents, which may involve nearby all infrastructure and signalling elements, also featuring stochastic variability, such as intermittent opening and closing of at-level crossings.

Asynchronous simulation is more suitable to support timetable construction, which can be performed stepwise and where low-priority trains are normally scheduled after other services. It is also simple to define new courses within a pre-defined timetable, obtaining a conflict-free timetable and an automatic allocation of station tracks.

Synchronous models do not always manage the **priority** order correctly, since priority is only based on pre-reservation of block section, causing unrealistic hindrances to other trains and increasing occupation of the conflict sections. On the other hand, in asynchronous tools trains with high priority are always preferred and displace trains with lower priority. Because of the strictly hierarchical structure of an asynchronous simulation, trains with low priority may receive more delays in an asynchronous simulation than they would in reality.

Another weakness of synchronous models is represented by **deadlocks**. They normally occur on single-track lines, when two or more trains are allowed to block and enter into the same line section at the same time, finally stopping in front of each other until the end of the simulation. Deadlocks are solved using simple algorithms, which are normally reliable and efficient, except for some rare situations (\Rightarrow 4.4)

4.2 Building a synchronous simulation model

To obtain the highest precision, a microscopic simulation model contains all characteristics of the real world, which have an influence on train movements and dynamics. On the other hand, since simulation and especially infrastructure modelling are quite time-consuming, parameters which have a influence smaller than the expected error could even be ignored, in order to obtain results more efficiently.

Regarding the line alignment, gradients have to be considered in detail if they are significant for train dynamics (normally $> 10\%$). While in many running time calculation softwares mean gradients for longer sections are used, each gradient has to be modelled especially when heavy freight trains are considered. The combination of low adhesive weight and wet rails reduces significantly freight train acceleration also if the ramp is just some hundred metres long. Tunnel position and kind (single or double-track, smooth or rough) are also inserted into the model, where their resistance is significant ($V > 60$ km/h).

On heavy railway lines, a detailed description of curve radii can be simplified considering mean curves on longer sections or adding the curve resistance to the gradients obtaining a total resistance due to line alignment.

The interlocking and block system have to be modelled in high detail, comprehending the complete station layout with all track circuits and all signal with their respective aspects associated to the possible routes. Also the way routes are released has to be modelled, considering release groups or other technical constraints, including release times or interdependencies among routes which do not appear analysing the track layout. Special attention has to be paid to model the braking supervision curves, which have an impact on both train behaviour and infrastructure occupation. Simulation softwares as default calculate braking

behaviour as continuous braking curves to the stop point: this is realistic considering lines without supervision and with many, but not all, systems.

For example on Italian conventional lines a cab signalling system (BACC) with four different aspects is applied. It allows a maximum speed of 250 km/h, which is reduced stepwise when the number of 1350 metre-long free track circuits in front of the train is lower (Figure 4.1).

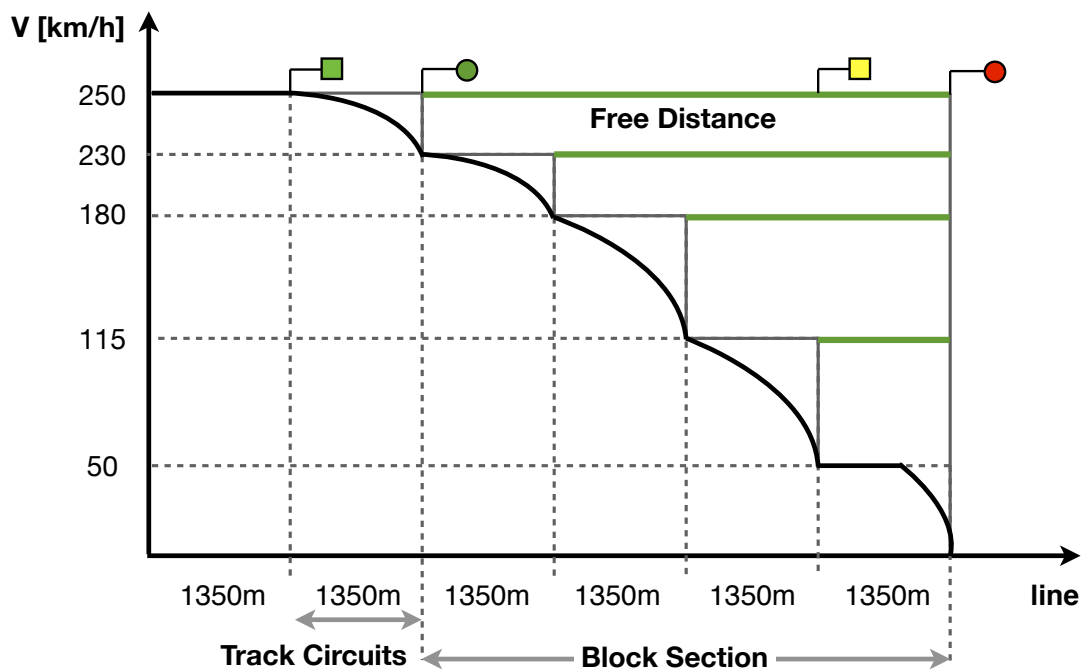


Figure 4.1: Braking supervision in BACC for high speed traffic.

4.2.1 Stochastic phenomena and secondary delays

When the infrastructure model is built and error-free, in order to simulate railway traffic it is necessary to define rules to represent the stochastic behaviour of trains. A detailed analysis of the sources of disturbances and their effects can be found in literature (Vromans, 2005), while in this section an approach to model train behaviour within a simulation tool is presented.

The complexity of modern railways, with the interaction among many processes and human factors leads to variability in the process times and to potential risks of disruptions due to malfunctions and deteriorations. In few cases devices (e.g.

switches or signals) can be found, which are particularly vulnerable to disruptions, and can be simulated. On the contrary, on most networks, breakdowns are more randomly distributed, significantly reducing the statistic relevance of the simulation of a specific case. Therefore, to validate a model, major disturbances within the node have to be excluded, and only “normal” variability is considered.

As a result, primary variability sources as the initial delays, the variability of stop times and of train running time are considered. The calibration of running times is explained in more detail in section 4.3, while this chapter is focused on the initial delays and the stop times.

Initial delays

Initial delays are the departure delays at the first station within the simulation area. Since micro-simulation reproduces in detail the interactions among trains, secondary delays are implicitly calculated within simulation, and if initial delay distributions are influenced by them a potential error is introduced. Thus, to simulate a given node, the border of the simulation is not formed by the door station on each line, but from the first station after it on each branch line, where the possibility of conflicts is lower, since trains are not diverging/converging. In other words, it is important to choose stations where potential conflicts among different services are limited. This rule cannot be respected at dead-end stations, where trains usually start (and therefore initial delays are applied) with potential conflicts with trains arriving. As a result, the departure times measured by track circuits are strongly influenced by conflicts with late arriving trains and therefore contain both primary and secondary delays. If such distribution is used for simulation, it leads to an overestimation of primary delays which, combined to the delays of the arriving trains, could lead to even later departures.

To explain the issues related to the initial delays, a simple example of railway node is considered (Fig. 4.2). Station A is a dead-end station, with one line coming from D, with some intermediate stations. In D two branch lines separate from the main one. To simulate the lines between A and D, the model is extended to E, F and G, since in D distributions could be affected by conflicts among services, as for example a train A-G with an E-A. In E the departure delay of an E-A service can be considered as primary for the simulation network, since it has been generated elsewhere outside the network.

Let's have a service in A with scheduled departure at 8.00, which is ready, but has to wait because of a conflict with delayed train scheduled to arrive at 7.58. The

train will start at 8.03, with 3-minute secondary delay. As a result, to simulate this example using departure delay in A, 3 minutes delay will be inserted instead of a punctual departure also if there is no conflict with the arriving train.

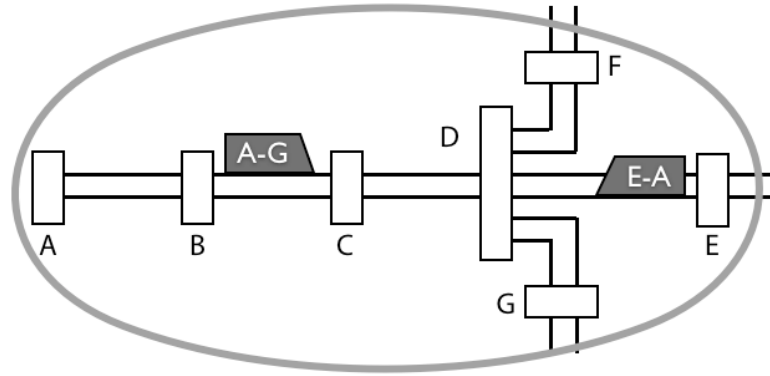


Figure 4.2: Simulation area to simulate railway operations between A and D

With the described simulation area, departure delay could be considered as a reasonable approximation to represent primary delays, except for the dead-end stations and other stations where conflicts with starting trains are significant. In such cases, and to avoid major disruptions to be considered within delays distribution, a filtering algorithm should be used.

This filtering should be performed using a model (e.g. Conte, 2008) to identify primary and secondary delays and then removing from the statistics all records corresponding to trains affected by major disruptions, and to trains which have consequently been delayed. In this study, a quicker and much simpler method has been used: the tail of each distribution is then shortened deleting all sparse values which are separated from the rest of the distribution by more than 1'. Such values are deleted from both the initial delay distribution and the statistics used for model validations, since it is impossible to reproduce disruptions, unless they are manually inserted into the model.

An extensive analysis of delay distribution has been performed, to obtain the mean percentage of sparse values within departure delay distributions, which is normally between 3 and 5% for local trains and between 2 and 7% for long distance services.

Stop time variability

Stop time variability reproduces two important phenomena affecting trains at stops: the departure imprecision and the variability of dwell time. Departure

imprecision takes place when a train, which arrived early or punctual at a stop does not depart punctually although passengers are already on board; it is generally short, with mean values between 10 and 20 seconds on Regional trains and often by more than 30" on long-distance services. Dwell time depends on the number of passengers effectively boarding/alighting, on the presence of late-arriving passengers and on the trainset.

To be able to model stop time variability correctly, it is necessary to separate the mentioned phenomena while analysing real traffic data, which normally include only the arrival and departure times. Instead of using a more detailed stop time calculation model (Buchmuller et al., 2008), a simple method based on two assumptions, valid for services with frequency higher than 10', has been used:

1. For late-running trains, stop time has no departure imprecision and is formed only by dwell time.
2. On early-running trains, the stop time includes both dwell time and departure variability.

The Dwell time is simply obtained analysing only the delayed trains, then, a departure variability is calculated considering early-running trains.

The departure time within the simulation model is the maximum between the latest between the departure time (which considers departure variability) and sum of arrival time and stop time (with dwell time variability). Therefore, given the distributions, the model will automatically considers dwell time variability and departure imprecision.

4.3 Calibration of motion equation

In the conventional softwares, running time calculation is performed on the basis of standard values representing good conditions. Influences on running times, such as human behaviour, weather and rolling stock condition which are not deterministic, are not explicitly considered, but covered by standard recovery times.

Using deterministic train movement calculation it is impossible to obtain a precise modelling of real train operations, where not only initial and station delays, but also running time variability is important. To compute the running variability still using the conventional empiric equations, a *performance factor* is implemented in

the micro-simulation tools. It is expressed as a percentage of the maximum speed, or acceleration, which is used in practice and is variable with a defined random function which can be different for each train. To calibrate this parameter, GPS trackings recorded on board of trains between March and May 2008 have been analysed and compared to running time calculation. For the analysis train motion has been divided into three phases: acceleration, full speed and braking; no analysis has been performed regarding drivers' behaviour at restrictive signal aspects, since in the records no systematic braking for signal has been recognised, which could be used for statistics.

4.3.1 Acceleration

Recorded acceleration is in some cases higher than the one theoretically calculated with running time calculator, although it is normally significantly lower with a variability of more than 30% (75-105%). In Figure 4.3, some very different acceleration speed-distance curves for loco-hauled Commuter trains at departure in Monfalcone are represented. Records are compared with the output of running time calculation with different performance factors, as indicated in the legend. While the continuous blue and light blue lines feature a 95% full-speed performance, the dashed line has a constant 60% for both acceleration and full-speed.

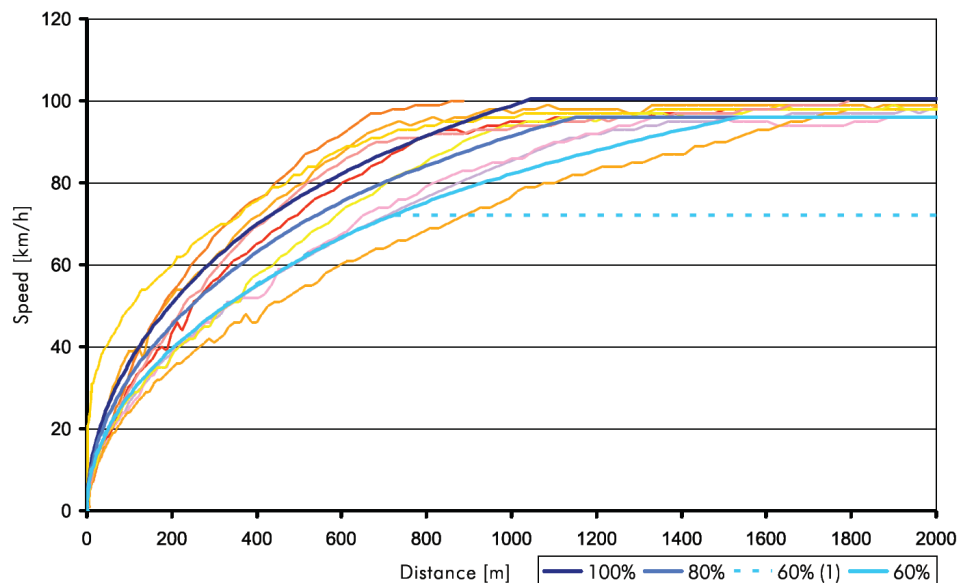


Figure 4.3: Recorded speed $v(s)$ profile at departure compared to calculated curves with different performance factors (blue and light-blue)

4.3.2 Full-speed behaviour

While during acceleration trains show quite variable performances, in most cases significantly lower than those theoretically possible, when running full speed their behaviour is radically different. The fine throttle regulation offered by modern inverter locomotives, such as the E464, allows all drivers to run at constant speed, just about 2-4 km/h below the limit. Within this range fall about 95% of records for line sections where the maximum speed can be kept for more than 1 km. It can also be noticed that the difference between the recorded speed and the limit is nearly independent of the speed on the analysed line.

Differently from what measured and suggested in other countries, such as in Germany, in Italy coasting before braking is normally not applied, also on early-running trains and when running times are much longer than necessary. In these cases trains normally have the same behaviour they have with a tight schedule, and the spare time is spent at the stops.

4.3.3 Braking

A variability even higher than in acceleration has been recorded in braking behaviour, with no train decelerating more than 90% of their calculated performance and less than 50% in more than 15% of records. Figure 4.3 shows some extremely different registered speed profiles at arrival in Monfalcone for loco-hauled Inter-Regional trains, compared with the UIC deceleration curves obtained as function of the Braking Weight Percentage used by the ATP.

It is not a simple task to decide with a simple analysis of the recorded trackings if the some very low deceleration rates are due to drivers imprecision in finding the best point to start the braking action or to the use of coasting on early running courses. Since fortunately such records are quite unfrequent (about 5% of the total) they can not be considered when calculating mean performances as input for a simulation model.

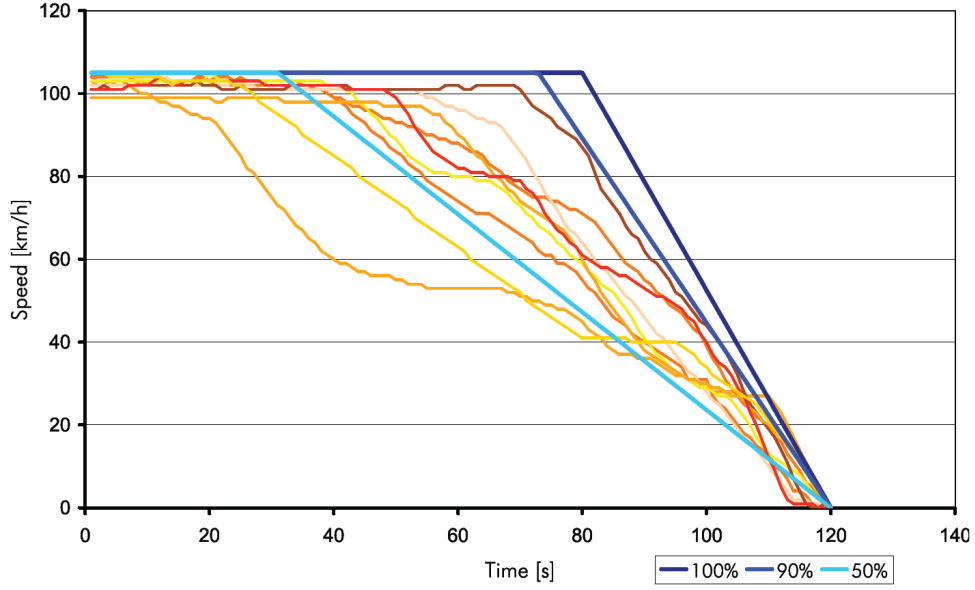


Figure 4.4: Recorded speed profile $v(t)$ at arrival compared to calculated curves with different performance factors (blue and light-blue)

4.3.4 “Global” calibration

Since only one stochastic performance parameter is provided, a simplified calibration procedure has to be used, where some average values have to be used for all motion phases. The resulting error has to be considered as model error and estimated when model calibration is performed. In this “global” calibration, total running time distribution between two stations is compared to running time calculation.

For each motion phase the same performance factor is used, obtaining a proportional reduction of the speed (at full-speed behaviour) or of the acceleration/deceleration. Once the minimum running time ($T_{100\%}$) is calculated, the distribution of p as a piecewise-linear function for each i of the n intervals of the aggregated distribution is

$$p_i = \frac{T_{100\%}}{T_i} \quad (\text{for } i = 1 \dots n) \quad (4.1)$$

It must be noticed that to capture train drivers behaviour and not to reproduce train traffic variability as a whole, all trains subject to route conflicts or having other hindrances for their movements have to be excluded from the set of data to be

analysed. This operation can be performed graphically by checking on the aggregate speed vs distance diagrams recorded on-board the presence of unplanned braking actions and excluding those records. In Figure 4.5, on-board recorded data are compared to running time calculations using different performance percentages. The difference in the intermediate stop is due to the distance between the simulated and real stopping point.

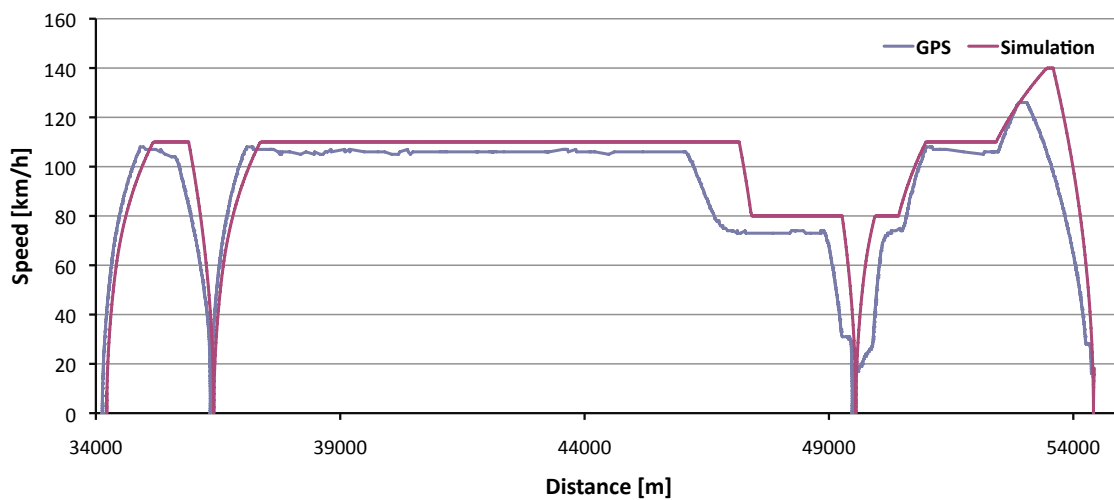


Figure 4.5: Speed vs distance diagram for a simulated and real train with the same running time

If on-board data are not available, an alternative is given by the use of station arrival and departure times, which are simpler to obtain, but require a special attention. Secondary delays have to be prevented, for example considering the first or last course of a day, where no traffic conflict should occur, and a check of the timetable graph with aggregate traffic data could point out possible conflicts, for example due to very late-running freight trains or on days with very bad weather conditions. In Figure 4.6 the aggregate timetable graph of a high-speed line, with large headway between trains: nevertheless, on some days with heavy snowfall conflicts due to very slowly running trains could not be prevented.

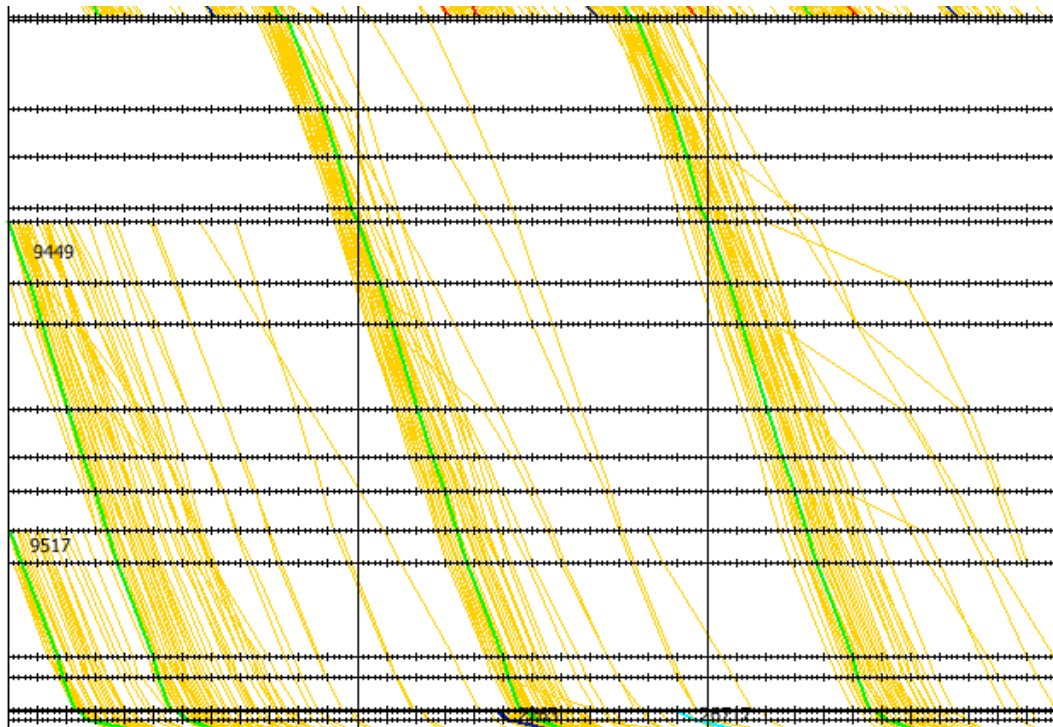


Figure 4.6: Aggregated timetable graph of a high-speed line (one direction) late in the evening.

If no time window and line section can be found that can be reasonably considered conflict-free, station data should be analysed with an algorithm to point out secondary delays, dramatically increasing the difficulty of this operation, otherwise very simple.

Since train drivers behaviour can be considered as nearly independent of the simulation scenario, and therefore it is not expected to vary as simply as changing the parameters of a theoretical distribution, it is not necessary to perform a fit of the obtained distributions.

Calibration aims at obtaining reasonably precise running times, but a model error still exists. This error in motion calculation can be divided into two levels:

a) **Trains used for calibration.** The difference between simulation and real data is clearly shown in figure 4.5. With the calibration, the performance factor distribution to obtain the real running times between station A and B is calculated, using a single parameter to represent the slightly different behaviour in the different motion phases. Therefore, at intermediate points, the difference between real and simulated timings can be significant.

b) **All trains of a simulation set.** Performance factor calibration is performed only for a small number of trains, whose results are then considered valid for all trains of the same family. All trains are expected to have the behaviour of off-peak-hour trains, supposing that the driving style is not influenced by traffic, number of passengers at stops and other variables.

4.4 Issues of synchronous simulation

The infrastructure, the interlocking and safety systems as well as train movements are represented in high detail in micro-simulation, theoretically leaving a (small) model error only considering the stochastic behaviour of trains, which is not perfectly represented using conventional equations, and using the described simplified methods to calculate initial delays and stop time variability. Nevertheless, if the error is lower than 5% for each train, synchronous simulation models show significant weaknesses in managing railway traffic.

It is possible to point out these weaknesses in the dispatching rules, in their effects on mixed-traffic lines and in event of platform change at terminal stations.

4.4.1 Dispatching rules

In micro-simulation models, no traffic management system is provided, which can reschedule trains on the basis of the current traffic status. Instead of this realistic control algorithm at network level, dispatching is based on the resolution of local traffic conflicts based on a weighted FIFO (First in-first out). Each train features a so-called pre-reservation distance (and time), which means that it can reserve a block section a given time interval before passing the corresponding distant signal. The pre-reservation distance is a function of the train category and delay, whose combination is usually realistic in modelling priorities. Therefore conflicts are solved on the basis of the pre-reservations and hence of train priorities, as established by railway legislation in many countries. But the fixed pre-reservation parameters do not fit well with the very different rules of thumb used by traffic controllers, who are able to foresee the consequences of their decisions.

For example, if a delayed train with low priority has stopped at a junction that is quite busy, it will not be able to enter the main line until on both tracks no train has

pre-reserved a route. In the corresponding real situation, the delayed train could probably pass immediately after the train he stopped for.

4.4.2 Complex dispatching on lines

The lack of a higher-level traffic control is particularly clear on lines where real-time dispatching systematically modifies the planned timetable in order to minimise delays. This situation can be found, for example, on high-density, mixed-traffic lines, where freight trains, and often also other low-priority passenger services, are continuously rescheduled, moving the overtakes to minimise delay for long-distance services.

Some extreme examples of such traffic could be found in the conventional lines in Italy before the corresponding high-speed lines entered into service. For example, between Bologna and Firenze the double-track "*Direttissima*" line was used until December 2009 to carry all the traffic between most part of northern Italy and Rome. To allow a higher flexibility in traffic management, overtakings are not fixed, but continuously moved to the best location, chosen by controllers thank to their experience. Moreover, the time-loss for them is often not concentrated at one station as stop time, but inserted as running time supplements among three consecutive stations, forming the so-called "*precedenze volanti*" (*flying overtakings*): the location of overtakings is autonomously decided by the controllers.

The planned and simulated timetable for the central part of the Bologna-Firenze conventional line is presented in Figure 4.7. A variety of differences between the deterministic simulation and the planned timetable can be found, as well as the "flying overtaking" of train 6599 in Vernio. In the example, the simulator is able to solve all conflicts and avoid nearly all time-losses for the green trains (high-priority Eurostar Italia services).

While in the deterministic example the simple dispatching routines can be found, in stochastic simulation the algorithm is not always able to avoid conflicts for high-priority services, and can also lead to unrealistically high waiting times if a train has to be overtaken and the overtaking tracks are already occupied. In Figure 4.8, the same timetable as in figure 4.7 is simulated under perturbed conditions but with the same dispatching rules. All rescheduling measures are correct, except for the highlighted 9312, which does not overtake the preceding freight trains.

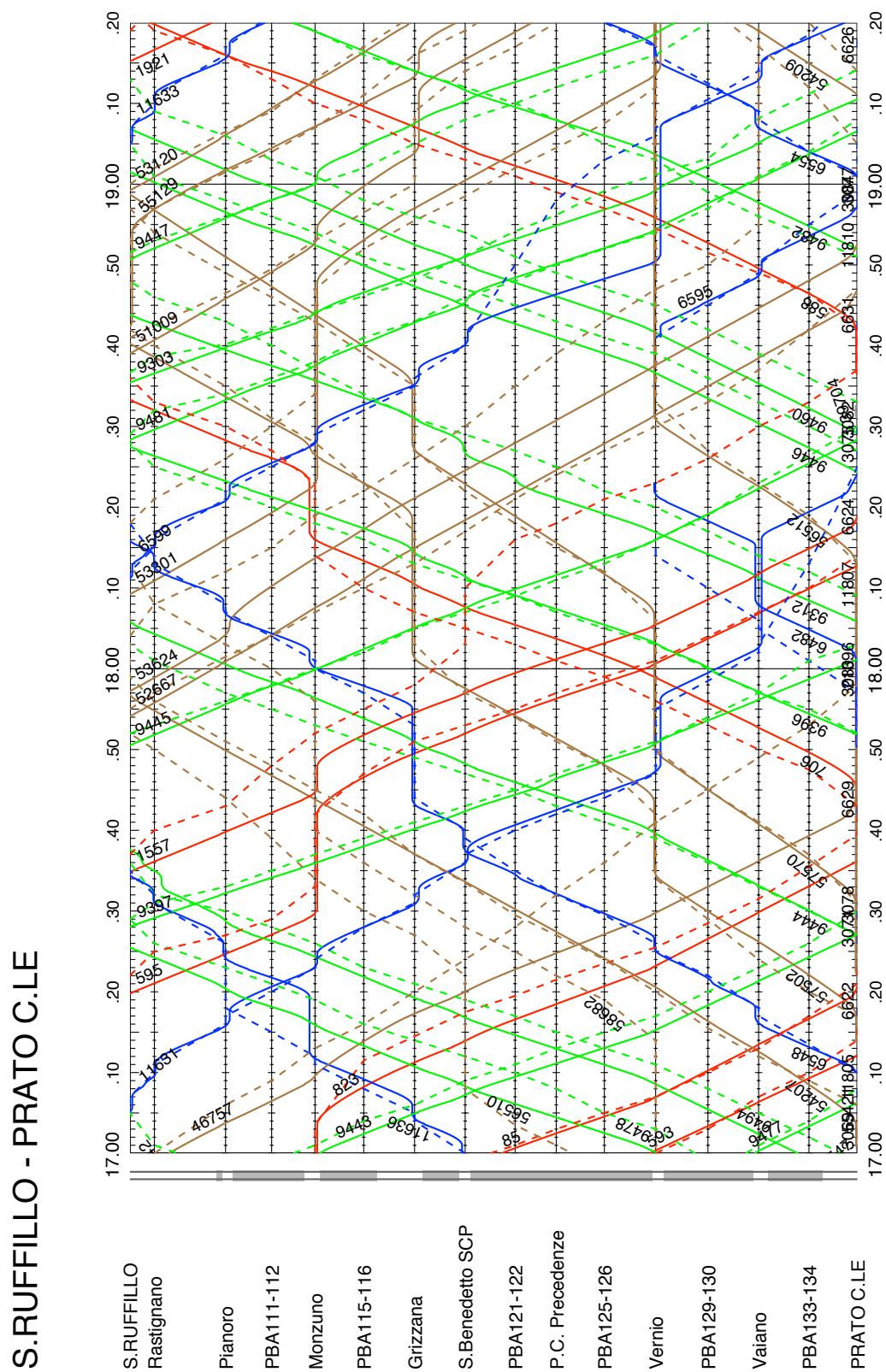


Figure 4.7: Simulation of a high-density, mixed-traffic line: differences between planned and simulated timetable.

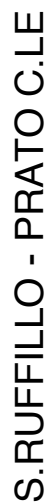


Figure 4.8: Simulation of a high-density, mixed-traffic line: time-losses for a high-priority service, highlighted in green.

4.4.3 Itinerary change and platform allocation

Another weakness in traffic management can be found in the choice of alternative itineraries. When a conflict is found, the simulator searches in a pre-defined itinerary list (with decreasing priority) for an itinerary which is still free and blocks it. The itinerary is changed even if the pre-defined one is released few seconds later and if other conflicts are caused by the new choice. Although this method to assign the itineraries is not realistic, the experience gained with a wide range of topologies shows that normally it does not lead to deadlocks nor to a significant increase in delays.

Not realistically representable with the existing algorithms is the platform change, since it works with the above-mentioned principle used to change the itinerary. The platform change forces all passengers to move, often just some minutes before train departure: therefore it is unfrequent in many european countries, but normally used in the largest terminal stations in Italy, such as Roma Termini and Milan Centrale. About 30% of trains in Roma Termini do not arrive/depart at the scheduled track: changes are decided by traffic controllers either to solve route conflicts or because the scheduled track is already occupied. Since turnover time is often quite long for long-distance services, the platform change of a train causes another change to a second train, scheduled to used that track. As a result, during the day, a high number of platform changes follow others in a sort of propagation.

No exact rule exists that leads to decide a platform change, and in the simulation it can be either performed anyway if the pre-defined track is occupied or totally deactivated. As a result, the pre-defined tracks are used anyway or tracks are very often changed unrealistically. Case studies have shown that using the pre-defined platforms realistic results in terms of delay per train are obtained, although the real track allocation process is not reproduced.

4.4.4 Strongly delayed trains

Strongly delayed trains represent a particular challenge for dispatchers, since a *de facto* new slot has to be inserted, also within nodes and large stations and with the lowest impact on other trains. Moreover, if such trains are freight services, overtakings have to be managed. If they are long-distance passenger services, platform have to be allocated at terminal stations, and it is necessary to decide wether the vehicle maintains the planned turnover (eventually delaying the course it will be used for) or not.

From the simulation point of view, the problem appears as a combination of the above-explained issues regarding complex dispatching on lines and platform allocation. With the look-forward rules, it is impossible to ensure that simulation represents such phenomena without generating high waiting times for other services. Therefore, on complex topologies combined with strong delays, simulation has to be tested to evaluate its accuracy, which may be significantly lower than in the calibration case studies presented in this thesis.

4.5 Calibration: Case Studies

Micro-simulation represents railway traffic in very high detail; nevertheless, some weaknesses can be found, which could limit the accuracy of micro-simulation, as explained in detail in section 4.4.

To estimate the difference between simulation and real traffic, and evaluate the possibility to use micro-simulation for ex-ante reliability studies, a number of case studies have been developed.

First, the Venice - Trieste line has been simulated to evaluate the impact of an increasing precision in the representation of stochastic phenomena (and therefore theoretically in model calibration) on the accuracy of representation of the real traffic.

Second, the Palermo - Punta Raisi single-track line, systematically affected by high delays has been simulated, in order to evaluate precision in the representation of the dispatching rules that are used in reality to reduce delay propagation. In a further development, the calibrated model has also been used to evaluate the impact on punctuality of timetable changes.

The third case study was focused on Turin node. After calibration, the model has been used as an ex-ante punctuality estimation tool, in presence of significant restrictions due to infrastructure works and of a new timetable. Simulation results for this scenario have also been compared with real traffic data after timetable change.

4.5.1 Venice – Trieste

In a first case study, calibration was applied to the Trieste - Venice line, to evaluate the impact of the different stochastic parameters and the benefits of the

use of GPS instead of station passings to estimate the distributions of performance factor. The double track, electrified line is about 130 km long and plays an important role in regional transport and as freight corridor between eastern Europe and Italy. Figure 4.9 shows a simple layout of the line; the analysed network was extended to approximately 30 km on each branch line, obtaining totally 280 km line. On the line, the impact of dispatching decision is very limited, since only two pairs of high-priority passengers trains are scheduled, most freight trains run during the night and during the day most services are Regional on following routes, with a regular-interval timetable:

- Venice - Trieste (Inter-Regional)
- Venice - Udine (Inter-Regional)
- Venice - Udine (Regional)
- Venice - Cervignano - Udine (Regional)
- Venice - Portogruaro (Regional)

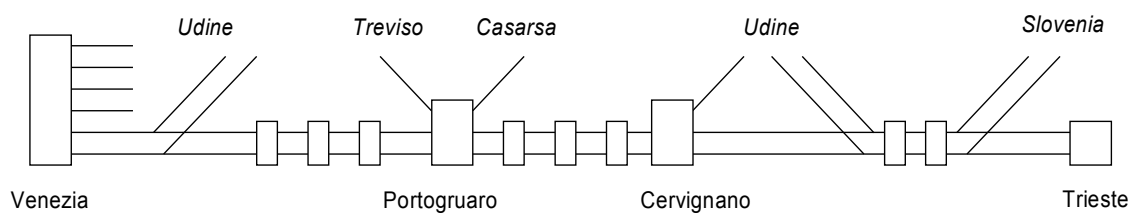


Figure 4.9: Simple layout of the Venice – Trieste line layout.

After performing the macroscopic analysis, the study focused on evaluating behaviour of trains in non-disturbed conditions, to estimate a standard performance percentage and a piecewise linear braking function for every train type. Performance variability was also defined as piecewise linear by fitting it to real distributions, as well as initial delay for every train family. GPS speed-time data were used for testing purposes. Stop time, represented by a mean time and a variability range, as defined in RailML standard, was automatically exported for every train and station.

The results of 100 stochastic simulations were analysed and compared to real data referred to 100 days. Different simulation scenarios were considered with growing calibration precision to point out the importance of various distributions and of micro and macroscopic analysis. In the first scenario only departure delay as negative exponential function per train family was considered. Then, in the second scenario, performance variability as empirically evaluated by timetable planners was

added. In a third phase stop time variability was included and in a fourth one departure delays were modelled using a piecewise-linear function.

Scenario results were compared using three simple punctuality parameters, similar to quality indicators commonly used in performance measurements. For each train, the differences among the percentage of trains arriving at each station within 1, 3 and 5 minutes and the mean delay in real and simulated data were calculated; the mean differences on all trains and for each indicator are shown in table 4.1. In the second column, "neg.exp." indicates that a negative exponential distribution was used to model the initial delays, while "pie.lin" indicates that real distributions as piecewise linear functions were inserted into the model.

Finally, in Scenario 5 the performance factor was calibrated using GPS-collected data. The approach was limited to a train type that represents about 3/4 of all passenger trains running on the line. Since, as already explained, the simulator considers only one "global" performance factor, a more realistic behaviour of trains was obtained modifying the acceleration and braking behaviour of trains, to emulate registered variability:

- the performance factor was calibrated using only full-speed behaviour on the open track.
- a second "virtual" parameter was introduced and calibrated, to let the simulator obtain realistic acceleration variability. Distribution of this second parameter was aggregated to 5 intervals. For each acceleration interval, a new model train was created, modifying the tractive effort of the original one to obtain the corresponding acceleration.
- a similar procedure was applied to the braking phase, modifying the braking deceleration of each of the 5 trains already created.
- the five model trains were randomly assigned to each course on the line and for each simulation day.

A significant precision increase was obtained in each scenario, showing that:

- negative exponential initial delay can be used only as first step;- Departure variability at all stations is the most important parameter for stopping trains; since it is often very different for various services and train types, the automatic import function is very useful;
- running time variability is important for long-distance and freight trains;

- on-board collected data effectively allow a better estimation train performance parameters; even better results could be obtained by a re-calibration of some more specific resistance factors, or with the adoption of different performance factors for each motion phase.

Scenario	Stochastic Parameters			GPS calibrated Motion	Results		
	Initial Delay	Running Time	Stop Time		1' [%]	3' [%]	Mean [s]
1	Yes (neg. exp)	No	No	No	14.1	11.3	68
2	Yes (neg. exp)	Yes	No	No	12.8	9.5	59
3	Yes (neg. exp)	Yes	Yes	No	9.8	7.2	51
4	Yes (pie.lin.)	Yes	Yes	No	7.2	5.3	40
5	Yes (pie.lin.)	Yes	Yes	Yes	5.3	2.9	22

Tab1e 4.1. Comparison between stochastic simulations. Calibration parameters source:
Scen. 1-4 timetabling points, Scen. 5 Collected data.

4.5.2 Palermo – Punta Raisi

The Palermo - Punta Raisi single-track sub-urban connection between the city centre and its airport has been simulated as second case study, aiming at evaluating the behaviour of microsimulation on a single-track line with high traffic density. A brief description of the line and the delay analysis of March and April 2008 are presented in section 3.1.

The challenge in the simulation of this simple line was represented by the very high delays of services coming from Cinisi. If the first trains in the morning are strongly delayed, their delay is propagated also to trains in the opposite direction, and changes in the location of crossings are not unfrequent. In Figure 4.10 is depicted the arrival distribution in Piraineto, at the beginning of the single-track section, for the first three services Cinisi - Palermo in the morning. Moreover, between Notarbartolo and Cardillo Zen the line passes through the densely inhabited suburbs of Palermo, with at-level crossings on important streets where trains are often forced to slow down.

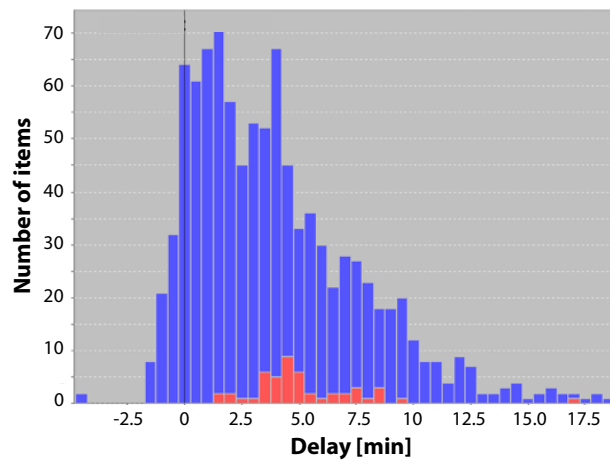


Figure 4.10: Departure distribution in Piraineto for all Cinisi - Palermo services.

Fortunately, despite the frequent arrival delays, departures both from Palermo Centrale and from Punta Raisi are quite punctual, thanks to long turnover times. Therefore, and since the first and last sections of the line are double-track, the real departure delay distributions can be considered the initial delays to be used in simulation. More complicated appeared the calibration of stop time variability, which is not only dependent on dwell time and time-loss due to crossings, but also by the at-level crossings, which in some stations force trains to longer stops.

Up to 200 multiple simulations with the same parameters were performed in order to estimate the number of simulations that are necessary to obtain reliable results. Given the results of 200 simulations, the difference among them and 40, 80, and 120 simulations were measured in terms of reliability indicators of selected trains. Results show that, on this line, 40 simulations are sufficient to obtain a difference inferior to 6% in those indicators with respect to 200 simulations, while with 80 this gap falls down to about 2%.

The model was validated comparing the mean delay and the 5' punctuality as well as the delay threshold of the 80% of trains on 5 selected stations along the line, for all trains on each direction between 7.00 and 14.00. Average indicators in three out of 5 stations are presented in Table 4.2. The results show that simulation normally overestimates delays in the first and last sections of the line, while they are lower than in reality in the central part. This is probably due to the distribution of departure imprecision, which was estimated using the simplified approach and considers also the effect of at-level crossings within the station area. Since a more detailed estimation of the phenomena at each station was not possible with the available

data (station passings), average distributions were used, which obviously do not consider each specific situation in detail.

In an overall comparison between simulation and reality, differences by less than 3% in punctuality, 30 seconds in mean delay and in the 80% threshold were obtained. The result can be considered quite satisfying, in particular because the difference between simulation and reality are distributed along the corridor, showing that the most relevant stochastic phenomena at each station were reproduced.

From Palermo		Real	Sim		Real	Sim	To Palermo
Palermo Notarb.	Mean	69	97		90	79	Piraineto
	5'	94	93		96	96	
	80%	90	128		120	118	
Piraineto	Mean	247	270		239	193	Capaci
	5'	70	73		69	74	
	80%	420	400		390	348	
Punta Raisi	Mean	76	104		194	227	Palermo C.le
	5'	85	82		79	80	
	80%	240	279		330	345	

Table 4.2: Real and simulated reliability measures.

The validated model was used to estimate the impact of speed restrictions on some line sections due to important infrastructure improvement works, which were scheduled to start in June 2008. Simulation results prognosticated a decrease of punctuality due to works by about 11%.

To improve service reliability during infrastructure works, simulation was used to point out weaknesses in the timetable, and to modify it without increasing travel times, that means only reallocating running time supplements. In 2008 timetable they are concentrated in the last sections, with no possibility to recover delays in the urban part of the line, where stop and running time variability is higher due to at-level crossings. The timetable was modified, considering a punctuality-goal of about 93%. After a presentation of the modifications to RFI, it was decided to increase running times, in order to obtain an even higher reliability. Reliability Indicators for the three scenarios are displayed in Table 4.3; "12/07" indicates the simulation of 2007

timetable in presence of restricted speed, used as comparison, “Mod1” indicates the first changes to the allocation of margins and “Mod2” the timetable that entered into service. The planned timetable of the first and the second scenario is also depicted in Figure 4.11.

This improved timetable entered into regular service in June 2008, but the civil works could not start due to technical problems. With a timetable designed for much stricter conditions, the overall 5’ punctuality reached a surprising 99%, representing one of the best values in Italy.

From Palermo		12/07	Mod1	Mod2		12/07	Mod1	Mod2	To Palerm
Palermo	Mean [s]	192	144	90		234	66	102	Piraineto
Notarb.	5’	82	89	94		86	96	96	
Piraineto	Mean [s]	312	138	78		216	144	0	Capaci
	5’	58	87	95		76	95	98	
Punta Raisi	Mean [s]	108	66	72		216	124	56	Palermo C.le
	5’	78	88	94		76	91	95	

Table 4.3: Comparison between 12/07 and modified timetables.

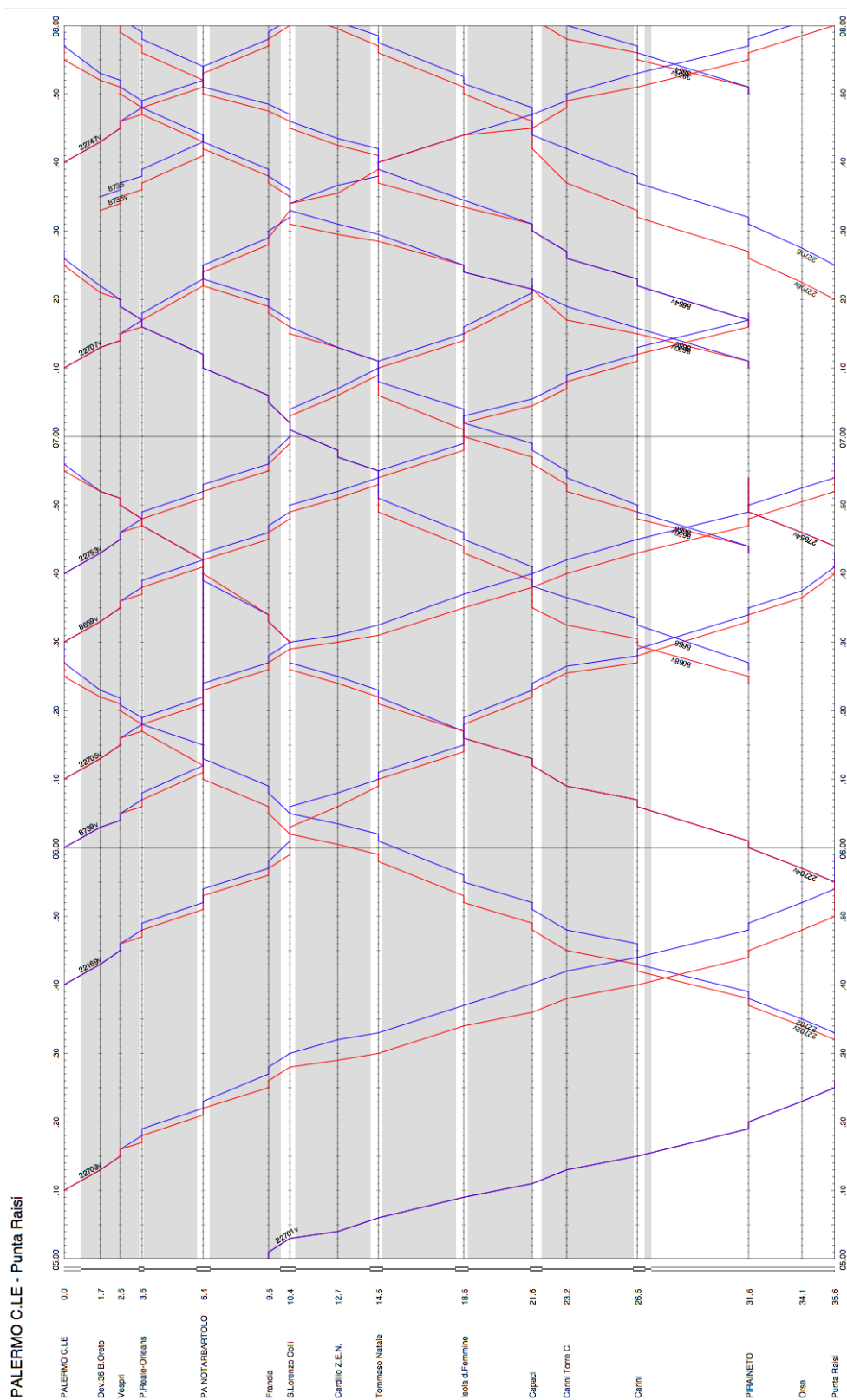


Figure 4.11: Comparison between 12/2007 and the modified timetable.

4.5.3 Turin Node

The node of Turin, a mid-size node in the second largest city of Northern Italy, was simulated in a third case study, aiming at evaluating the potential of simulation in a more complex network. A brief description of the node as well as its layout are presented in section 3.1.

The potential hindrances for an accurate representation of the real world with the simulation are represented by the large terminal station Turin Porta Nuova and by the traffic management rules used in event of delays. These are important in particular between Trofarello and Lingotto, where the four tracks are used in a flexible way to reduce traffic conflicts, and between Chivasso and Turin, with different train categories and priorities and frequent delays especially for trains coming from the branch lines.

Departure delay of each train bound for Turin at the first station outside the node represented in Figure 4.12 were used as initial delay. On the contrary for services from Turin Porta Nuova, the automatic import was not possible, since the departure delays from the terminal station also incorporate secondary delays, as already noticed in section 4.4.3.

To estimate the impact of secondary delays on train departures in Turin Porta Nuova quickly, standard negative exponential distributions were used for all trains. Simulated departure delay were compared to the input delay, obtaining the secondary delays, which were about 20% of the total amount of the simulated departure delays. This estimation was considered valid also for the real distributions: Real departure delays in Turin Porta Nuova therefore were reduced by 20% and used as initial delays to validate the model.

The model was tested and validated comparing the F and P indexes for each train family and line section. An average 3.4% error has been registered, using calibrated parameters and using 15-second block release time to compensate a difference between minimum real and simulated headway time. Calibration results by train family on the Turin - Chivasso line are shown in Table 4.4.

Train Family	Train Cat.	From	to	Real		Simulated	
				P %	F	P %	F
601n	IC	Turin	Milan	94,4	36,8	97,7	32,4
2001n	IR	Turin	Milan	75,6	65,9	76,4	67,1
4001n	R	Turin	Chivasso	86,1	51,7	83,6	55,2
9301n	HS	Turin	Milan	90,4	36,2	93,3	31,4
20001n	R	Trofarello	Settimo	74,8	81,2	79,2	84,2
600n	IC	Milan	Turin	76,3	67,1	78,5	62,3
2000n	IR	Milan	Turin	85,7	52,9	82,3	56,6
4000n	R	Chivasso	Turin	84,7	56,2	88,3	54,2
9300n	HS	Milan	Turin	72,7	59,3	76,2	53,2
20000n	R	Settimo	Trofarello	87,8	49,3	92,6	42,3

Table 4.4 Comparison of reliability measures for real and simulated train families

The calibrated model was used to simulate operations after the timetable change in December 2007, when the tunnel between Lingotto and Porta Susa was closed to allow major infrastructure improvement works (Figure 4.12). This had heavy consequences on the whole node, since a higher number of trains was scheduled to pass the Quadrivio Zappata at-level double junction, not only with movements to and from Porta Nuova, but also from Lingotto. The timetable was adapted to the modified infrastructure, but only few trains were withdrawn, leading to the saturation of Quadrivio Zappata. Therefore simulation was performed, paying special attention to delay propagation at Quadrivio Zappata.

The simulation output was an average $F=56$ ($P=92\%$) in the peak hour between 6 and 9 am: not only delay propagation appeared to be limited, but also punctuality was at the same level of 2007 timetable, using the same input distributions.

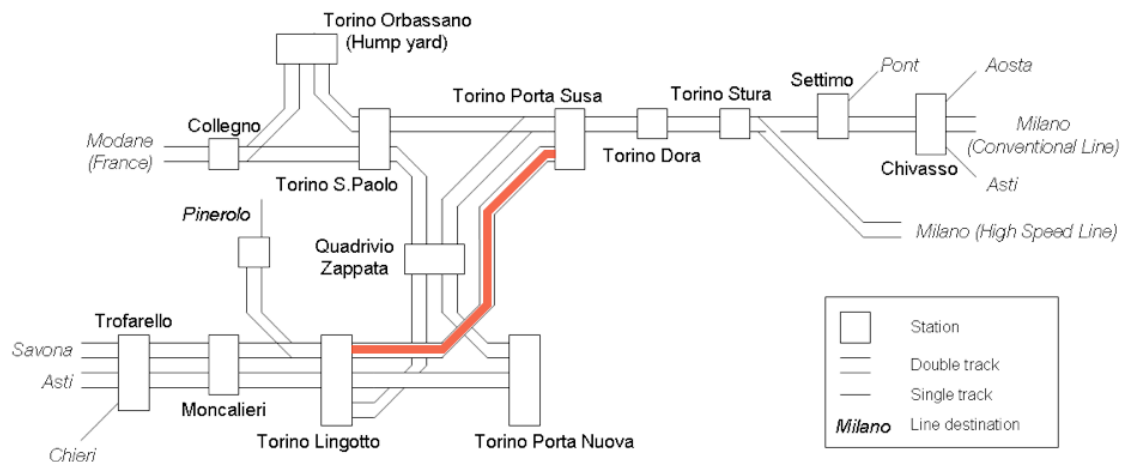


Figure 4.12: Simple Layout of Turin Node. in red is highlighted the Lingotto -Porta Susa city tunnel

After the timetable change, simulation results were compared with real ones, showing unexpected higher delays especially on the Chivasso - Turin line, the most busy and important-one . Punctuality dropped down from 89 to 68%, with significant delay propagation in Quadrivio Zappata.

A detailed delay analysis showed systematic perturbations involving trains between Milan and Chivasso. These caused propagation in the node extended to most lines and the opposite direction. Between the 10th and the 22nd December, an average $F = 140$ was measured (Fig 4.13).

To verify the validity of the model, the registered initial delay distributions were used as input, obtaining only a 4.2% mean error.

After some weeks, traffic precision rapidly increased again in January 2008, with punctuality rates similar to winter 2007 as forecasted by the model.

In picture 4.14 the mean delay of 2008 timetable is depicted for 4 train families on the Chivasso – Turin line as a function of the mean initial delay. Real delays in December 2007 and February 2008 are compared to simulations using winter 2007 and December 2007 initial delay distributions. Intermediate scenarios between these extreme values were simulated using proportionally defined delay distributions. Real records are represented with black symbols, while the coloured ones indicate the simulated scenarios.

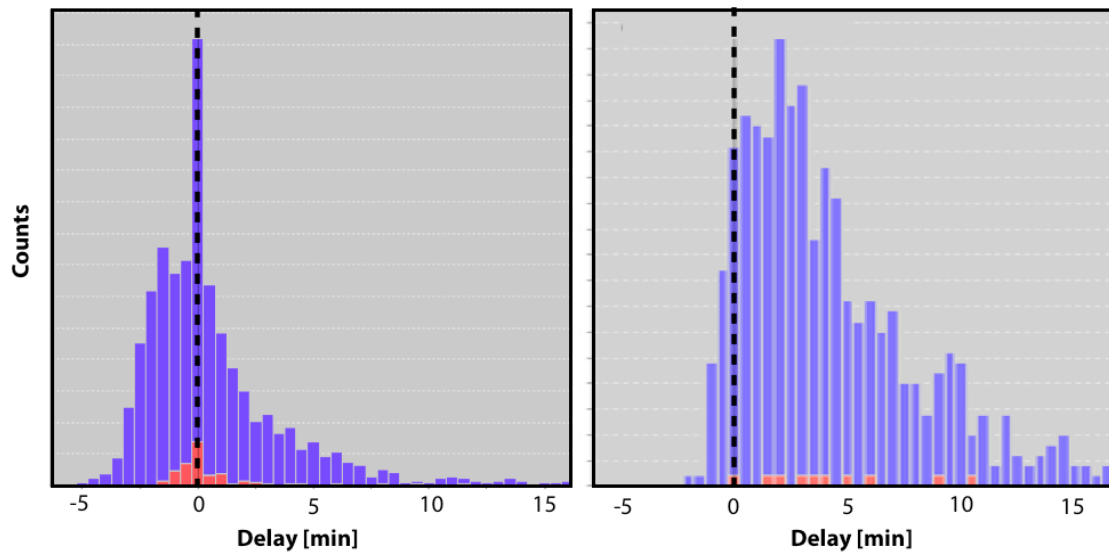


Figure 4.13: Departure distribution in Chivasso for all Chivasso - Turin Regional services, in november 2007 (on the left) and after timetable change in December 2007

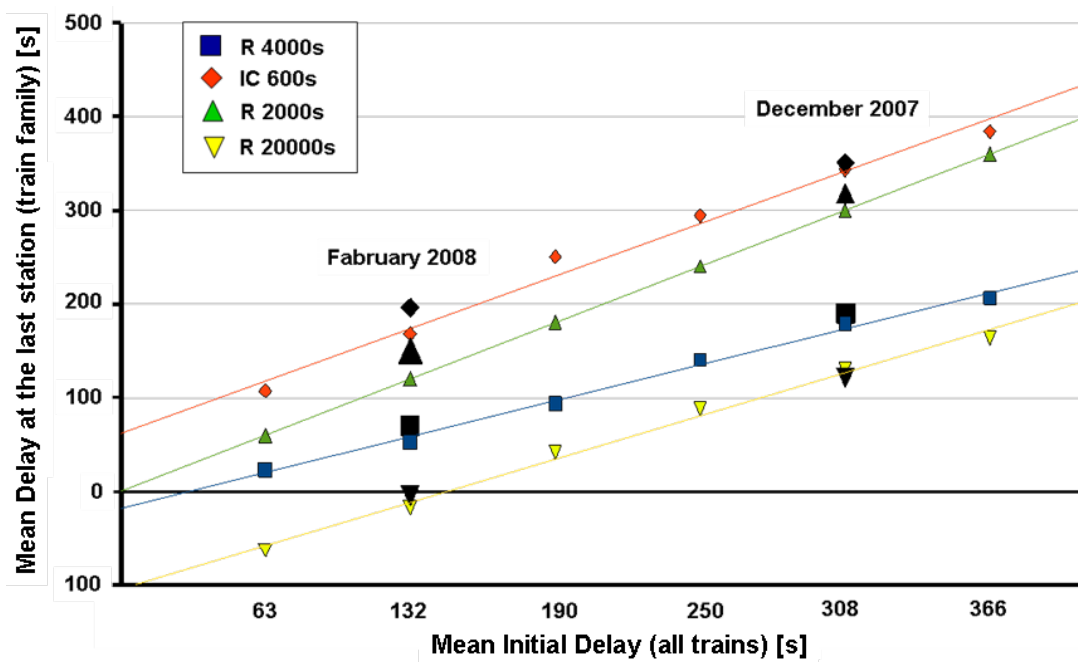


Figure 4.14: Simulated and real F for different train families compared to real data.

*Oh quei fanali come s'inseguono
accidiosi là dietro gli alberi,
tra i rami stillanti di pioggia
sbadigliando la luce su 'l fango! [...]
Già il mostro, conscio di sua metallica
anima, sbuffa, crolla, ansa, i fiammei
occhi sbarra; immane pe 'l buio
gitta il fischio che sfida lo spazio.
Va l'empio mostro; con traino orribile
sbattendo l'ale gli amor miei portasi.*

(G. Carducci, Alla stazione in una
mattina d'autunno, 1877)



The trade-off between capacity and reliability

Carrying the maximum number of trains with a satisfying reliability level is a fundamental goal for an infrastructure manager, since it allows to divide the high maintenance costs over a higher number of slots.

To reach this objective without improving the signalling and control systems a key role is played by the timetable, which could first avoid conflicts and then recover delays and minimise delay propagation.

In many companies, great attention is paid to the calculation of the minimum headway times, which can also be viewed as occupation stairways directly in the timetable planning software. The use of such tools, which are based on a microscopic infrastructure, allows the construction of conflict-free timetables, reproducing each train movement also within stations.

Less frequent in practice, although already investigated in literature, is the use of quantitative approaches to the determination of running time supplements and buffers, which have direct influence on both capacity and reliability of the system.

This is mainly due to the satisfying results that are currently obtained thanks to the planners' experience, and to the objective difficulty in developing a close formulation that enables to determine the buffer time and the supplements to be used in each situation. Moreover, while the proven rules of thumb are applied by planners to a wide variety of cases, it is not simple to find what parameters should be considered in such a model to be easy to use and precise at the same time.

To consider all the variables involved in real traffic and combine them with supplements and margins, a calibrated micro-simulation model could be used, since it allows a precise real-world representation, already tested under very different conditions. In particular, a simple procedure was developed, to enable a standardised evaluation of the sensitivity of the system to a wide range of factors.

The methodology is explained in the first part of the chapter, followed by its application to two case studies. In the third part, the main results referred to the case studies are reported and discussed. Finally, the results obtained using micro-simulation and using stochastic differential equations by Stok (2008) are briefly compared.

5.1 Estimation of the trade-off between capacity and robustness

The basic element for any estimation of the service reliability using micro-simulation is the timetable, which is requested by the tool, but also contains the train-mix, on which the capacity will depend. To evaluate the impact of different factors and the maximum capacity, given the train mix, it is necessary to start creating a dense timetable.

5.1.1 Creating a dense timetable

A dense timetable is basically a timetable where no buffer time is inserted between slots, and no running time supplement is provided. Obviously, since different train services have different running times on a line, free "space" is still present, for example after departure of the slowest train. But between each couple of services, there is at least a block section, where the reservation of the second trains starts immediately after the first has released it (Fig 5.1). When a delay occurs, for

example because a delayed train enters into the considered network, or because of a longer stop at a station, it is propagated to the following train, and not recovered until the train reaches its final destination.

The dense timetable shows the maximum capacity of a line section with a given train mix, and therefore represents the extreme bound for any reliability measure. Since no delay compensation is possible, the output delay of a simulation is significantly higher than the input, the highest possible with the given train-mix.

Since absolutely no supplement has to be provided, no rounding has to be inserted, for example in the departure time; the dense timetable is simply constructed defining the departure time at the first station and the stop time for each intermediate halt. Once the train order is defined, and the occupation stairways for each train-type are calculated, they are moved close to each other, checking that no train is shown a restrictive signal aspect by simulation.

The procedure is similar to the *compression* suggested by UIC within the constructive method for capacity calculation (UIC, 2004), but there are some fundamental differences. Unlike in the dense timetable, where the simulation tool has to be already calibrated to incorporate all reservation and release times, and absolutely no buffer time has to be inserted, the UIC recommends to insert buffer times, and other so-called “indirect” occupation times.

Once the dense timetable is constructed for the first series of trains, and contains the train-mix and order realistically, the departure time of each train (D_i^d) represents the *basic period* of the complete timetable, which is not different from the *interval* in regular interval timetables. The period is then repeated, ensuring that no buffer time is provided between the last train of a period and the first of the following one. Obviously, the obtained interval has a specific value, which is normally not as simple as 60 or 30 minutes and is not realistic to operate such a timetable; the *minimum period* (P_{min}) is defined as the time interval between two following trains of a given train family in a dense timetable.

Simulated station passings and departure times are used as planned timetable for all timetabling points.

As a result, the capacity bound is simply obtained as

$$C_{\max} = \frac{N}{P_{\min}} \quad (5.1)$$

where N is the number of trains in the period

Buffer times and running time supplements can now be inserted in the timetable.

5.1.2 Buffer times

To insert buffer time, the headway time between two trains is increased. The departure time of the first trains remains fixed, while the entire timetable of the second train is delayed by the *buffer time* T_b . The buffer time is added before the third train train delaying it by the doubled buffer. For each i train in the period, the departure time is:

$$D_i = D_i^d + (T_b \times (I_i - 1)) \quad (5.2)$$

where I is the position of the train within the interval (e.g. for the 1st train $I = 1$).

The corresponding time interval P_b becomes:

$$P_b = P_{\min} + N \times b \quad (5.3)$$

and the departure time for each train of the n th period:

$$D_{i,n} = D_i^d + (T_b \times (I_i - 1)) + (n \times N \times b) \quad (5.4)$$

It can be noticed that the procedure can be simply automated once the dense timetable is defined; timetables are then quickly modified to generate different scenarios, for example with increasing buffer time.

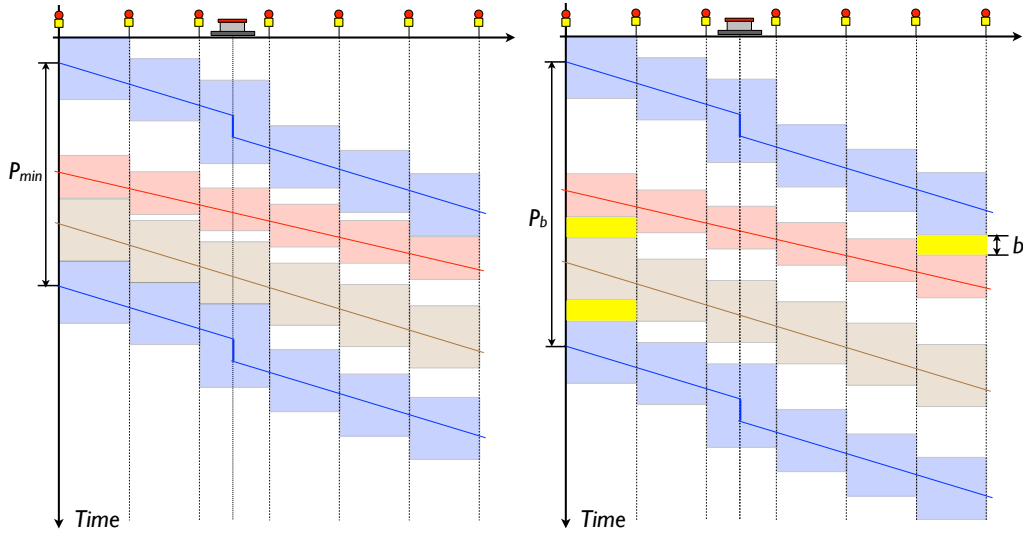


Figure 5.2: Dense timetable (on the left) modified by inserting buffer times.

5.1.3 Distributed running time supplements

Distributed running supplements are inserted along the train route to compensate the primary delays caused by the different driving styles of drivers, by longer dwell times at station and other phenomena.

Distributed running times cause an increase in running time between two timetabling points which is directly proportional to the minimum running time itself. Therefore, distributed supplements are normally inserted as percentage of the total running time (UIC, 2000). However, since the total running time is not the same for all services, a proportional supplement would result in a number of conflicts in the dense timetable. If we consider a regional train, which has a longer running time compared to the intercity service that follows it, this will also have longer supplements. This causes a conflict in the last block sections, where the headway time between the trains is the minimum headway.

To overcome this problem, maintaining a contact with the UIC leaflets and the practice, distributed running time supplements are first inserted as percentage $S_{d[\%]}$ on a single course: the resulting total amount at the last station S_d is then taken as total supplement for each other course. Considering a reference course r , its running time is obviously:

$$T_r = T_r^{\min} \times (1 + S_{d[\%]}) \text{ and consequently} \quad (5.5)$$

$$S_d = T_r - T_r^{\min} = S_{d[\%]} \times T_r^{\min} \quad (5.6)$$

and for each other course i :

$$T_i = S_d + T_i^{\min} \quad (5.7)$$

Seemingly, all station passings are then proportionally modified for each train. It is also possible to define other rules to assign the distributed supplements (e.g. Vromans, 2005); however, this has not been performed yet.

The procedure can be considered realistic for trains running the same distance; thus it leads to unrealistic distributions if the same supplements are applied to significantly shorter services. In fact, on such courses the supplements would be proportionally higher than on longer ones. To cope with this problem, it is necessary to consider nearly the same length (or the equivalent running time) on each branch line. If this is not possible, the proportional running time supplement $S_d[\%]$ can be used on such courses, which are modified starting from the block section where the occupation steps of the considered train are close to those of other trains. Suppose, for instance, that we have a train entering a line at the second-last station before the main one, but which cannot be simulated on the branch line. The proportional supplement will be used, inserting it backwards from the last station, where the timetable cannot be modified.

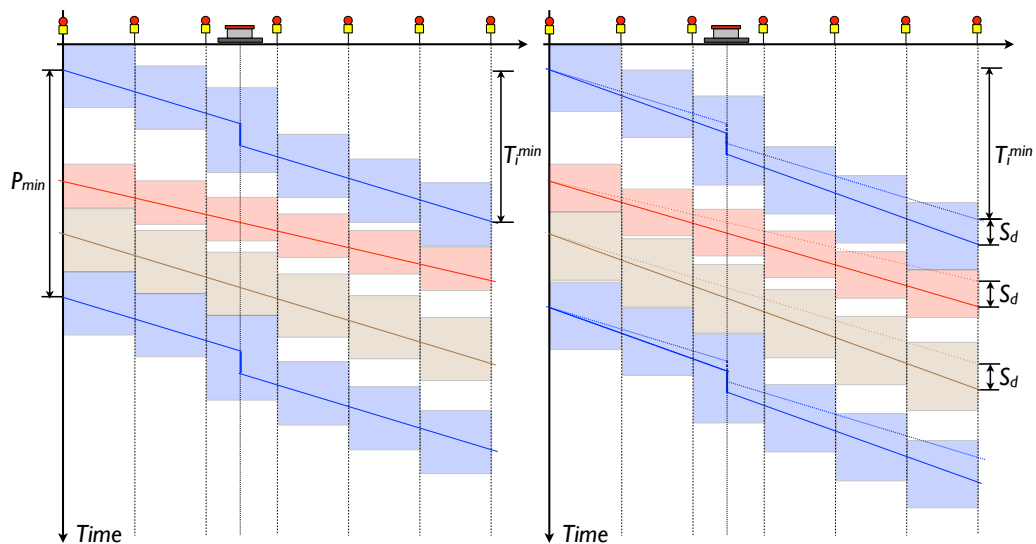


Figure 5.3: Dense timetable (on the left) modified by inserting distributed running time supplements.

5.1.4 Concentrated running time supplements

Concentrated running time supplements are usually inserted at node-entrance and at main stations, to take into account possible conflicts and compensate them. This leads to a proportionally high increase in running time between two timetabling points, and consequently to longer (planned) occupation of the critical sections of the node.

It is possible to insert such supplements into a timetable in two different ways: increasing the headways to consider the longer occupations or allowing the overlap of occupations, without modifying the headways.

If the first approach is used, buffer times are implicitly inserted within the timetable since, in some sections, longer occupations are assumed to compensate possible conflicts. As a result, it would be difficult to evaluate the impact of buffer times and supplements separately, since they are inserted in the timetable at the same time.

On the other hand, with the second approach, although longer occupations (and therefore conflicts) are theoretically planned, they do not occur during simulation if no stop on such sections is planned or if the early departure is allowed. In this case, concentrated running time supplements can be seen as mere performance increase measures, since they have no effect on traffic.

In this study, concentrated supplements are inserted using the second approach. A supplement (S_c) is inserted at the last timetabling point, without modifying the rest of the timetable nor the period. Therefore, when no stochastic phenomena are inserted, trains arrive earlier than planned by an interval corresponding to S_c . Moreover, capacity remains unchanged inserting running time supplements.

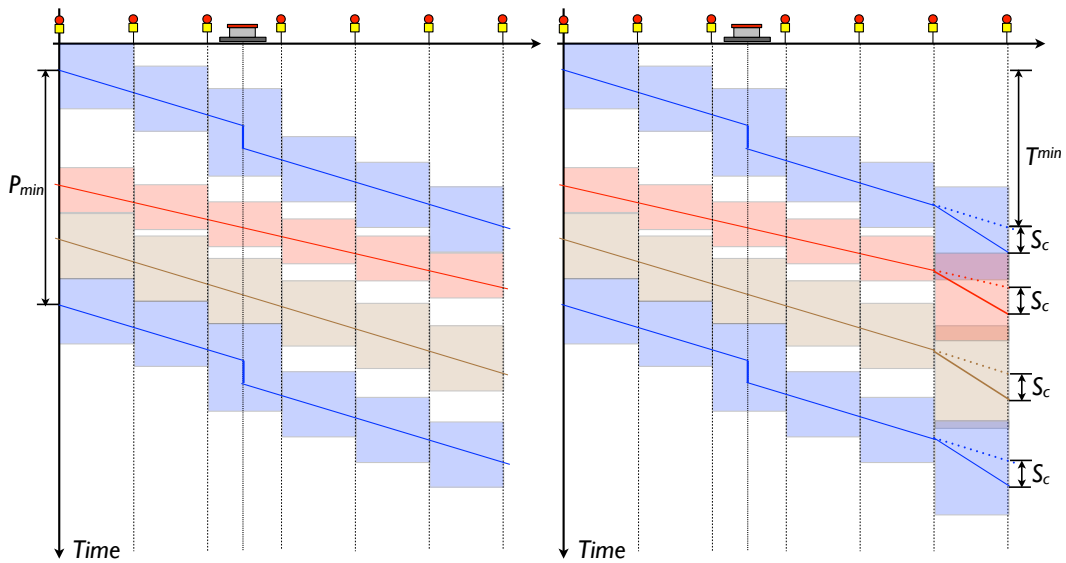


Figure 5.4: Dense timetable (on the left) modified by inserting concentrated running time supplements at the last timetabling point.

5.1.5 Stop time supplements

Stop time supplements are inserted at intermediate stops to increase departure punctuality by compensating both dwell time variability and arrival delay. The insertion of stop time supplements leads to an increase of the total running time, which is normally different for each service. For example, in Figure 5.5, only the stopping trains have an intermediate stop where the supplement S_c is added.

This difference in the increase of total running time on a line leads to conflicts in the dense timetable, which are avoided increasing the headway times. Since supplements are normally not completely used to compensate primary delays, they partially act as buffer times, reducing the risk of conflicts. Therefore, while concentrated and distributed running time supplements are inserted without modifying the capacity, stop time supplements implicitly increase the headway time between trains, reducing the effective capacity.

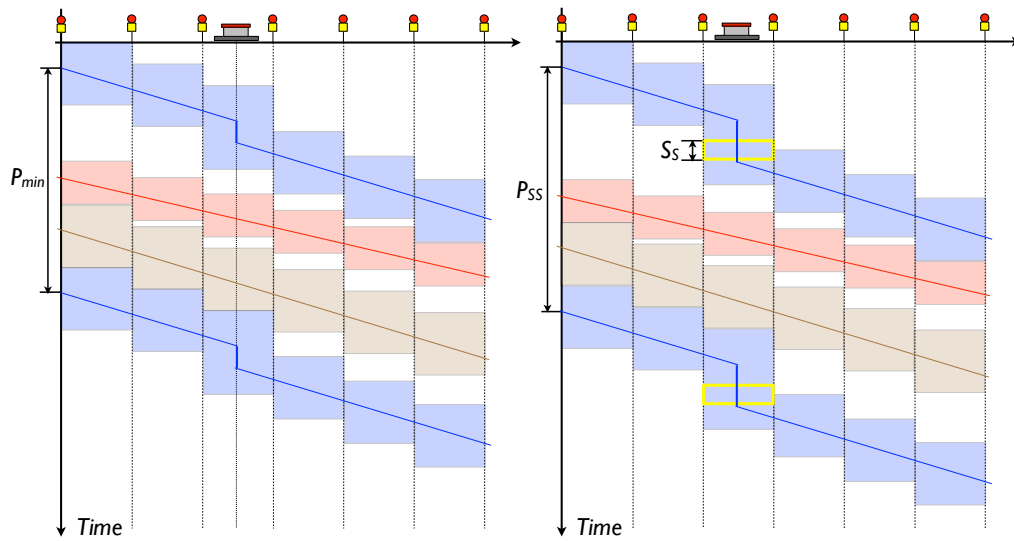


Figure 5.5: Dense timetable (on the left) modified by inserting stop time supplements at the intermediate halt.

As for the other supplements and for the buffers, stop time supplements can be easily inserted in a dense timetable. New departure times and the new period have to be calculated. Consider a line between stations A and D, with two intermediate stops B and C (Fig. 5.6). The stops are not shown, but the time loss is included in the running times. All trains stop at least at one intermediate station, except for the red one, which therefore has no supplement ($S_2 = 0$). Two assumptions can be identified:

- 1) departure time has to be delayed for each service faster than the preceding one;
- 2) for each train that has to be delayed, the delay is the difference between the total supplements of the preceding train and its own total supplements.

Defining:

D_i = Departure time for train i in the timetable with stop time supplements

S_i = total stop time supplement for train i (sum of supplements at each station)

the departure time for a train, considering only its supplements and the supplements of the preceding service:

$$D_i = D_i^d + S_{i-1} - S_i \quad \text{if} \quad S_{i-1} > S_i \quad (5.8)$$

$$D_i = D_i^d \quad \text{if} \quad S_{i-1} < S_i \quad (5.9)$$

When more trains are considered, supplements of all the preceding trains have to be added. In the example, for each train:

$$D_2 = D_2^d + S_1 \quad \text{since} \quad S_2 = 0 \quad (5.10)$$

$$D_3 = D_3^d + S_1 \quad \text{since} \quad S_3 > S_2 \quad (5.11)$$

$$D_4 = D_4^d + S_1 \quad \text{since} \quad S_4 > S_3 \quad (5.12)$$

$$D_5 = D_5^d + S_1 + S_4 - S_5 \quad \text{since} \quad S_5 < S_4 \quad (5.13)$$

$$D_6 = D_6^d + S_1 + S_4 - S_5 \quad \text{since} \quad S_6 > S_5 \quad (5.14)$$

Since 6 is the first train of the second period, the increase in the period is:

$$P_{ss} = P_{\min} + S_1 + S_4 - S_5 \quad (5.15)$$

This very simple procedure can be used to insert supplements even if overtakings take place at an intermediate station, as in the example in Figure 5.7. Consider a train, which is overtaken by train 2 in station C, and is followed by train 3. Total stop time supplement for train 1 (S_1) consists of two parts, which include the supplements inserted in b (S_{1ac}) and in c (S_{1cd}). Train 2 has no supplements, while train 3 has its supplement (S_3).

Train 2 is delayed by the supplement of train 1 before the overtaking (S_{1ac}). Train 3 must be delayed at least as 2, to consider headway at departure in A, but also headway time at arrival in D after 1 must be respected. As a result:

$$D_2 = D_2^d + S_{1ac} \quad \text{since} \quad S_2 = 0$$

$$D_3 = D_3^d + S_{1ac} \quad \text{if} \quad S_3 > S_{1cd} \quad \text{and}$$

$$D_3 = D_3^d + S_{1ac} + S_{1cd} - S_3 \quad \text{if} \quad S_3 < S_{1cd}$$

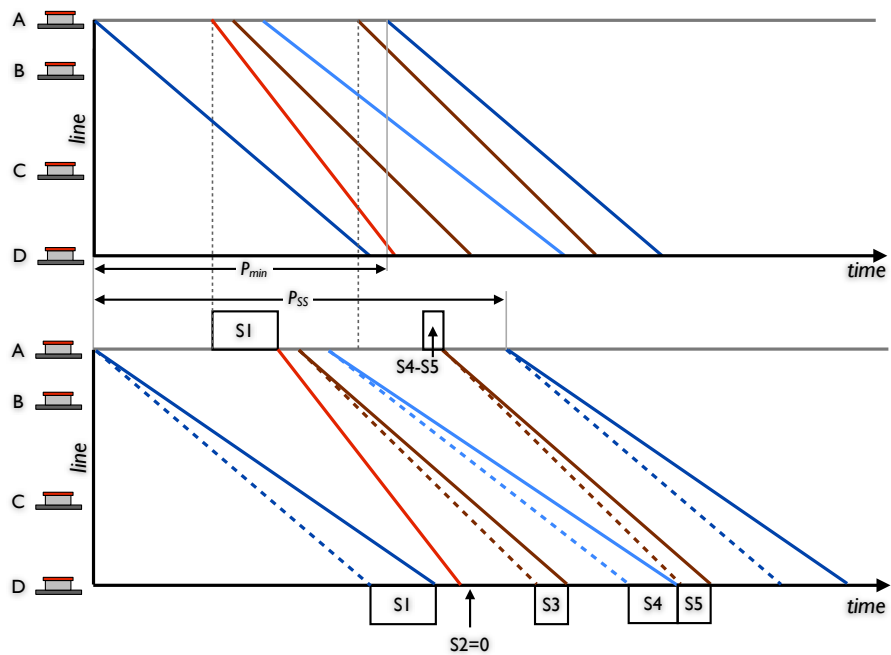


Figure 5.6: Dense timetable (up) modified by inserting stop time supplements, and consequently shifting the courses.

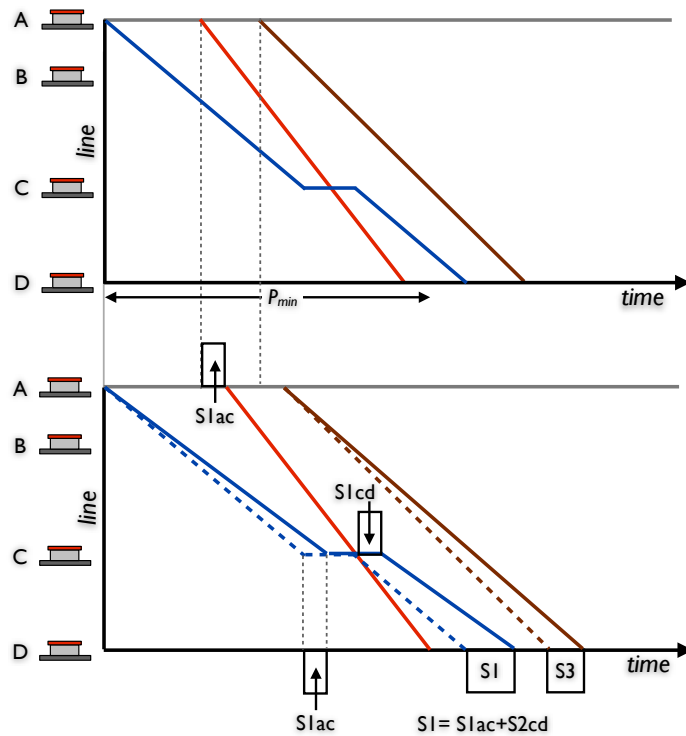


Figure 5.7: Dense timetable (up) in presence of an overtaking modified by inserting stop time supplements, and consequently shifting the courses.

5.1.6 Stochastic phenomena

As for the calibration and the validation of existing timetables, to evaluate the impact of different buffer and running times supplements on reliability it is important to define the stochastic phenomena that characterise train movements. As in (4.1.2) three kinds of stochastic phenomena were considered: the initial delay, the running time variability and the stop time variability.

Initial delays

Initial delays are not inserted importing or modifying the specific recorded distributions for the simulated lines, since these data are the result of an given timetable. To obtain a more general validity of the simulations, typical distributions found in literature have been used instead. In particular, log-normal distributions have been chosen, since they are the best-fitting distributions on a number of tests carried out on 2007 and 2008 data for Turin, Palermo and Roma nodes. An example of fit is presented in Figure 5.8.

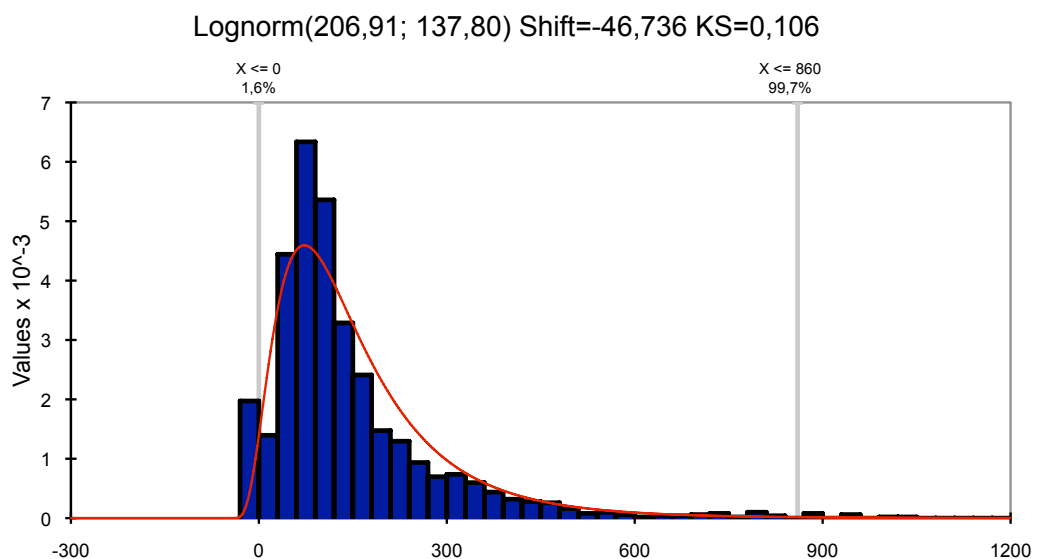


Figure 5.8: Log-normal fit and corresponding parameters for the departure of Regional trains in Turin Porta Nuova.

A number of hypothesis were assumed, in order to reproduce distributions realistically under very changeable conditions:

Different delay levels: Once the distribution to be used for a first scenario (with the lowest delay) has been defined, it is necessary to set up a rule to increase delays and generate other scenarios. In other words, once the first distribution is fixed by its mean and variance, the rule defines how the variance reasonably grows with increasing mean delays. To obtain a realistic value, a mean parameter was calculated on the above mentioned data to relate variance in different scenarios.

The simple expression is the linear regression of the standard deviation as a function of the mean delay for at least 150 trains on 100 days for each category, using a linear trend-line. All trains were measured at the beginning of the simulation areas or at departure, if they start within the area. Very different services were considered, in order to obtain sparse values from 0 to more than 600 seconds mean delay. Only trains with less-than-30-minute delay were considered, to avoid the effects of major breakdowns on mean delay. However, the presented relationships (Table 5.1) are not proposed as general rules to link standard deviation and mean delay, but can be considered as very representative in the Italian contest. The data plots that were used for the definition of the empiric relationships are depicted in Figures 5.9-13.

Regional trains	$SD=0.6 M + 140$
Intercity, Departing	$SD=0.9 M + 110$
Intercity, Passing	$SD=0.4 M + 220$
High Speed, Departing	$SD=0.1 M + 320$
High Speed, Passing	$SD=0.7 M + 120$

Table 5.1: Relationship between Mean Delay (M) and Standard Deviation (SD) for different train categories.

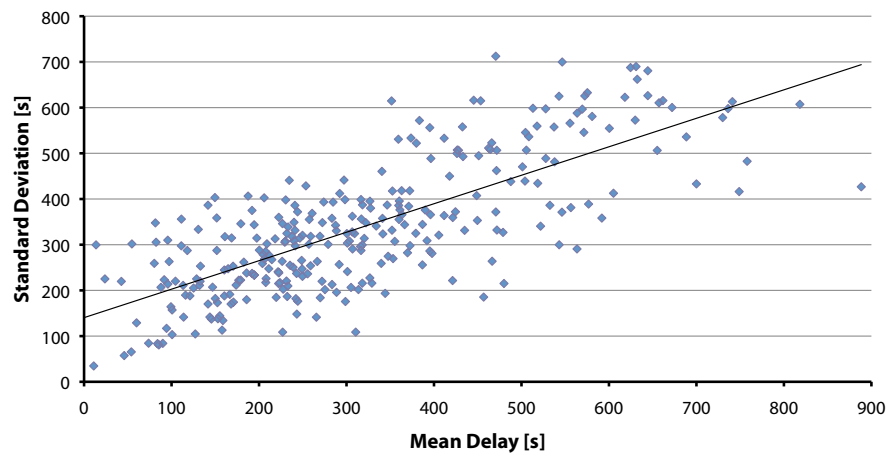


Figure 5.9: Standard Deviation vs Mean Delay for Regional trains.

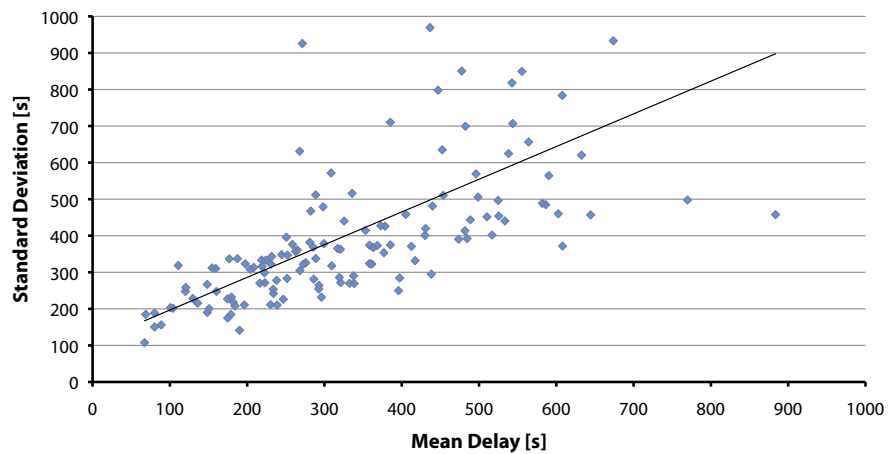


Figure 5.10: Standard Deviation vs Mean Delay for departing Intercity trains

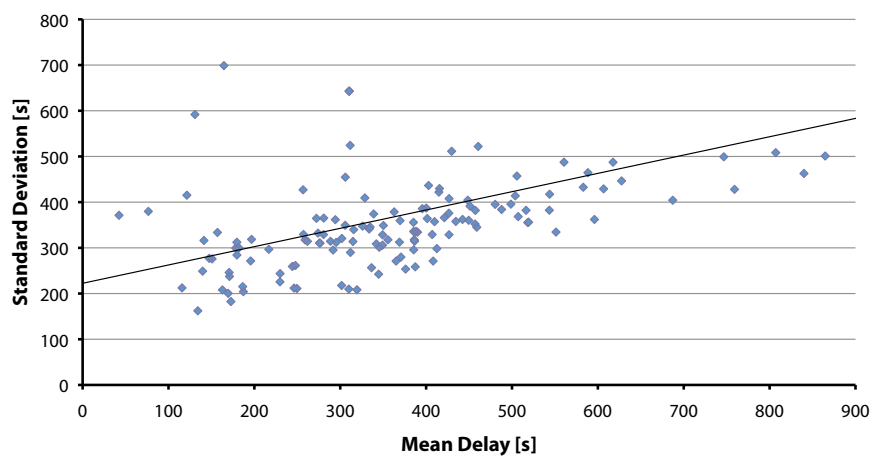


Figure 5.11: Standard Deviation vs Mean Delay for passing Intercity trains

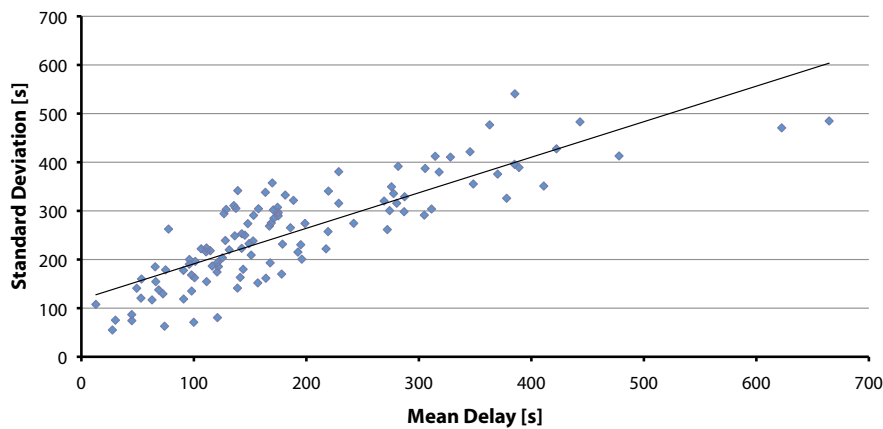


Figure 5.12: Standard Deviation vs Mean Delay for departing High-Speed trains

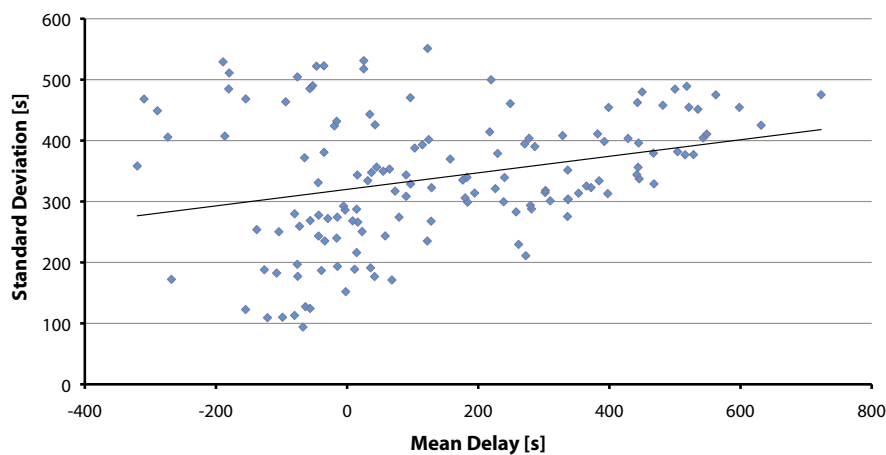


Figure 5.13: Standard Deviation vs Mean Delay for passing High-Speed trains

- *Early-running trains*: Intercity and High-speed services often do not stop at the first station of a simulation area; as a result, trains entering early in the system have also to be considered. Theoretically, since no running time supplements are used to build a dense timetable, it could be possible to exclude early-running trains (which are produced by supplements outside the simulation area). However, early-running trains were considered, in order to focus on a node without modifying the real conditions at the border. The advance was obtained by splitting the corresponding category into two parts (one for trains stopping at the entrance into the model and one for passing trains) and then calculating new distributions, separated for the trains departing and passing at the first station. Distributions for passing trains were moved to start at -200 seconds (for high-speed trains, Fig. 5.13) and the mean delay was consequently adapted.

- *Freight trains*: in Italy, freight trains always have the lowest priority and their schedules contain a lot of slack time that can be used as synchronisation time among different lines or for unplanned overtakings. Therefore, it is impossible to define a real schedule for such trains, whose distribution is often very wide, with many trains running up to three hours in advance and others with heavy delays. However, freight trains are not negligible on the most important mixed-traffic lines; to cope with their variability, they were inserted with the same variability of Intercity trains, but they were not inserted in the statistics, as in most reliability statistics in Italy.

Stop time variability

Dwell time is normally not function of a particular timetable structure, but of the number of passengers boarding and alighting and the train-type. As a result, it is more difficult to assume a change in dwell time variability (and therefore in stop time) when simulating a realistic scenario. However, in a more theoretical approach, stop time variability can be modified, for example to evaluate the impact of different strategies for the distribution of supplements. In this study, dwell time was modelled as a fixed part (*min. stop time*) plus a normal distribution, whose mean value is independent of the planned stop time.

Running time variability

More complex to estimate ex-ante if compared to stop time variability is the variability of running times, which is determined by train performance variability (4.3.4). Since it depends most on rolling stock conditions and on the driving styles, it is impossible to assume under what conditions running times can change stochastically. In other words, it is possible to evaluate the impact of different rolling stock, or of a new ATC system, but not of the way the driving style changes consequently. Even more unrealistic could be, for example, to decide a priori the benefits of driving support systems or other technologies that should reduce running time variability.

Train performance distributions, were therefore considered as fixed and correspond to the values obtained during model calibration on the Trieste - Venice line.

5.2 Model application

The approach is suitable for a wide range of estimations concerning the impact on reliability of variations in one or more factors. The approach maximises the benefits of micro-simulation, simplifying the definition of new scenarios and the analysis of results. Once the dense timetable has been defined, nearly all factors can be modified to estimate their impact on reliability.

It is possible to quantify the relationship among capacity (buffer times), running time supplements, initial and station delays and reliability, and then compare them for different timetable structures or in combination with infrastructure improvements, such as ETCS, shorter block sections, different speed profiles and others.

The basic element of any evaluation is a capacity-reliability curve, which is obtained for a given infrastructure and given stochastic phenomena, starting with a dense timetable and then inserting buffer times: capacity is therefore implicitly reduced. Once the first curve has been obtained, running time supplements can be added to each scenario, obtaining a second curve. In the same way, initial and stop time delays can be modified to define other curves.

Since the insertion of buffer times and running time supplements can be automated (5.1), the editing of timetables to obtain the relationship between running capacity and reliability can be performed quickly. Moreover, since the initial and station delay distributions are standardised, also the definition of different delay scenarios can be performed semi-automatically, for example by doubling or adding 60 seconds to the mean delays per train category. Corresponding distributions are automatically generated and inserted in the simulation tool.

As for all micro-simulations, reliability can be measured with any indicator to point out different characteristics:

Punctuality (P), for example with 5' threshold, can be used to obtain results comparable to the quality measures currently performed by most companies. As clearly explained in (2.5), punctuality shows a number of weaknesses, which become particularly relevant when representing more abstract phenomena. However, it can be used to evaluate some timetable-related measures, such as the running time supplements required to obtain a certain punctuality rate.

Mean Delay (M) offers a better representation of the phenomena, also considering heavy delays.

The *F index (F)* (3.1) features a number of strengths compared to punctuality and mean delay but, being new, it might not be very easy to understand, if not clearly explained.

From these three indicators, others can be easily calculated to point out different aspects, for example:

Δ *Mean Delay (ΔM)* is the difference, calculated for each train, between output and input delays: it represents delay propagation within the model. If the mean output delay is lower than the mean input delay, delay propagation is lower than running times supplements, normally indicating a stable system.

(ΔF) is the difference, calculated for each train, between output and input *F* indexes. It represents a more precise alternative to ΔM , clearly showing not only if the capacity to absorb delays, but also weighting early running trains, which may cause conflicts in nodes.

5.3 Case Studies

The model was applied to different case studies in Italy, first to test and validate it with a number of traffic engineers in Italy and then to use it as a decision support tool in both investments and timetable planning.

First, the method was applied to the node of Turin, to test it on an infrastructure whose model had already been validated. Moreover, by testing the method on a node instead of doing it on a line it was possible to point out some potential weaknesses in the coordinated management of more lines and of a large dead-end station.

The second case study was the Rome - Formia line, considered saturated by RFI and affected by frequent and high delays. Third, at a more applicative level, the approach has been used to evaluate the benefits on capacity and reliability of two possible improvements on the block system.

5.3.1 Turin Node

The Node of Turin (3.2) contains a number of characteristics, which make it an interesting case study. It contains, among others, a mixed-traffic double track line (Turin - Chivasso), equipped with continuous ATC, a large dead-end station (Turin Porta Nuova), where the traffic from different lines has to be coordinated, and a four track line (Turin Lingotto - Trofarello).

The dense timetable was constructed on the basis of the 2008 timetable keeping the same services and the most frequent train order, since the timetable was not always periodic.

Different buffer times were added to the timetable after each train, obtaining 5 scenarios, with 0-, 30-, 60-, 90- and 120-second buffer time. To obtain the first reliability-capacity relationship, the timetable was simulated using the 2008 mean delays per category, calculating the correspondent standard deviation and consequently the distribution. The mean delay among all trains was rounded up to 120 seconds, to make it easier to classify and modify it.

Concentrated and distributed running time supplements were then inserted in various combinations from 30 to 180 seconds, to evaluate their impact on punctuality and mean delay.

Each simulation set was then combined with different stochastic variability, represented by increasing initial delays and stop time variability. Mean initial delays of 30, 60, 90, 120, 180 and 240 seconds were simulated, to estimate the impact of growing delays on reliability under different capacity usage and delay compensation measures.

5.3.2 Rome – Formia

The 214-kilometre-long Rome - Formia - Naples line is one of the fastest and most modern lines in Italy, excluding the new high-speed lines. The line entered into service in 1927 as fast connection between the two cities, with operating speed of up to 180 km/h, that could be reached on most part of the line also thank to a number of tunnels. In recent years the line has been equipped with the most advanced continuous ATP system in Italy, the so-called "4-aspect BACC" (4.2 and Fig. 4.1), to allow a maximum speed of 200 km/h for tilting trains on some sections. Block sections are 1550 metres long on the entire line, except for two tunnels, where sections are longer than 2000 metres. Until 2007, the line was used by fast tilting

Eurostar services, as well as Intercity, Intercity Night, Inter-regional and stopping services. When the new high-speed line Rome - Naples entered into service, the Eurostar services gradually started using it, reducing the difference in commercial speed among different train categories on the Rome - Formia - Naples line.

In this study, the 128 kilometre-long section between Rome and Formia was considered, in order to include in the simulation the central part of the line, only used by Rome - Naples Intercity and Inter-Regional services, and the northern part, saturated by Regional trains to/from Rome. In Figure 5.14 is represented the simple layout of the Rome - Formia line section

The line is affected by relevant delays, which significantly reduce the punctuality of services; on the other hand, the new high-speed line has reduced both the number of trains and the speed difference among them, clearing capacity for the regional services. In order to obtain an equilibrium between the goals of maximising the infrastructure usage and the service reliability, the trade-off between capacity and reliability has been studied and represented.

The same scenarios described above (5.3.1) with reference to the node of Turin were prepared and simulated, obtaining a number of results that were successfully used to assess the maximum number of slots that could be planned.

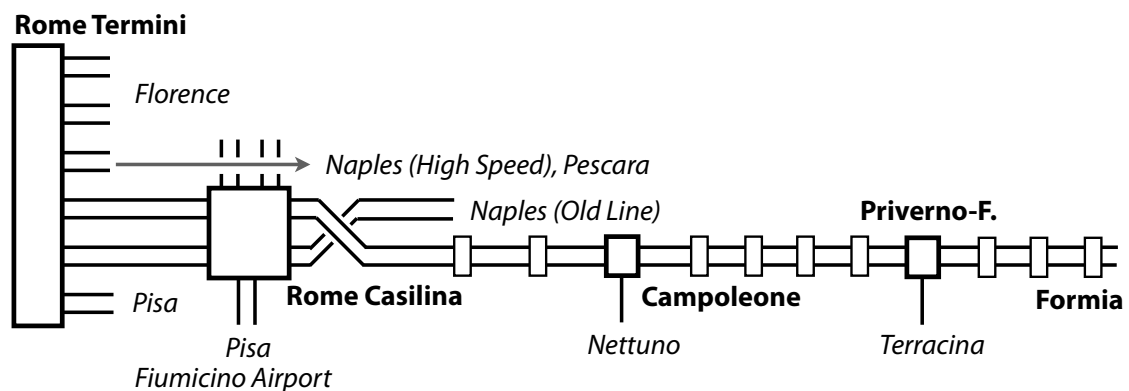


Figure 5.14: Simple layout of the Rome - Formia line

5.3.3 Venice - Trieste

On the Venice - Trieste line (4.5.1), capacity in the less-used central part is limited by two long block manual block sections (Fig. 5.15). RFI developed two possible alternatives to increase capacity on the line without modifying the block

system on the other sections. In Italy the automatic block based on track circuits can be used only on primary lines, while on all others more affordable axle counters are used. However, regulations limit the number of intermediate sections between two stations to 3. In order to obtain a capacity increase maintaining the axle-counter system, the benefits of adding one block section on the critical part of the line, eventually combined with a continuous ATP system were simulated. Realistic and standardised delay distributions were used, and no running time supplement was inserted, in order to evaluate reliability more precisely.

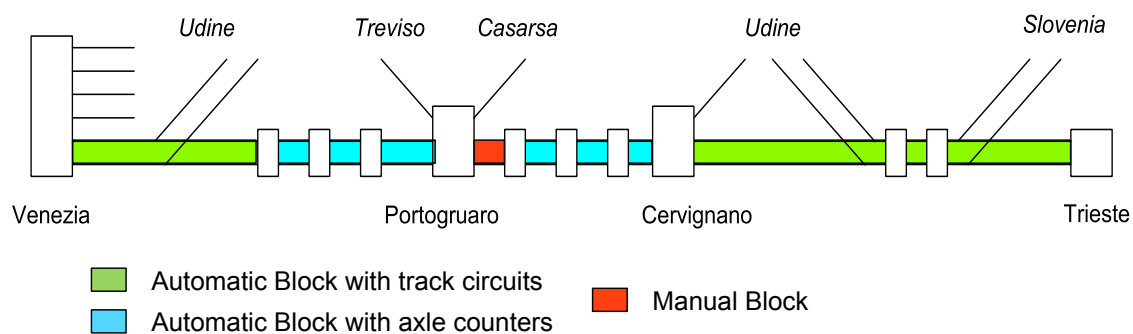


Figure 5.15: Block systems on the Venice - Trieste line

5.4 Main Results

The very wide set of results is the combination of the variation in buffer times, running time supplements and initial delays with different reliability measures. Results can be aggregated at different levels, for example to include only a train family on a line. Since the study is based on a standardised timetable where the single service does not have its specific delay distribution and behaviour, an aggregation of all services on one line and direction gives a more synthetic and representative indication, to be used to compare different scenarios. At an even higher level, to evaluate the performances of major stations with many interconnected lines, an aggregation of all trains arriving or departing could be chosen.

The first result that is obtained is the representation of the relationship between buffer times and expected reliability.

5.4.1 Buffer time – Timetable Robustness

The relationship between buffer time and reliability represents one of the most interesting outputs of the approach. It gives a synthetic representation of the timetable stability in presence of an increasing level of saturation of the infrastructure. This way, a part of the capacity balance (Figure 2.5) defined by the UIC (2004) for a real line or network is depicted.

When the mean delay is used as a reliability measure, and real delay distributions are used as input, simulation results can be compared to real traffic and timetable analysis. Consequently, an implicit validation of the results is obtained together with an indication of the capacity margins on the line. Depending on the expected increase in delays, it is possible to decide if an increase in traffic density is sustainable. In Figure 5.16 the variation of mean arrival delay (with 2' mean initial delay) is depicted, corresponding to increasing headway times for all Formia - Roma services.

While using the maximum line capacity a 300-second mean delay is measured in Roma Termini, corresponding to a delay propagation of about 180 seconds per train. Already inserting 30- or 60-second buffer time, and consequently increasing the mean headway time from 6.3 to 7.3 and 7.8 minutes respectively, mean arrival delays of 240 and 190 seconds are obtained. Increasing buffer times by further 30 seconds, delays fall to 160 seconds. The improvement in reliability becomes then gradually less significant, while the increase in headway time is constant.

Analysing the curve, it is possible to notice that the increase in reliability by adding buffer times decreases constantly: a threshold could be defined, after which the reduction of capacity is more relevant than the punctuality improvement. The global shape of the curve is similar to the same qualitative evaluations that can be found in literature, and to the findings of different studies, based on stochastic delay propagation or on queueing theory. This confirms the validity of the method on one side and, on the other, underlines the advantages of micro-simulations, which combines the possibility to obtain the same results of the most advanced approaches with the use of proven tools, already widely used in most countries.

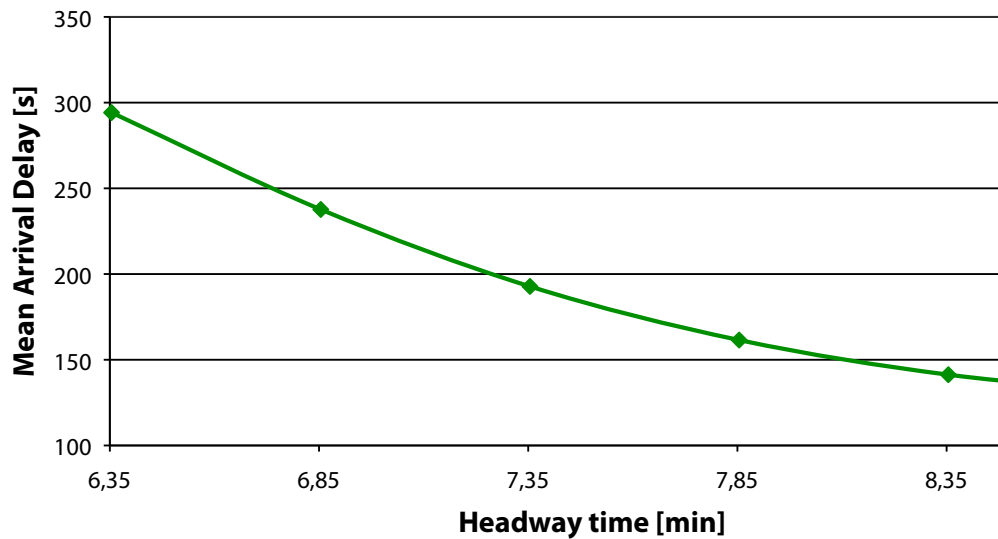


Figure 5.16: Mean arrival delay as a function of headway time on the Rome - Formia line.

Using different indicators, other aspects can be focused. In figure 5.17 the trend of capacity and reliability on the Turin - Chivasso line is depicted. The percentage reduction of available capacity increasing the headway time is represented by the red line, while the green line shows the corresponding percentage reduction in delay propagation by $\Delta Mean Delay$ (5.4). The different shape of the curves gives a clear representation of the impact of buffer times on reliability: while in a first phase they lead to a strong reduction of delay propagation (up to 80% with the first 60 seconds), benefits become even lower than the decrease of the available capacity. Compared to the headway time vs mean delay diagram (Figure 5.16), this representation offers a simpler identification of the trade-off between capacity and reliability, being simple to understand for non-experts.

Figure 5.18 contains the same trade-off, but concerning all Formia - Rome trains. Hence, the diagram represents the same scenario already shown in picture 5.16. Compared to the results of the Chivasso - Turin line described above, this line presents a significantly different behaviour. This can be first referred to the difference in percentage increase in headway time. In fact, while in Turin the dense timetable leads to headway times of less than 5 minutes, in Rome it grows to about 6.3 minutes: since 30-second buffer times are added each step, the percentage increase is clearly higher in Turin, leading to corresponding more positive effects. On the other hand, the differences between the lines, the timetable structure and the rolling stock cannot be ignored. In fact, the lower speeds, lighter (and therefore more performing) trains and a timetable structure, in which most trains run for shorter distance on the main line, increases the possibility to recover delays

Some considerations can also be drawn considering the real timetables. In Turin a headway time of 6.5 minutes is used in the peak-hour, and delay propagation is normally very low (4.5), confirming the capacity-reliability relationship, in which with 6.5 minutes headway time delay propagation is reduced by 90%. Moreover, considering the diagram, the capacity used in the timetable is nearly the maximum one to avoid high delay propagation.

In Rome, 17 trains are scheduled in the two peak hours, with a mean headway time of about 7 minutes. To increase usable capacity, the difference in commercial speed among different services is reduced, adding significant running time supplements to the fastest trains. Hence, the timetable structures cannot be compared precisely. However, with 7-minute headways the line is not only considered saturated by RFI, but it is affected by significant delay propagation: this can be found also on the diagram, where with 7-minute headways delay propagation is very high.

A more accurate comparison with real results can be obtained considering also running time supplements: in the following paragraph, different running time supplements are inserted in the diagrams.

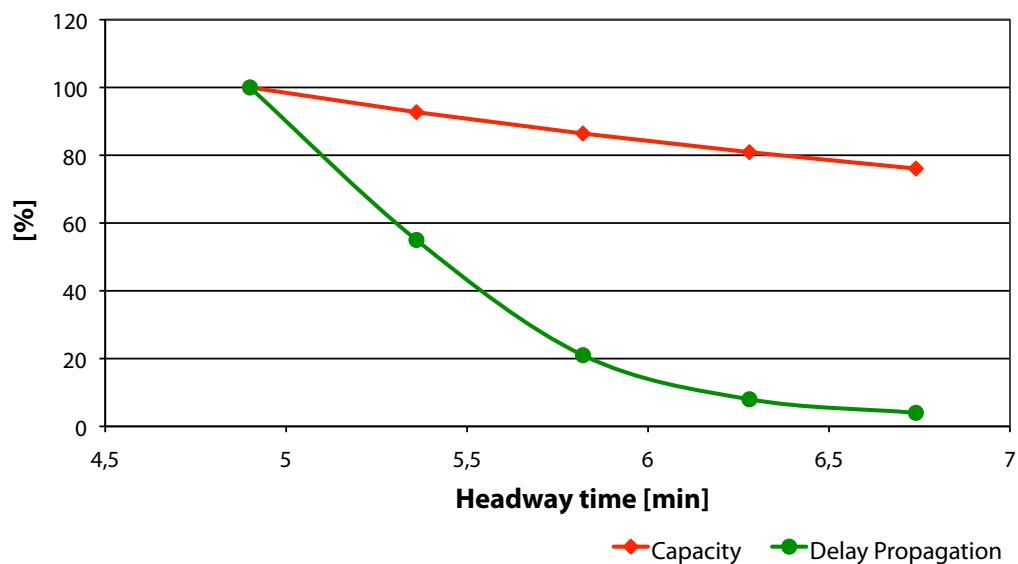


Figure 5.17: Percentage variations in available capacity and delay propagation corresponding to increasing buffer times on the Turin - Chivasso line.

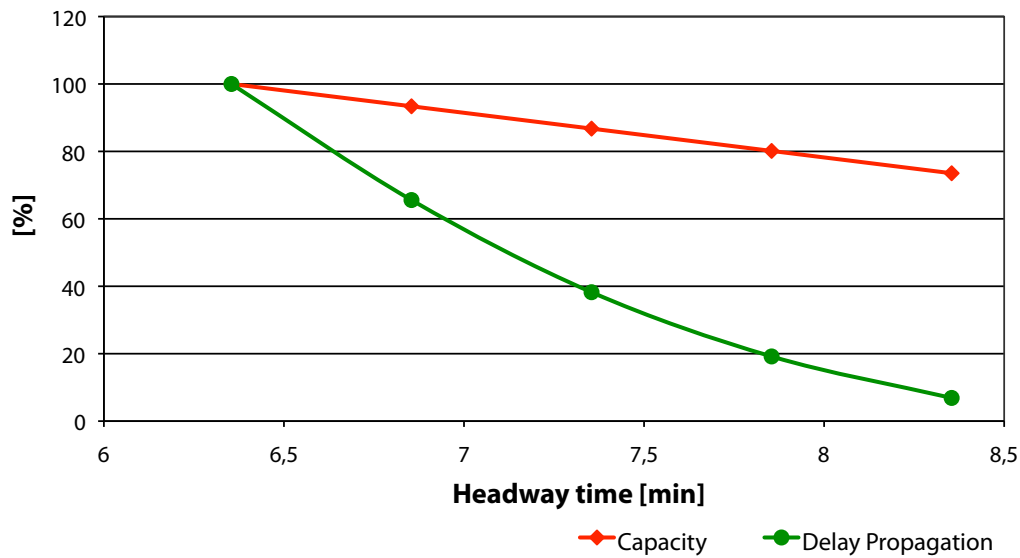


Figure 5.18: Percentage variations in available capacity and delay propagation corresponding to increasing buffer times on the Rome - Formia line.

5.4.2 Running time supplements – Reliability

Concentrated running time supplements are inserted in the last timetabling points to increase punctuality. Their effect on reliability depends not only on their location, but also on their combination with buffer times. In fact, the more trains are delayed, the more running time supplements are effective; hence, if propagation is limited by buffer times, also delays to be recovered will be lower.

High supplements lead to frequent early-arriving trains, which are not useful for the global stability of operations. Therefore, to evaluate the impact of supplements a measure has to be used, which is not positively influenced by early runnings. Since the supplements are normally inserted to increase punctuality at important stations, the same index was used in this study for all evaluations.

In Figure 5.19, the relationship between buffer times and 3-minute punctuality on the Turin - Chivasso line is represented: different supplements correspond to the data series, whose values in seconds are indicated in the legend. Analysing the difference in punctuality in different series, the effect in terms of punctuality appears nearly proportional to the supplements for each buffer time scenario. However, benefits decrease significantly when buffer time is increased but, unlike buffer times, the decrease appears more gradual. The effects of the combination of supplements and buffers can be clearly observed determining their amount to obtain a given

punctuality level, such as for example 95%. It cannot be reached with the dense timetable - obviously not considering very high supplements - but a number of different combinations is possible: for example, inserting 2' supplements and 30-second buffer times, or even 60-second supplements and 90-second buffer times.

The same diagram for the Rome - Formia line is represented in Figure 5.20. Despite the very different conditions, benefits of supplements are similar to the Turin case. As already noticed for the capacity-reliability diagram, punctuality on this line is significantly lower. As a result, the decrease in the benefits of supplements is more gradual, with a difference by more than 20% between 0- and 150-second supplements even in combination with the highest buffer time. Moreover, punctuality with the highest buffer time and supplements is 93%, compared to the 99% of Turin: higher buffer times could be added to complete the diagram.

A comparison with the real timetables appears more complete in this scenario, since punctuality is the result of a combination of running times and buffers. In Turin about 2-minute supplements are used together with about 1.5-minute buffer: 3' punctuality is about 5% lower than the results of simulation. The difference is due to the fact that in Turin some Regional trains have significantly higher delays compared to other services. The impact of these hourly trains on global punctuality is relevant, also because they are among the fastest services within the node; however, aiming at representing the mentioned relationships, the error can be considered acceptable. Moreover, if the services to/from Milan are excluded from the statistics, real punctuality grows to about 94.5% that is only 3% lower than simulation.

In Rome, the real buffer time is about 60 seconds, and 3' running time supplements are inserted at arrival. Real 80% in the peak hour punctuality is not significantly different from the results of simulation and the gap between them is due to the impact on real services of two Intercity-Night trains coming from southern Italy normally with relevant delays, forcing timetable managers to rescheduling measures in order to minimise their impact on operations. Such trains were not inserted in the dense timetable, since it is not appropriate to evaluate the potentials of a line on the basis of one or two spot services normally delayed by more than 15 minutes.

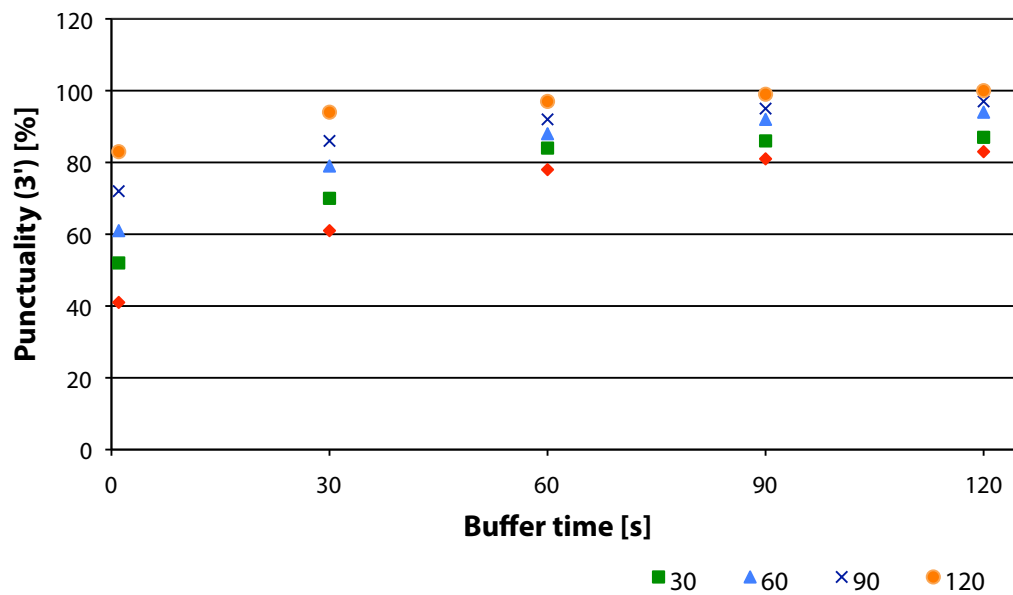


Figure 5.19: Arrival punctuality as a function of buffer time (x-axis) and running time supplements (data series) in Turin node.

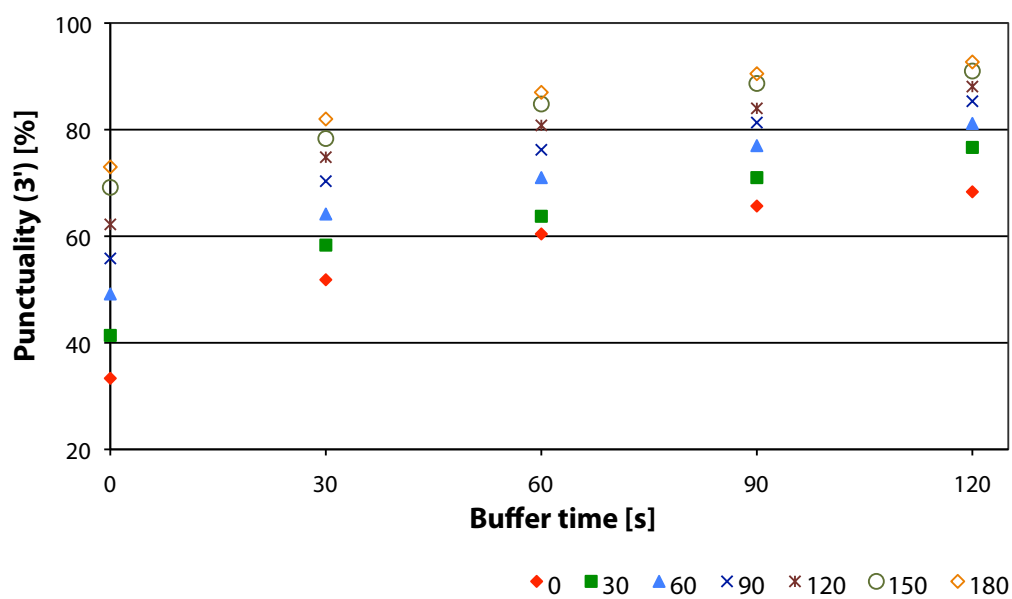


Figure 5.20: Arrival punctuality as a function of buffer time (x-axis) and running time supplements (data series) on the Rome - Formia line.

5.4.3 Stochastic delays – Reliability

The comparison between real and simulated reliability in the previous paragraph highlights the importance of the delay distributions to obtain realistic results. For this purpose, it is possible to represent in diagrams also the variation in reliability corresponding to different delay scenarios.

In Figure 5.21, the delay propagation as a function of buffer times (on the x-axis) and of mean initial delay on the Turin - Chivasso line is represented. A number of considerations can be drawn by accurately analysing the diagram. First, the benefits of higher headway times are proportionally decreasing with all delay values. Second, the differences among curves are higher with low buffer times, which means that high buffer times are useful only in presence of significant delays. For example, if the mean initial delay is 30 seconds, buffer times longer than 60 seconds lead to no increase in reliability; if delays are 180-seconds a significant reduction in delay propagation can be obtained even with higher buffer times. Moreover, it is possible to notice that with 180-second delays, propagation is over-proportionally higher than with lower delays: the difference with the lower curve (120 seconds) remains nearly constant even with high buffer times.

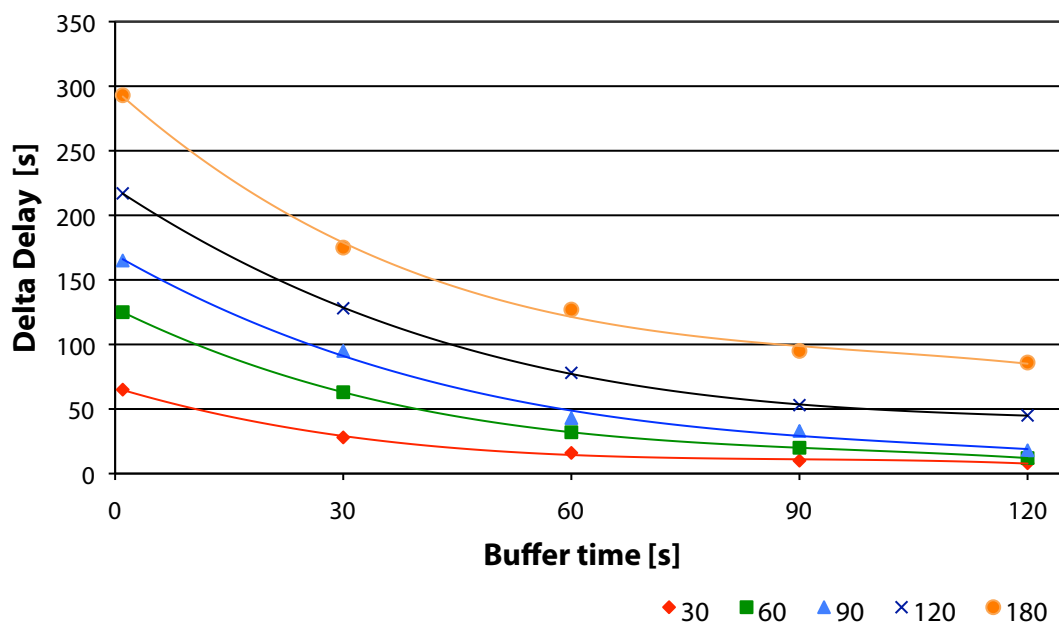


Figure 5.21: Mean delay propagation per train as a function of buffer time (x-axis) and mean initial delay (data series) in Turin node.

To explain this difference it is necessary to focus on the consequences of delays on train traffic. When most delays are very low, they can be compensated by the continuous ATP system. In fact, finding a yellow signal, the train is allowed to enter the following block section and starts braking; as soon as the preceding train has released the block section, the following one is allowed to accelerate again. As a result, measured headway at the end can be even lower than the minimum, reducing the secondary delay of the second train. When delays are higher, trains are forced to stop and the relevant time-losses for re-accelerating produces a further increase of delays.

This means that, in presence of continuous ATP and trains with relatively high acceleration (like nearly all passenger trains) small delays (1-2 minutes) are compensated more than proportionally than higher ones. Therefore, when the amount of trains that are forced to stop is high, this produces an over-proportional growth of delays, which is compensated only with very high buffer times. As a result, when initial delays are significant it is very difficult to obtain high reliability, even if the number of trains is reduced.

5.4.4 Evaluation of Infrastructure improvements

The comparison among different scenarios is performed at best on a capacity-reliability diagram, with pre-defined delay distributions and with no running time supplement, which could modify the results of different scenarios. "2009" represents the current configuration of the block system, while in "4500m" scenarios one block section was added, the manual ones were replaced by axle counters and three other sections were relocated, in order to obtain homogeneous 4500metre long sections. In the second scenario (4500m+ATP), this intervention was combined with the adoption of a continuous (instead of discrete) ATP system. In Figure 5.22 the blocking time stairways in "2009" (above) and "4500m" are depicted.

In Figure 5.23, the effects on capacity and reliability of the improvement scenarios are clearly represented. The diagram shows that the mixed-traffic timetable structure limits the benefits of the interventions in terms of capacity to less than 1 train/hour, but it is very significant in reliability: remarkably, with the homogenisation of block sections (4500m) delay propagation is nearly halved. With the adoption of the continuous ATP, it is reduced to about 1/4 of the initial value.

The increase in reliability leads indirectly to higher usable capacity: if the headway time is defined to limit propagation to a given amount, the usable capacity

in the scenarios is notably different. For example, considering a maximum propagation of 40 seconds per train, one more train per hour can be scheduled in the second scenario.

Although the application appears very simple, the results were used to define the improvements on the line, validating the approach for a more complex use.

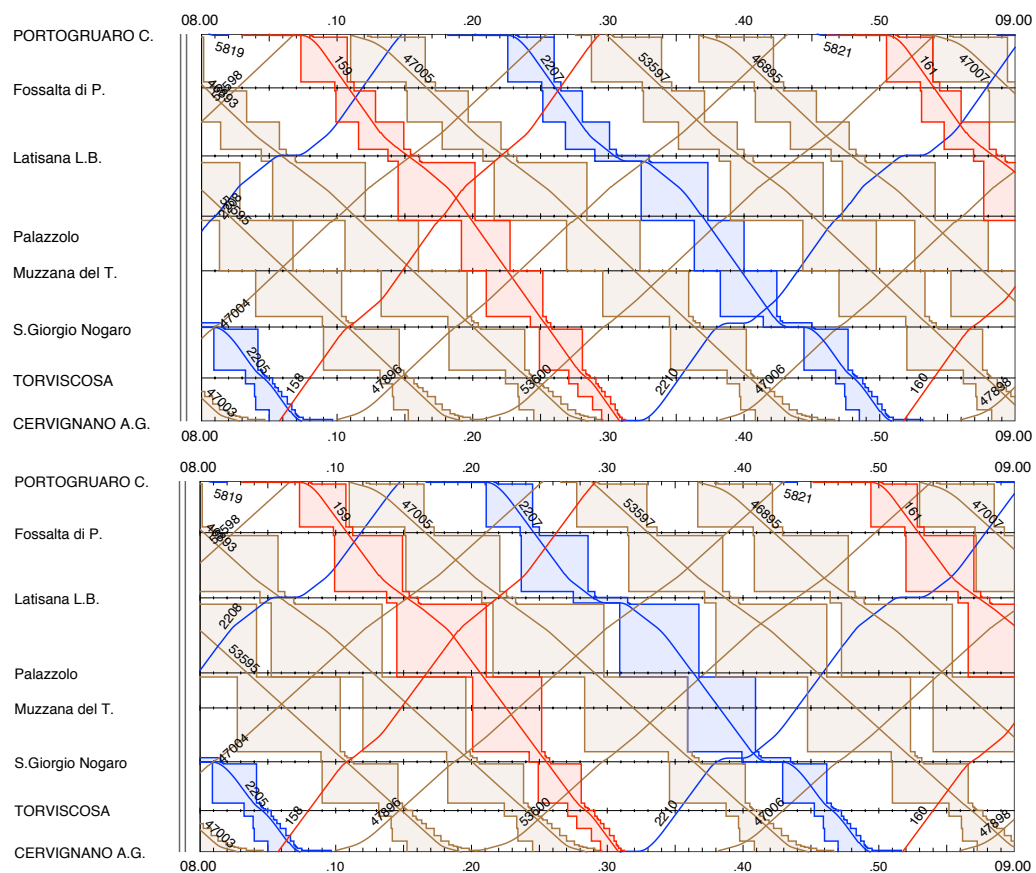


Figure 5.22: Blocking time stairways in scenario 2009 (above) and "4500m".

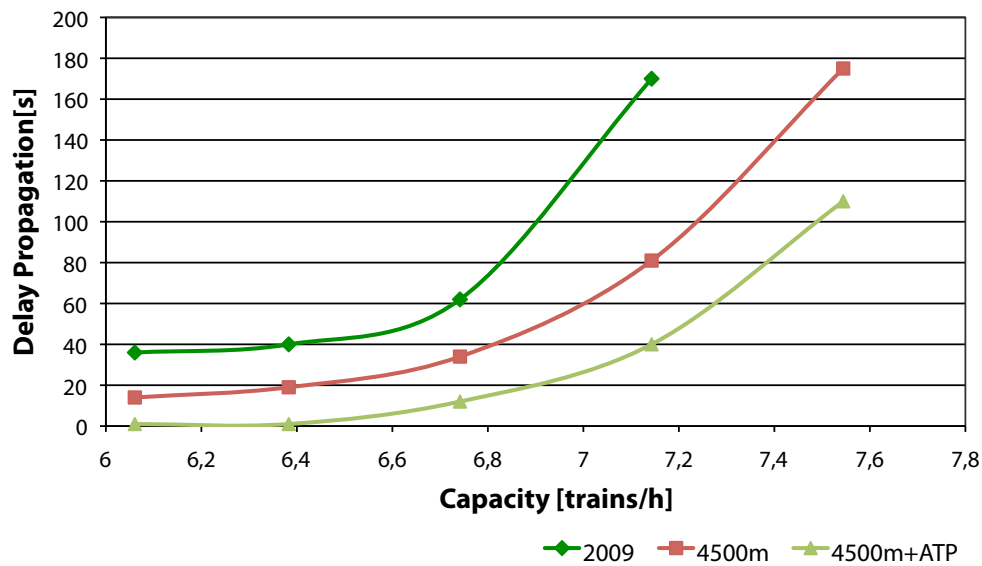


Figure 5.23: Mean delay propagation per train as a function of buffer time (x-axis) and mean initial delay (data series) in Turin node.

5.5 Comparison with stochastic differential equations

Different approaches can be found in literature, that aim at representing the relationship between capacity and reliability on railway networks. Among them, Stok (2.2.3, and *Stok*, 2008) developed a model that uses stochastic differential equations to model train traffic on a line, obtaining a capacity-reliability relationship. The model can be considered analytic for the robust mathematic formulation, but is simulative in the technique used to calculate train motion. Therefore, the same components investigated in this chapter are inserted:

train motion: trains are moved by means of the motion equation, as in all micro-simulators, but an optimal control function that reproduces the driver is introduced. This control function aims at reducing delay or energy consumption depending on the goal assigned to the driver.

stochastic phenomena: instead of using initial delay distributions, the stochastic elements are inserted by adding a Brownian motion to the motion equation, leading to a stochastic differential equation.

Capacity-reliability relationship: Compared to simulation, where the common reliability indicators are used, with Stok's method a conflict probability (called Risk probability) is obtained as output, which gives a measure of the probability to have delay propagation on the line (Figure 5.22).

Although some key elements are comparable, they appear focused on very different issues. Stok takes the advantages of analytical models in terms of computation times, replicability of results and possibility to include any control algorithm. However, it was tested only on a simplified double-track line with no timetable: since it is still timetable-independent, and complex interlocking systems have never been modelled, it cannot be deployed on real large-scale studies.

On the other hand, with the approach proposed in this thesis, a proven and reliable tool can be used to obtain similar results. Although more rigid in the algorithmic part, this method allows simpler model calibration, adaptability to very different interlocking and block systems and a very wide variety of outputs that can be evaluated using conventional reliability measures.

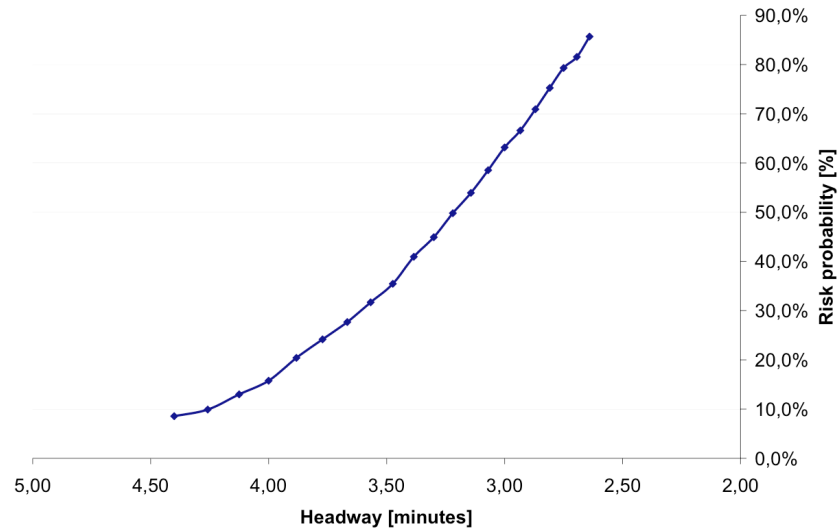
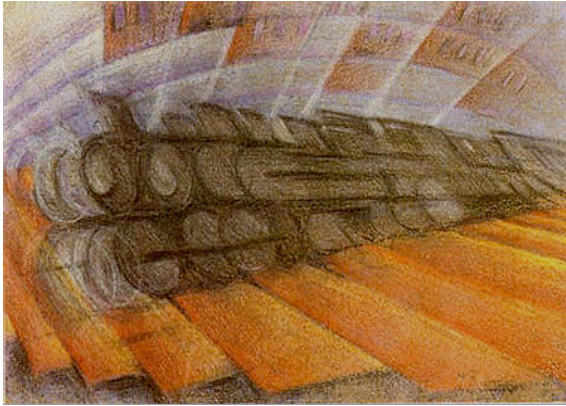


Figure 5.24: Headway time vs conflict probability curve.



(L. Russolo, *Dinamismo di un treno che corre*, 1912)

VI Conclusions

Maximising the usage of the existing railway network and assuring satisfying reliability is an important challenge to increase the competitiveness of railways against the other transport modes. An increasing precision is therefore required not only in the real-time traffic management, but also in timetable planning, where the basis for smooth operations are set up.

In this field, a number of tasks are still performed relying on the experience of planners instead of adopting quantitative methods. In this thesis flexible and reliable micro-simulation is combined with the analysis of real data and an original approach, in order to support planners in a number of critical issues in timetable planning.

In this chapter, the most important achievements of the thesis are summarised; furthermore, some directions for the future research are drawn.

6.1 Main Results

The comparative evaluation of reliability measures led to the development of a new index, that overcomes the main weaknesses of the most used measures. The

new index, called *Delay Frequency Index (F)* extends some key concepts already common in other public transport systems to railways, defining an absolute scale and corresponding quality-of-service levels. Moreover, the proposed definition of levels of service, is simpler to understand and especially focused on the needs of the Infrastructure Managers. The measure was tested not only in some selected stations, but also as network-wide measure, proving suitable to evaluate not only reliability, but also the quality of a timetable considering all timetabling points.

The strengths and the weaknesses of micro-simulation were pointed out applying it on a wide range of case studies, covering very different situations: from single-track lines to nodes and mixed-traffic lines. The applications allowed the fine tuning of dispatching rules and the definition of some of practical advice concerning the construction and management of micro-simulation even on complex networks.

Starting from wide-spectrum real data analysis and including also high-frequency on-board-collected GPS trackings, the behaviour of trains was studied in detail and compared with conventional running time calculations based on the solution of the motion equation. A method for the calibration of the motion equation is presented, which allows an efficient definition of the performance factor starting from station passing data or GPS trackings. The accuracy of the calibrated model was tested on the Venice - Trieste line to evaluate the impact of an increasing precision in the representation of stochastic phenomena on the output quality. The second case study was the Palermo - Punta Raisi single-track line, where the results of simulations supported timetable planning. The third case study was focused on the Turin node; after calibration the model was used as an ex-ante punctuality estimation tool, in presence of significant restrictions due to infrastructure works and of a new timetable. Ex-post analysis of real data confirmed the results of simulation, proving its effectiveness as a reliability estimation tool.

It must be noticed that since the calibration of the model, and especially of the stochastic delays, is strongly related to the data set used as input. Thus, precision in the ex-ante estimation of punctuality can be only assured as a function of the input distributions and not as an absolute prediction, which is not possible. Simulation of timetable and infrastructure changes in Turin clearly demonstrates that strong variations in initial delays lead to unexpected punctuality rates, but a well-calibrated model accurately represents very different scenarios. However, the possibility of obtaining any commonly used or innovative reliability measure as output makes calibrated models very useful to compare timetable or infrastructure scenarios.

To overcome this limitation in the use of micro-simulation, the coordinate method for the evaluation of capacity and reliability was developed. Starting with the construction of a dense timetable based on a realistic offer model, a number of different scenarios are obtained by inserting running time supplements and buffers semi-automatically. Delay distributions are generated and inserted on the basis of a number of assumptions, defined after a wide analysis of real data; also different delay scenarios are defined and inserted in the model semi-automatically.

Perhaps the most significant contribution of the present work is the representation of the relationship between capacity and reliability, which helps planners not only in defining a suitable headway time, but also in understanding the potential benefits of different choices. Inserting delays and running time supplements together with capacity and reliability, the relationships among them are clearly shown, allowing not only a number of general considerations, but also a precise and coordinated definition of buffer times and supplements as a function of the expected delays. Moreover, the model is suitable to evaluate the impact of infrastructure improvements on capacity and reliability, representing a significant precision increase compared to conventional capacity evaluations, but maintaining the proven flexibility of micro-simulation.

In such an approach, the use of proven micro-simulation models is paramount, since the real-world validation of capacity-reliability relationships is normally not a viable solution. This constraint highlights the importance of a well-calibrated model as a starting point to seek reliable results even in complex networks.

6.2 Recommendations and Future Research

The novel method was applied to proven calibrated models and validated by the experience of capacity and timetable planners; however, it has not yet been tested in presence of severe time constraints, which often strongly limit the possibility to perform evaluations in detail. Despite the enormous total computational load to simulate a sufficient number of scenarios, the time necessary to obtain reliable results seems to be competitive with the needs of the Infrastructure Managers, especially on important lines where the search for an optimum between capacity and reliability is an important task. Performing this real test, the method

could be definitely validated and improved, for example by inserting new functions or simplifying some procedures in order to make it more efficient.

A model to identify primary and secondary delays could be used to filter the input delay distributions especially at complex and terminal stations, where at present only very rough estimation of secondary delays has been performed. Concerning the calibration of train motion, a further accuracy increase could be reached using a wider set of GPS trackings, extended to all train-types. These data could also lead to the definition of separated performance factors for the acceleration, full speed and braking phases.

The method for the evaluation of the relationship between capacity and reliability could be applied to very different infrastructure improvement scenarios, covering the most common measures. The results on a standard infrastructure could be used to pre-select the most effective interventions to be used under different conditions, reducing the number of scenarios in each real application.

Results of a number of case studies could be studied from a statistical point of view, fitting the relationships among factors and finding out possible generalisations. These could lead to an analytic formulation of the relationship between reliability, buffer times, running time supplements and delays, which could represent a valuable compromise between time-consuming and precise simulations and the “rules of thumbs” commonly used in practice.

Finally, the results in terms of capacity and reliability could be combined with econometric estimations to be used as parameters for multi-criteria or cost-benefit analysis.

Bibliography

- Abril, M., F. Barber, et al. (2008). "An Assessment of Railway Capacity." *Transportation Research Part E: Logistics and Transportation Review* 44(5): 774-806.
- Akermann, H. R. and T. Bueker (2008). "Ziel Pünktlichkeit: Fahrplanstabilitätsuntersuchung für die SBB." *Eisenbahntechnische Rundschau (ETR)* 57(11): 744-747.
- Albrecht, T. (2009). "The Influence of Anticipating Train Driving on the Dispatching Process in Railway Conflict Situations." *Journal of Networks and Spatial Economics* 85-101: 63-84.
- Barber, F., M. Abril, et al. (2007). Survey of automated Systems for Railway Management. U. P. d. V. Departamento de Sistemas Informaticos y Computacion. Valencia. (*unpublished*)
- Barber, F., P. Tormos, et al. (2006). A Decision Support System for Railway Timetabling (MOM): The Spanish Case. *In: J. Allan, Brebbia, C.A., Hill, R.J., Sciutto, G., and Sone, S.. Computers in Railways X*, WIT Press, Southampton: 235-244.
- Brown, S. and R. Watson (2005). Rail Operations Modelling and Analysis. *Conference on Excellence in Railway Systems Engineering and Integration*. Derby, UK.
- Brünger, O. (1996). *Konzeption einer Rechnerunterstützung für die Feinkonstruktion von Eisenbahnfahrplänen*, RWTH Aachen, PhD Thesis
- Buchmueller, S., U. Weidmann, et al. (2008). Development of a dwell time calculation model for timetable planning. *In: J. Allan, E. Arias, C. A. Brebbia et al., Computers in Railways XI*, WIT Press, Southampton, 525-534.
- Bueker, T. and E. Wendler (2009). Efficient modelling of delay distribution functions. *3rd International Seminar on Railway Operations Modelling and Analysis*. Zurich.
- Burkolter, D. (2005). *Capacity of Railways in Station Areas using Petri Nets*, ETH Zuerich, PhD Thesis.
- Burkolter, D., T. Herrmann, et al. (2008). Generating Dense Railway Schedules. *Advanced OR and AI Methods in Transportation*, Publishing House of Poznan University of Technology, 290-297.
- Cacchiani, V., A. Caprara, et al. (2005). *PARTNER WP3 Deliverable*. European Project PARTNER.

- Caimi, G. (2009). *Algorithmic decision support for train scheduling in a large and highly utilised railway network*, ETH Zurich, PhD Thesis
- Caimi, G., D. Burkolter, et al. (2005). Finding Delay-Tolerant Train Routings through Stations. *Operations Research Proceedings 2004*, Springer, 136-143.
- Caimi, G., D. Burkolter, et al. (2007). Design of a Railway Scheduling Model for Dense Services. *2nd International Seminar on Railway Operations Modelling and Analysis (RailHannover 2007)*. Hannover.
- Caimi, G., M. Fuchsberger, et al. (2009). Conflict-free train scheduling in a compensation zone exploiting the speed profile. *3rd International Seminar on Railway Operations Modelling and Analysis*. Zurich.
- Caimi, G., M. Laumanns, et al. (2009). The Periodic Service Intention as a Conceptual Frame for Generating Timetables with Partial Periodicity. *3rd International Seminar on Railway Operations Modelling and Analysis*. Zurich.
- Camus, R., G. Longo, et al. (2005). "Estimation of Transit Reliability Level-of-Service Based on Automatic Vehicle Location Data." *Transportation Research Record: Journal of the Transportation Research Board No 1927 TRB*, National Research Council, Washington, D.C.(No 1927).
- Capek, K. (2004). *FlexFahrplan - A new approach to railway operational planning*. Technical University of Praha. Master Thesis (unpublished)
- Carey, M. and S. Carville (2003). "Scheduling and platforming trains at busy complex stations." *Transportation Research Part A: Policy and Practice* 37(3), 195-224.
- Conte, C. (2007). *Identifying Dependencies Among Delays*. Georg-August-Universitaet zu Goettingen, Georg-August-Universitaet zu Goettingen. PhD Thesis.
- Corman, F., A. D'Ariano, et al. (2009). Effectiveness of dynamic reordering and rerouting of trains in a complicated and densely occupied station area *3rd International Seminar on Railway Operations Modelling and Analysis*, Zurich.
- Crenca, D., G. Malavasi, et al. (2006). Effects on carrying capacity of railway lines equipment and configuration. *Conference on "Improvement and broadening of the Rail Service"*. Turin.
- Curchod, A. and L. Lucchini (2001). *CAPRES: General Description of the Moede*, Ecole Polytechnique Federale de Lausanne (EPFL), Lausanne. Document ITEP 788/6.
- D'Ariano, A. (2006). *Improving Real-Time Train Dispatching: Models, Algorithms and Applications*, TRAIL Thesis Series no. T2008/6, The Netherlands TRAIL Research School PhD Thesis.
- D'Ariano, A., D. Pacciarelli, et al. (2007). "A branch and bound algorithm for scheduling trains in a railway network." *European Journal of Operational Research* 183: 643-657.
- D'Ariano, A. and M. Pranzo (2009). "An Advanced Real-Time Train Dispatching System for Minimizing the Propagation of Delays in a Dispatching Area Under Severe Disturbances." *Journal of Networks and Spatial Economics* 9, 63-84.
- Daamen, W., R. M. P. Goverde, et al. (2009). "Non-Discriminatory Automatic Registration of Knock-On Train Delays." *Journal of Networks and Spatial Economics* 9, 47-61.

- de Fabris, S., G. Longo, et al. (2008). Automated analysis of train event recorder data to improve micro-simulation models. In: J. Allan, E. Arias, C. A. Brebbia et al., *Computers in Railways XI*, WIT Press, Southampton, 575-585.
- de Keizer, B. (2005). Analysis of the capacity use of the rail yard of Rotterdam Centraal, *1st International Seminar on Railway Operations Modelling and Analysis*. Delft.
- De Kort, A. F., B. Heidergott, et al. (1999). *Train Movement Analysis at Railway Stations: Procedures & Evaluation of Wakob's Approach*, Research Report, Delft University Press.
- De Vita, M. G. and G. Malavasi (2006). Evaluation of network extension in railway traffic simulation models by means of operational indicators *Conference on "Improvement and broadening of the Rail Service"*. Torino.
- Dekker, R., L. G. Kroon, et al. (2003). Reliability and heterogeneity of railway services. ERIM Report Series Research in Management, Rotterdam, Erasmus Research Institute of Management (ERIM)
- Fénix, J., T. Graffagnino, et al. (2005). User centred design applied to increase timetable stability, *36. Tagung Modernschnenfahrzeuge*. Graz (A).
- Fenner, W. and P. Naumann (1998). *Verkehrssicherungstechnik*. Erlangen, MCD Verlag.
- Ferchland, C. and T. Körner (2004). "Analytische Verfahren der Eisenbahnbetriebswissenschaft." *Eisenbahntechnische Rundschau (ETR)* 53(7/8): 499-505.
- Ferchland, C. and C. R. Warninghoff (2004). "Nutzung von Simulationen zur Unterstützung der betrieblichen Infrastrukturplanung." *Eisenbahntechnische Rundschau (ETR)* 53(7/8): 499-505.
- Flier, H., T. Graffagnino, et al. (2009). Finding Robust Train Paths in Dense Corridors. *3rd International Seminar on Railway Operations Modelling and Analysis*. Zurich.
- Gooßmann, R. and T. Polzin (2006). RTCSIM: Ein innovativer, erweiterbarer Rechenkern für Fahrplanvalidierung und Simulation. *Eisenbahntechnisches Kolloquium Darmstadt*.
- Goverde, R., W. Daamen, et al. (2007). Automatic Identification of Secondary Delays based on Train Descriptor Systems. *Forms/Format '07 Conference*. Branschweig.
- Goverde, R. M. P. (2005). *Punctuality of Railway Operations and Timetable Stability Analysis*, TRAIL Thesis Series no. T2005/10, The Netherlands TRAIL Research School PhD.
- Goverde, R. M. P., Bovy, P. H. L. et al. (1999). *The max-plus algebra approach to transportation problems*, *8th World Congress on Transport Research (WCTR)*, July 1998, Antwerp
- Goverde, R. M. P., Daamen, W. et al. (2008). Automatic identification of route conflict occurrences and their consequences. In: J. Allan, E. Arias, C. A. Brebbia et al., *Computers in Railways XI*, WIT Press, Southampton, 575-585.
- Groeger, T. (2002). *Simulation der Fahrplanerstellung auf der Basis eines hierarchischen Trassenmanagements und Nachweis der Stabilität der Betriebsabwicklung*. Veröffentlichungen des Verkehrswissenschaftlichen Institutes der RWTH Aachen. PhD Thesis.
- Hansen, I. A. (2004). Increase of capacity through optimised timetabling. In: J. Allan, Brebbia, C.A., Hill, R.J., Sciutto, G., and Sone, S., *Computers in Railways IX*. Southampton, WIT Press, 529-538.

- Hansen, I. A. (2005). Current State of the Art of Railway Operations Research. *Conference on Excellence in Railway Systems Engineering and Integration*. Derby, UK.
- Hansen, I. A. and Nie, L. (2005). "System analysis of train operations and track occupancy at railway stations." *European Journal of Transport and Infrastructure Research* 5(1): 31-54.
- Hansen, I. A. and J. Pacht (2008). *Railway Timetable & Traffic*. Hamburg, Eurailpress.
- Hauptmann, D. (2000). *Automatische und diskriminierungsfreie Ermittlung von Fahrplantrassen in beliebig großen Netzen spurgeführter Verkehrssysteme*. Hamburg, Eurailpress, PhD Thesis.
- Hedinger, E. (1996). *Qualitätsmanagement von Eisenbahnstrecken*, ETH Zuerich, PhD Thesis.
- Herrmann, T. (2005). *Stability of Timetables and Train Routings through Station Regions*. ETH Zuerich, PhD Thesis
- Hofman, M. A. and Madsen, L. F. (2005). *Robustness in train scheduling* IMM. Copenhagen, Technical University of Denmark. Master Thesis
- Huerlimann, D. (2001). *Objektorientierte Modellierung von Infrastrukturelementen und Betriebsvorgängen im Eisenbahnwesen*, ETH Zuerich, PhD Thesis
- Huisman, D., L. G. Kroon, et al. (2005). Operations Research in Passenger Railway Transportation ERIM Report Series Research in Management N. 32(ERIM). Rotterdam
- Huisman, T., R. J. Boucherie, et al. (2002). "A solvable queueing network model for railway networks and its validation and applications for the Netherlands." *European Journal of Operational Research* 142(1): 30-51.
- IVE (2005). *NEMO - Netz Evaluations Modell*. Institut fuer Verkehrswesen, Universitaet Hannover and I. f. V.-u. E. mbH. Hannover.
- Jacobs, J. (2003). *Rechnerunterstützte Konfliktermittlung und Entscheidungsunterstützung bei der Disposition des Zuglaufs*. Veroeffentlichungen des Verkehrswissenschaftlichen Institutes der RWTH Aachen. PhD Thesis.
- Jacobs, J. (2004). Reducing Delays by Means of Computer-Aided 'on-the-Spot' Rescheduling. In: J. Allan, Brebbia, C.A., Hill, R.J., Sciutto, G., and Sone, S., *Computers in Railways IX*. . Southampton, WIT Press, 615-625.
- Jansen, D. N., Klages, S.G. et al. (2008). The impact of GSM-R on railway capacity. In: J. Allan, E. Arias, C. A. Brebbia et al., *Computers in Railways XI*, WIT Press, Southampton, 223-241.
- Jentsch, E. (2007). "Aufwandsreduzierte Fahrzeitermittlung." *Der Eisenbahningenieur* 57(8): 10-16.
- Jespersen-Groth, J., D. Potthoff, et al. (2007). *Disruption management in passenger railway transportation*. IMM-Technical Report-2007-3, Informatics and Mathematical Modelling, Technical University of Denmark.
- Jochim, H. E. (2001). Die Kapazität von Engpassstrecken der Eisenbahn bei der Trennung von Netz und Betrieb. 18. *Verkehrswissenschaftliche Tage* Dresden.
- Kaminsky, R. (2001). *Pufferzeiten in Netzen des spurgeführten Verkehrs in Abhängigkeit von Zugfolge und Infrastruktur*. Hamburg, Eurailpress. PhD Thesis

- Kanacilo, E. M. and A. Verbraeck (2005). Simulation Services to support the control design of rail infrastructures. Proceedings of the 2006 Winter Simulation Conference, Monterey (California)
- Kittelson & Associates, I., I. KFH Group, et al. (2004). *TCRP Report 100: Transit Capacity and Quality of Service Manual*, 2nd ed. Washington, D.C.
- Klahn, V. (1992). "Die betriebliche Simulation von Eisenbahnstrecken." *Eisenbahntechnische Rundschau (ETR)* 41(3): 163-170.
- Klahn, V. (1992). "Graphnetheoretische Grundlagen zur Abbildung grosser Eisenbahnbereiche." *Eisenbahntechnische Rundschau (ETR)* 41(10): 659-664.
- Klahn, V. (1994). "Die betriebliche Simulation grosser Eisenbahnnetze." *Eisenbahntechnische Rundschau (ETR)* 43(10), 667-672.
- Koesters, D. and Limmer, B. (2006). Synopsis of Capacity Management and Analysis in Air and Rail Traffic. *GARS Amsterdam Junior Researchers' Workshop*.
- Kohls, H. (2007). *Eastfields Station - Railsys Performance Modelling*, RWA Rail, (unpublished)
- Korte, K. (1998). *Entwicklung eines Infrastruktur-Datenbanksystems für den spurgeführten Verkehr*. Universitaet Hannover, PhD Thesis
- Kroon, L. G., Dekker, L. et al. (2006). *Stochastic Improvement of Cyclic Railway Timetables*. ERIM Report Series Research in Management (ERIM). Rotterdam.
- Kroon, L. G., Huisman, D. et al. (2007). *Railway Timetabling from an Operations Research Perspective*. ERIM Report Series Research in Management (ERIM), Rotterdam.
- Kroon, L. G. and Peeters, L. W. P. (2003). *A Variable Trip Time Model for Cyclic Railway Timetabling*, *Transportation Science* 37 (2), 198-217.
- Kroon, L. G., Vromans, M. J. C. M. et al. (2005). *Cyclic Railway Timetabling: a Stochastic Optimization Approach* ERIM Report Series Research in Management (ERIM), Rotterdam.
- Kuckelberg, A. (2004). Component Based System Architecture for Railway Capacity Management Systems In: *Proceedings of the Fourth International Conference on Quality Software (QSIC'04)*
- Kuckelberg, A. and Wendler, E. (2008). Real-time asynchronous conflict solving algorithms for computer aided train dispatching assistance systems. In: J. Allan, E. Arias, C. A. Brebbia et al., *Computers in Railways XI*, WIT Press, Southampton, 555-563.
- Landex, A. and Kaas, A. H. (2005). Planning the most suitable travel speed for high frequency railway line, *1st International Seminar on Railway Operations Modelling and Analysis*. Delft.
- Landex, A. and Kaas, A. H. (2009). Examination of Operation Quality for High-frequent Railway Operation, *3rd International Seminar on Railway Operations Modelling and Analysis*. Zurich.
- Landex, A., Kaas, A. H. et al. (2006). Practical use of the UIC 406 capacity leaflet by including timetable tools in the investigations. In: J. Allan, Brebbia, C.A., Hill, R.J., Sciutto, G., and Sone, S., *Computers in Railways X*. Southampton, WIT Press: 235-244.

- Landex, A., Kaas, A. H. et al. (2006). Evaluation of railway capacity. *Trafikdage på Aalborg Universitet Aalborg* (DK).
- Laube, F., S. Roos, et al. (2007). "PULS 90 – Ein systemumfassender Ansatz zur Leistungssteigerung von Eisenbahnnetzen." *Eisenbahntechnische Rundschau (ETR)* 55 (3): 105-107.
- Liebchen, C. (2006). "Fahrplanoptimierung im Personenverkehr – Potenzial mathematischer Optimierung." *Eisenbahntechnische Rundschau (ETR)* 55(7/8): 504-510.
- Liebchen, C., Schachtebeck, M. et al. (2007). *Computing Delay Resistant Railway Timetables*. Reports from the group "Combinatorial Optimization and Graph Algorithms" T. B. Department of Mathematics. 2007/17.
- Lindfeldt, O. (2007). Quality on single-track railway lines with passenger traffic - Analytical model for evaluation of crossing stations and partial double-tracks. Division of Traffic and Logistics, Department of Transport and Economics Stockholm, KTH Royal Institute of Technology. PhD Thesis.
- Lindfeldt, O. (2008). Evaluation of punctuality on a heavily utilised railway line with mixed traffic. In: J. Allan, E. Arias, C. A. Brebbia et al., *Computers in Railways XI*, WIT Press, Southampton, 545-554.
- Longo, G. and Saffi, F. (1999) Elevate densità di circolazione e regolarità di marcia sulle linee ferroviarie - *Ingegneria Ferroviaria* n. 5/1999.
- Luethi, M. (2008). Optimizing traffic flow in heavily used railway networks: influence factors and potential strategies. *8th Swiss Transport Research Conference*, Ascona (CH).
- Luethi, M. (2009). *Improving the Efficiency of Heavily Used Railway Networks through Integrated Real-Time Rescheduling*. Dissertation Nr. 18615, ETH Zuerich. PhD Thesis
- Luethi, M., Huerlimann, D. et al. (2005). Understanding the Timetable Planning Process as a Closed Control Loop. *1st International Seminar on Railway Operations Modelling and Analysis*. Delft.
- Luethi, M., Medeossi, G. et al. (2007). Evaluation of an Integrated Real-Time Rescheduling System for Railway Networks. *2nd International Seminar on Railway Operations Modelling and Analysis* (RailHannover 2007). Hannover.
- Luethi, M., Medeossi, G. et al. (2009). "Structure and Simulation Evaluation of an Integrated Real-Time Rescheduling System for Railway Networks." *Journal of Networks and Spatial Economics* 9(1): 103-121.
- Luethi, M., Nash, A. et al. (2007). Increasing railway capacity and reliability through integrated real-time rescheduling. *Proceedings of the 11th World Conference on Transport Research*. Berkeley.
- Luethi, M., Weidmann, U. et al. (2007). Rescheduling and train control a new framework for railroad traffic control in heavily used networks. *86th Transportation Research Board Annual Meeting*. Washington D.C.
- Lukaszewicz, P. (2001). Energy Consumption and Running Time for Trains. Department of Vehicle Engineering. Stockholm, Royal Institute of Technology. PhD.
- Mattsson, L. G. (2005). Train service reliability. A survey of methods for deriving relationships for train delays. R. I. o. T. Department of Infrastructure. Stockholm.

- Mazzarello, M. and E. Ottaviani (2005). A Traffic Management System for Real-Time Traffic Optimisation in Railways. 1st International Seminar on Railway Operations Modelling and Analysis. Delft.
- Medeossi, G. (2006). Una Nuova Metodologia per l'incremento della Capacità di Reti Ferroviarie. Trieste, University of Trieste. MSc.
- Meester, L. E. and S. Muns (2005). Stochastic delay propagation in railway networks and phase-type distributions. 1st International Seminar on Railway Operations Modelling and Analysis. Delft.
- Middelkoop, L. G. and M. BAouwman (2000). Train network simulation for support of network wide planning of infrastructure and timetables. Computers in Railways VII. J. Allan, R. J. Hill, C. A. Brebbia, G. Sciutto and S. Sone, WIT Press, Southampton: 267-276.
- Montigel, M. (1994). Modellierung und Gewährleistung von Abhängigkeiten in Eisenbahnsicherungsanlagen, ETH Zuerich. PhD Thesis
- Morlok, E. K. and Nitzberg, B. F. (2003). "Speeding up Commuter Rail Service: Comparative Actual Performance of Different Train and Station Platform Designs." *Transportation Research Record: Journal of the Transportation Research Board TRB*, National Research Council, Washington, D.C.
- Musto, M. (2007). *The Role of Simulation in Railway Capacity Allocation*. Department of Civil and Environmental Engineering, University of Helsinki. Master Thesis.
- Nyström, B. (2008). *Aspects of improving punctuality: from data to decision in railway maintenance*, Lulea University of Technology, PhD Thesis
- Ogawa, H., Kawai, S. et al. (2008). Day-to-day tuning of the Tokaido Shinkansen timetable – carried out for every single operating hour throughout the year. In: J. Allan, E. Arias, C. A. Brebbia et al., *Computers in Railways XI*, WIT Press, Southampton, 515-524.
- Pachl, J. (2002). *Railway Operations and Control*. VTD Rail Publishing, Mountlake Terrace.
- Pachl, J. (2005). Modelling Specific Signaling Features in Computer-based Scheduling Systems. 1st International Seminar on Railway Operations Modelling and Analysis. Delft.
- Pachl, J. (2008). *Systemtechnik des Schienenverkehrs: Bahnbetrieb planen, steuern und sichern* Teubner Verlag, Wiesbaden.
- Pachl, J. and White, T. (2004). Analytical Capacity Management With Blocking Times, 83th *Transportation Research Board Annual Meeting*. Washington D.C.
- Pead, M. (2007). *The Impact of Boarding and Alighting Passengers on the Dwell Time at Railway Stations*. Ashton University.
- Peeters, L. W. P. (2003). *Cyclic Railway Timetable Optimization*. TRAIL Thesis series T2003/5, The Netherlands TRAIL Research School. PhD Thesis.
- Putallaz, Y. and R. Rivier (2004). Strategic evolution of railways corridors infrastructure: dual approach for assessing capacity investments and M&R strategies In: J. Allan, Brebbia, C.A., Hill, R.J., Sciutto, G., and Sone, S. *Computers in Railways IX*. Southampton, WIT Press,

- Putillaz, Y. (2007). *Gestion Strategique de la Maintenance et de la Capacité d'un Reseau Ferrè*. Faculté de l'environnement naturel, architectural et construit, EPFL Lausanne. PhD Thesis.
- Radtke, A. and D. Hauptmann (2002). Automated planning of timetables in large railway networks using a microscopic basis and railway simulation techniques. In: J. Allan, Brebbia, C.A., Hill, R.J., Sciotto, G., and Sone, S. *Computers in Railways IX*. Southampton, WIT Press,
- Radtke, A. and R. Rudolph (2006). Optimisation of Allowances in Railway Scheduling. *World Conference on Railway Research 2006*, Montreal.
- RMCon (2008). *RailSys - Network wide Timetable and Infrastructure Management Integrated network wide Operational and Timetable Simulation*. R. M. Consultants. Hannover.
- Roos, S. (2006). *Bewertung von Knotenmanagement-Methoden fuer Eisenbahnen*. Institut fuer Verkehrsplanung und Tranportsysteme, ETH Zurich. Master Thesis.
- Rudolph, R. (2003). Allowances and Margins in Railway Scheduling. *World Conference on Railway Research 2003*, Edinburgh.
- Rudolph, R. (2004). *Entwicklung von Strategien zur optimierten Anordnung und Dimensionierung von Seitzuschlaegen im Eisenbahnbetrieb*. Hamburg, Eurailpress. PhD Thesis
- Rudolph, R. and J. Demnitz (2003). Simulation of large railway networks. Robustness test for the timetable 2003 in NorthRhine Westphalia. *Proceedings of World Conference on Railway Research 2003*. Edinburgh: 644-652.
- Schaafsma, A. A. M. (2005). Dynamic Traffic Management – innovative solution for the Schiphol bottleneck 2007. *1st International Seminar on Railway Operations Modelling and Analysis*. Delft.
- Schaer, T., Jacobs, J. et al. (2005). "DisKon – Laborversion eines flexiblen, modularen und automatischen Dispositionsassistenzsystems." *Eisenbahntechnische Rundschau (ETR)* 54(12): 809-821.
- Schultze, K. (1985). Modell für die asynchrone Simulation des Betriebes in Teilen des Eisenbahnnetzes. Veroeffentlichungen des Verkehrswissenschaftlichen Institutes der RWTH Aachen. PhD Thesis.
- Schwanhäusser, W. (1974). Die Besessung der Pufferzeiten im Fahrplangefüge der Eisenbahn, Veroeffentlichungen des Verkehrswissenschaftlichen Institutes der RWTH Aachen. PhD Thesis.
- Sewcyk, B., A. Radtke, et al. (2007). Combining Microscopic and Macroscopic Infrastructure Planning Models *2nd International Seminar on Railway Operations Modelling and Analysis (RailHannover 2007)*. Hannover.
- Stalder, O. (2006). Bahn 2000 der SBB: Prozessveränderung als Wettbewerbsvorteil. *24. Horber Schienen-Tage*. Horb.
- Stalder, O. and F. Laube (2004). The Efficient Railway - Field of Action for formal Methods. FORMS / FORMAT 2004: Formal Methods for Automation and Safety in Railway and Automotive Systems. Braunschweig.

- Stalder, O., Laube, F. et al. (2003). Increasing Performance of the Rail Network in the heart of Europe: a Program for the Swiss Federal Railways. *International Symposium on Speed-up and Service Technology for Railway and Maglev Systems 2003 (STECH'03)*. Tokio.
- Stock, R. and Longo, G. (2007). Estimation of railway capacity using stochastic differential equations. *2nd International Seminar on Railway Operations Modelling and Analysis (RailHannover 2007)*. Hannover.
- Stok, R. (2008). *Estimation of railway capacity consumption using stochastic differential equations*. Dept. of Civil and Environmental Engineering. Trieste, University of Trieste. PhD Thesis.
- Suh, S. D., Yang, K. Y. et al. (2005). "Allocation Of Rail Line Capacity Between Ktx And Conventional Trains Under Different Policy Goals With Mathematical Programming." *Proceedings of the Eastern Asia Society for Transportation Studies* 5: 175-189.
- Theeg, G. and S. e. Vlasenko (2009). *Railway Signalling and Interlocking - International Compendium*. Eurailpress, Hamburg.
- Uhlmann, M. and Mutschink, K. (2004). "Bemessung komplexer Eisenbahninfra- struktur – Die konstruktive Methode." *Eisenbahntechnische Rundschau (ETR)* 53(7/8): 506-514.
- UIC, (1996). *Leaflet 405 OR - Links between Railway Infrastructure Capacity and the Quality of Operations*. International Union of Railways (UIC)
- UIC, (2000). *Leaflet 451 OR - Timetable recovery margins to guarantee timekeeping - Recovery margins*. International Union of Railways (UIC)
- UIC, (2000). *Leaflet 451-2 OR - Coordination of work sites and operating measures to be taken on main lines particularly for international traffic*, International Union of Railways (UIC)
- UIC, (2004). *Leaflet 406 - Capacity*. International Union of Railways (UIC)
- Ullius, M. (2004). *Verwendung von Eisenbahnbetriebsdaten für die Schwachstellen- und Risikoanalyse zur Verbesserung der Angebots- und Betriebsqualität*, ETH Zuerich. PhD Thesis
- Vakhtel, S. (2002). *Rechnerunterstützte analytische Ermittlung der Kapazität von Eisenbahnnetzen*. Veröffentlichungen des Verkehrswissenschaftlichen Institutes der RWTH Aachen. PhD Thesis.
- Vromans, M. J. C. M. (2005). *Reliability of Railway Systems*. TRAIL Thesis series T2005/7, The Netherlands TRAIL Research School, PhD Thesis.
- Vromans, M. J. C. M., Dekker, R. et al. (2006). "Reliability and heterogeneity of railway services." *European Journal of Operational Research* 172(2): 647-665.
- Waas, K. and Bueker, T. (2008). "Effiziente Qualitätsanalyse von Fahrplankonzepten." *Eisenbahntechnische Rundschau (ETR)* 57(11): 744-747.
- Wakob, H. (1985). *Wartemodells zur Ermittlung der planmäßigen Wartezeiten im Eisenbahnbetrieb unter besonderer Berücksichtigung der Aspekte Leistungsfähigkeit und Anlagenbelastung*. Veröffentlichungen des Verkehrswissenschaftlichen Institutes der RWTH Aachen. PhD Thesis.
- Watson, R. (2008). *Train Planning in a fragmented railway - A British Perspective*. Loughborough, University of Loughborough, PhD Thesis.

- Weeda, V. A. and Hofstra, K. S. (2008). Performance analysis: improving the Dutch railway service. In: J. Allan, E. Arias, C. A. Brebbia et al., *Computers in Railways XI*, WIT Press, Southampton, 463-472.
- Wegele, S., Corman, F. et al. (2008). Comparing the effectiveness of two real-time train rescheduling systems in case of perturbed traffic conditions. In: J. Allan, E. Arias, C. A. Brebbia et al., *Computers in Railways XI*, WIT Press, Southampton, 535-544.
- Wegele, S., Slovák, R., et al. (2006). "Echtzeitoptimierung für die Disposition im Schienenverkehr." *Signal+Draht (S+D)* 98(6): 6-10.
- Weigand, W. and T. Körner (2007). "Das Wachstum des Schienengüterverkehrs – Herausforderung für die DB Netz AG." *Eisenbahntechnische Rundschau (ETR)* 57(3): 94-103.
- Wende, D. (2001). "Fahrdynamische Bewertung innovativer Bremstechnik." *Eisenbahntechnische Rundschau (ETR)* 50(11): 685-691.
- Wende, D. (2003). *Fahrdynamik des Schienenverkehrs*. Wiesbaden, Teubner Verlag.
- Wendler, E. (1999). Analytische Berechnung der planmässigen Wartezeiten bei asynchroner Fahrplankonstruktion, Veröffentlichungen des Verkehrswissenschaftlichen Institutes der RWTH Aachen. PhD Thesis.
- Wendler, E. (2002). "Bemessungsmethoden für große Eisenbahnknoten." *Eisenbahntechnische Rundschau (ETR)* 51(7/8): 418-424.
- Wendler, E. (2007). "The scheduled waiting time on railway lines." *Transportation Research Part B: Methodological* 41(2): 148-158.
- Wendler, E. (2008). "Stochastische Modelle in der Eisenbahnbetriebswissenschaft." *ZEVrail Glasers Annalen* 132(1-2): 40-48.
- Weymann, F., A. Kuckelberg, et al. (2008). Coupling of synchronous and asynchronous methods in the simulation of railway operation. *10th International Conference on Applications of Advanced Technologies in Transportation (AATT)*. Athens.
- Yuan, J. (2006). *Stochastic modelling of train delays and delay propagation in stations*, TRAIL Thesis Series no. T2006/6, The Netherlands TRAIL Research School, PhD Thesis.
- Yuan, J. and Hansen, I. A. (2007). "Optimizing capacity utilization of stations by estimating knock-on train delays." *Transportation Research Part B: Methodological* 41(2): 202-217.