Quality on single-track railway lines with passenger traffic - Analytical model for evaluation of crossing stations and partial double-tracks

Olov Lindfeldt

Licentiate thesis
Division of Traffic and Logistics
Department of Transport and Economics
Stockholm 2007
Cover picture: Crossing at Tällberg (Sweden), February 2007. Note that the exit signal at the station limit is already clear, and so the locking time for the coming exit train path is minimized.
Abstract
Railway transportation is showing a substantial increase. Investments in new infrastructure, new fast and comfortable vehicles, and high frequency of service are important factors behind the increase.

Infrastructure configuration and timetable construction play important roles in the competitiveness of railway transportation. This is especially true on single-track lines where the travel times and other timetable related parameters are severely restricted by crossings (train meetings). The crossings also make the lines’ operation more sensitive to disturbances.

Since the major part of the Swedish railway network is single-track it is of great interest to examine the relationships between operation properties, such as travel times and reliability, and infrastructure configuration on single-track lines. The crossings are the core feature of single-track operation and this thesis focuses on the crossing time, i.e. the time loss that occurs in crossing situations.

A simplified analytical model, SAMFOST, has been developed to calculate the crossing time as a function of infrastructure configuration, vehicle properties, timetable and delays for two crossing trains. The effect of possible surrounding trains is not taken into account and all kinds of congestion effects are thus excluded from evaluation. SAMFOST has been successfully validated against the simulation tool RailSys, which shows that this type of simplified model is accurate in non-congested situations.

A great advantage of disregarding congested situations is that analysis is independent of timetable assumptions. The model also explicitly shows the effect of punctuality, which is of particular importance on single-track lines where the interdependencies between trains are strengthened by the crossings.

For the same reason, the timetable is severely constrained. Nonetheless, there is often a need for changes of the timetable (crossing pattern). The thesis proposes three simple measures of timetable flexibility, all based on assigned crossing time requirements. Together, these measures can be used to evaluate how infrastructure configuration, vehicle properties, punctuality etc affect possibilities to alter the timetable.

As an example of its application, SAMFOST has been used to evaluate the effect of shorter inter-station distance, partial double-track and combined crossing and passenger stop. These measures affect the operational properties quite differently.

More crossing stations result in a minor decrease in travel time (lower mean crossing time) but significantly higher reliability (lower crossing time variance). These effects are independent of punctuality, which is a valuable property.

A partial double-track results in shorter travel times and in some cases also higher reliability. Both effects are strongly dependent on punctuality and high punctuality is needed to achieve high effects.
A combined crossing and passenger stop results in a situation similar to that of a partial double-track. In this case it is important to point out that the assignment of time supplements in the timetable should be directly correlated to punctuality in order to achieve good operation.
Preface

This licentiate thesis is based on research performed between 2004 and 2007 at the Division of Traffic and Logistics at the Royal Institute of Technology (KTH).

The thesis consists of two parts; an introductory essay and three papers that form the basis of the thesis:


The results are also presented in the research report “Effekter av partiella dubbelspår och fler mötesstationer på enkelspår”, Lindfeldt 2007 (in Swedish).
Acknowledgements

Banverket (the Swedish National Rail Administration) funded this research.

I would like to thank everyone who has been involved in this project. My supervisors Professor Bo-Lennart Nelldal and Professor Lars-Göran Mattsson, thank you for supervision, encouragement, support and advice. The members of the reference group, Magnus Wahlborg, Magdalena Grimm, Hans Dahlberg, Håkan Berell from Banverket (the Swedish National Rail Administration) and Sirpa Holmroos from SJ AB have given me valuable guidance and support.

Dr Piotr Lukaszewicz from KTH Railway Technology provided inspiring comments on my thesis at the final seminar and Dr Björn Kufver from Ferroplan supported me at the conferences in Prague and London. Anders Lindfeldt MSc, my youngest brother, showed me the usefulness of Matlab and helped me in several tricky Matlab-programming situations. Carl Grauers, train driver at Inlandsgods AB, provided valuable information about train driving, signalling etc.

Thanks also to my parents who introduced the railway to me in my early years. I also want to send a special thought to my father who died during the final phase of the thesis work. His skills in mathematics and science really inspired me to study at KTH, both for my master’s and my doctor’s degree.

Stockholm, November 2007

Olov Lindfeldt
Contents

1 Introduction ......................................................................................................... 1  
  1.1 BACKGROUND ........................................................................................................ 1  
  1.2 OBJECTIVES ........................................................................................................... 2  
  1.3 DELIMITATIONS ...................................................................................................... 3  
2 Related research .................................................................................................. 5  
  2.1 SINGLE-TRACKS ....................................................................................................... 5  
  2.2 KNOCK-ON DELAYS ................................................................................................. 6  
  2.3 RESCHEDULING AND DELAY MANAGEMENT ............................................................. 7  
  2.4 CONCLUDING REMARKS .......................................................................................... 8  
3 Methodology ........................................................................................................ 9  
  3.1 MODELLING SINGLE-TRACK OPERATION ................................................................. 10  
4 Results .................................................................................................................. 13  
  4.1 INFLUENCES OF STATION LENGTH (PARTIAL DOUBLE-TRACK) AND INTER-STATION DISTANCE ON DELAYS AND DELAY PROPAGATION ON SINGLE-TRACK LINES WITH REGIONAL RAIL TRAFFIC ................................................... 13  
  4.2 SAMFOST – A TIMETABLE-FREE WAY OF ANALYSING SINGLE-TRACK RAILWAY LINES ................................................................................................................. 16  
  4.3 CROSSING TIMES ON SINGLE-TRACK RAILWAY LINES – DEPENDENCIES OF DIFFERENT INFRASTRUCTURE AND TRAFFIC FACTORS ............................................. 23  
5 Discussion of the main contributions of the thesis ............................................. 31  
  5.1 SIMPLE MODELS ...................................................................................................... 31  
  5.2 PUNCTUALITY DEPENDENCIES .............................................................................. 31  
  5.3 INFRASTRUCTURE AS A VARIABLE ......................................................................... 32  
  5.4 TIMETABLE FREE ANALYSES AND TIMETABLE FLEXIBILITY .................................. 32  
  5.5 TOOL FOR INFRASTRUCTURE AND TIMETABLE PLANNING .................................. 33  
6 Final remarks and future work ........................................................................ 37  
7 References ............................................................................................................ 39
1 Introduction

1.1 Background

Transportation is a large and important part of the economy and the need for transportation increases continuously. Road traffic is the major form of transportation. However, a combination of investments in railway infrastructure, new, fast and comfortable rail vehicles, and increased frequency of service has led to an extensive expansion of rail passenger traffic. The Mälar valley and the Sound region are good examples of geographical areas where travel by train has increased dramatically.

Along with increasing environmental awareness and a political desire to reduce emissions, the railway has a good opportunity to increase its market share and contribute to a sustainable society. Such an increase is strongly dependent on its competitiveness and so rail services need to be fast, frequent, comfortable and reliable.

These competitive factors, in turn, depend on technical properties in the railway system:

- Infrastructure layout and operational reliability.
- Vehicle design and operational reliability.
- Timetable.

In order to achieve fast, frequent and reliable services it is necessary to understand the relationships between infrastructure, vehicle, timetable and disturbances (lack of reliability). This thesis deals with some of these relationships on single-track lines.

Regarding speed, frequency and reliability of services, single-track lines exhibit special properties, most of them tightly connected to crossings. On single-track lines, with only ordinary crossing stations, each crossing implies longer running times. The crossings also imply decreased reliability since delays propagate between crossing trains. The limited crossing possibilities also constrain capacity and thereby also the frequency of services on single-track lines.

In order to increase the competitiveness of the railway transportation system, measures that limit these effects are of special interest. Examples of such measures are:

- Transformation of single-track into complete double-track.
- Partial transformation of single-track into double-track.
- Increased number (higher density) of crossing stations.
- Crossing combined with regular passenger stop.
Greater punctuality.

A complete transformation of a single-track line into a double-track means a complete elimination of the problems related to crossings and results in a great leap in capacity as well as a decrease in travel time and disturbance sensitivity.

Since this type of investment is extensive, it is not always justifiable to go directly from a single-track to a complete double-track. In order to decrease the leaps and achieve a better match between investments and demand it is of great interest to consider less expensive measures that still, completely or partly, result in the desired effects. This is possible in some cases, though at the cost of different kinds of restrictions on the traffic. As traffic demand develops such restrictions will become more and more troublesome and at some point the complete measure is motivated.

Partial double-tracks, a greater number of crossing stations and combined crossing/passenger stop, all affect the characteristics of the railway line so that the travel times and/or disturbance sensitivity decreases. Capacity is also affected, but this change may be difficult to utilize without losing some of the travel time gained and/or the decrease in disturbance sensitivity. In order to understand how these measures alter the railway system and under what conditions they actually work as required, it is necessary to examine the operational properties of both the original system and the adjusted system.

The railway suffers from low punctuality due to frequent primary delays. A major reduction in these delays would help to make the operation of single-track lines more reliable and contribute to a decrease in average travel times. This measure is systematically examined in the thesis through the use of different levels of primary delays.

1.2 Objectives

This thesis has several objectives. The overall objective is to clarify the operational properties of single-track railway lines, i.e. how the traffic is affected by infrastructure configuration, timetable and disturbances. This includes examination of some important factors in infrastructure and vehicles as well as punctuality and how these factors interact.

Punctuality, in the form of primary delays, is a factor of particular importance in the operation of single-track lines. A special objective is therefore to show and analyse sensitivity to disturbances that occur on single-track lines.

A more specific objective is a further examination of three measures that can be used to decrease the negative effects of crossings. The work aims to show how time for crossing, and its variance, is affected by partial double-track, shorter inter-station distances and crossing combined with passenger stop. These measures are all strongly linked to the infrastructure. Therefore, unlike in many other studies, the infrastructure has to be treated as a variable. A secondary objective that follows from this is to show how the infrastructure can be treated practically as a variable by fictive model lines that do not exist in reality.
The three measures not only affect the properties at isolated points on the studied line. Properties along the whole line, or on continuous sections of line, are changed, thereby also changing the conditions for alternative timetables. In order to take this change into account, timetable flexibility also has to be studied.

It is beyond the scope of this thesis to give a strict definition of timetable flexibility. However, one important objective is to make a first attempt to give some sort of definition of this important concept, including punctuality to emphasise its importance for timetable construction.

A methodological objective is to develop a simplified, mathematical traffic model and show an example of how such a model can be used. Simplifications require assumptions about traffic density, punctuality, dispatching rules, signalling systems, train movements etc. Given these assumptions the model helps to give important insights about system properties at a basic level. An important part of the work is to show under which conditions the model is accurate.

Using a mathematical model, fictive infrastructure configurations with a high degree of symmetries, as well as existing configurations with unique, non-symmetric properties, may be analysed. This type of comparison is also an aim, since analysing both an idealised case and a “real” case greatly increases understanding of the operational properties.

### 1.3 Delimitations

Only single-track railway lines, where crossings take place, are taken up. Nodes and all types of network and network effects are excluded. The analysed lines are assumed to be long (> 120 km) and without connections to other lines. Only passenger services are modelled and two different vehicle types are used: X50, a regional train with high acceleration and X2, a long distance train with lower acceleration. Both traffic directions have been operated by the same vehicle and mixed crossings are not analysed.

Capacity utilisation is considered to be moderate and the frequency of services does not exceed two trains/hour and direction.

All fictive model lines are constructed to obtain results that are easy to understand. Therefore, one standard station design is used and the inter-station distances are equal within each model line\(^1\). No gradients are modelled.

One type of signalling system is used: ERTMS, level 2. Continuous updating of driving permissions makes the trains behave in a logical and deterministic way. This delimitation eliminates the need to model ATC track antennas and their locations.

Punctuality is varied within the interval 100 – 350 seconds mean arrival (primary) delay. Better and poorer punctuality is not analysed. These limits are derived from

---

\(^1\) Partial double-tracks influence the flanking inter-station distances in a special way.
real delay distributions on the Swedish railway. In the analyses negative exponential distributions are principally used since they are easily adjusted for different punctuality levels.

Only situations without congestion effects are treated. This means that two following trains are always assumed to be more than approximately two inter-station distances apart, which implies that they are independent in crossing situations that appear with trains in the opposite direction. These delimitations in service frequency and punctuality make the probability of congestion very low (<1/200).

The modelled timetables contain no slack (time supplements) and so no recovery from delays is possible within the model. This delimitation is natural since the objective is to model crossings with disturbances. The need for supplements is rather a result of the analysis, since they are necessary for recovery from the delay propagation that occurs in crossing situations.

Driver behaviour is not modelled. All trains behave according to deterministic vehicle data and follow the same assumption regarding acceleration courses, use of maximum speed and deceleration courses.

Series of crossings are not modelled. However, this would be of great interest for further research. Such analyses need modelling of, or more detailed assumptions about, time supplements between the crossings.
2 Related research

The research in railway operation is extensive. Most modern research concerns double-tracked systems or entire networks of double-tracked lines and stations that connect them.

In many countries single-track lines are of minor interest since the traffic volumes correspond to double-track systems. For this reason the infrastructure related research on single-tracks is limited. Instead, most modern single-track literature is orientated towards timetable construction and reliability in existing infrastructure configurations.

Many studies deal with knock-on delays (delay propagation) and rescheduling. Although most of these studies are not applicable to single-tracks, they give a useful basic picture of the railway system, operation and possible modelling techniques.

Several types of analytical models are presented in the literature whereas simulation methods are more rarely discussed. Queuing theory and other types of statistical methods, where delay distributions are combined, are used to model interactions between delayed trains.

2.1 Single-tracks

Research on single-track systems generally concentrates on either infrastructure or timetable. This is a simplification since infrastructure and timetable are very closely interconnected. Analyses concerning just one of them therefore imply considerable assumptions as regards the other.

In [9] Petersen presents a simple model that can be used to describe the delay time as a function of the traffic intensity. The main timetable assumption, that makes the study general, is that departing times for trains are independent random variables that are uniformly distributed over the defined time period. Given this randomised timetable the time costs (delays) for crossings and overtakings are calculated. One conflict at a time is identified and resolved and so the trains are treated pair-wise. The most important advantage of models like this is that effects of changes in infrastructure parameters are easy to examine.

In [10] both timetable and line alignment are assumed to be known. Train performance is assumed to be deterministic and from this a method to find the best locations of crossing stations is presented. The track length at crossing stations is also considered and conditions for minimum crossing delay are examined. The method focuses on frequent small delays that can be managed by longer crossing stations (i.e. partial double-tracks) and time supplements. Longer delays are handled by secondary crossing stations. These may also be used for slower trains with low priority. The study concludes that single-tracks work quite well, as long as infrastructure and timetable are coherent and delays limited.

Higgins et al [6] describe a decomposition procedure that for a given cyclic timetable (day or week) for high-speed trains finds the numbers and positions of crossing
stations that minimise both the risk of delays and the delays caused by train conflicts. The timetable is specified only by information about earliest possible departure time of the trains. Each conflict is divided into two parts:

- Conflict delay: the part that is included in the timetable.
- Risk of delay: expected amount of delays due to unforeseen events.

The output from the model is both an optimal infrastructure and an optimal timetable. This combination gives a complete technical solution to the entire single-track problem. In most practical cases, however, there are other timetable-related constraints that ought to be included in the model.

In [5] Higgins et al develop their model further, resulting in a model that can be used both in dispatching and in long-term infrastructure planning. However, the model does not allow random delay events.

There are several examples of studies where the infrastructure design is fixed and the timetable is somehow constructed according to infrastructure constraints (and market demand). An early example of this is the mathematical treatment of two-way traffic on a single-track presented by Frank [4]. Using simplified models for train movements he calculates the capacity of a single-track line both for one-way traffic and for certain (fleet) systems of two-way traffic. The results are most applicable on freight or military transport systems but may also serve as a starting point for further studies.

Chen and Harker [3] present a sophisticated model for estimation of mean delays and delay variance for trains that operate on a single-track. In this model the inter-station distances are assumed to be even and the actual departure time of each train is randomised around a specified timetable time. The conflict resolution is handled through calculation of probabilities of conflict between every pair of trains. The study shows that shorter inter-station distances lead to lower mean delays and delay variances. The number of trains also influences the delays significantly.

### 2.2 Knock-on delays

Some modern research has also been carried out with an explicit focus on knock-on delays, or secondary delays as they are referred to in this thesis. Although these studies mostly concern stations and/or double-track lines, they clearly show the effects of delay propagation. Several studies in this area are timetable-independent, which makes them very useful in long-term planning and in other situations where the timetable is not known.

Carey [2] takes up different measures of reliability. He discusses the advantages and disadvantages of measures based on probabilities (i.e. observed delays) and measures not using probabilities. Almost all measures for prediction of reliability involve headways (time space between two consecutive trains) since longer headways generally reduce knock-on delays. Apparently there are several advantages of using simple measures that are not based on probabilities, although mathematical methods for more exact calculations are available.
In [8] Oetting proposes a model for calculating knock-on delays that appear in a system of serial infrastructure elements. This type of chaining is based on the fact that outgoing delay in one segment is equal to ingoing delay in the following section and that disturbances tend to spread in the opposite direction. This type of calculation technique, based on convolution, may be used for studies of infrastructure elements as well as studies of train-paths (timetables).

In [7] Huisman et al use queuing theory to analyze the dependencies and interactions between the individual components in a railway system. The operation is here defined by frequencies of service and no specific timetable is defined. Stations, junctions and sections (lines) are modelled according to their special properties. The model seems to be a good alternative to simulation and the result is, in some sense, mean values of all possible timetables that can be constructed from the frequencies that are given as input data. In the model occupation times as well as minimal headway times are assumed to follow negative exponential distributions. Therefore the calculated waiting times are slightly overestimated.

Other examples of literature concerning knock-on delays are given by Wendler [15], Carey [1] and Yuan et al [16].

2.3 Rescheduling and delay management

While the research on knock-on delays concerns small delays that occur with high frequency, the focus in rescheduling is on large delays and disturbances. In a rescheduling problem the aim is to restore the traffic to the timetable in such a way that the knock-on delays are minimised. The infrastructure is then a given constant as is the planned timetable. The literature on rescheduling is extensive and only a few examples are mentioned here.

In [13] and [14] Törnquist presents an optimisation approach to the rescheduling problem. A mathematical formula which allows an n-tracked network to be modelled is constructed. Alternative objective functions, such as total final delay and total cost associated with delays are used and four different rescheduling strategies are tested. The study shows that it is possible to find rescheduling solutions that limit the knock-on delays and/or cost associated with these delays. The most complete rescheduling strategy is sometimes too time-consuming for practical use. However, a more limited optimization strategy is often good enough.

A special field within the rescheduling research concerns connections between trains (most often passenger trains). One example is Schutter [12] who examines the possibilities to recover from delays by breaking connections. In the presented model the connections are represented by different kinds of synchronization constraints. In case of delay the so-called soft constraints may be broken, but at a cost that represents compensation activities and dissatisfaction for passengers. The method is feasible for real-time dispatching since the system uses a moving horizon in which the model is continuously updated.
2.4 Concluding remarks

Infrastructure configuration, timetable and delays are essential ingredients in railway operation. Modelling the operation of a railway system therefore requires assumptions of these parameters. Depending on these assumptions the analysis will fit into one (or more) of the boxes in the following matrix. Using this type of matrix it is possible to group the models presented in literature.

<table>
<thead>
<tr>
<th>Infrastructure</th>
<th>Constant (existing)</th>
<th>Variable (no constraints)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant (existing)</td>
<td>Reliability, rescheduling and delay management</td>
<td>Do not apply</td>
</tr>
<tr>
<td>Known demand (frequencies)</td>
<td>Timetable optimisation (knock-on delays)</td>
<td>Combined infrastructure-and timetable optimisation</td>
</tr>
<tr>
<td>Variable (no constraints)</td>
<td>Examination of infrastructure potential</td>
<td>General examination</td>
</tr>
</tbody>
</table>

When both timetable and infrastructure are constant (upper left) the main variable will be delays (punctuality) and different kinds of rescheduling and reliability models apply.

In cases where the timetable is a variable that is guided by some sort of constraints (e.g. demand requirements) we turn into the field of optimization, either of the timetable layout only, or of a combination of infrastructure and timetable. Knock-on delays play an important role in objective functions in this type of optimization.

If the timetable is completely unknown, and the aim is a general examination, rather than optimization, we are out for an investigation of infrastructure potential or a complete general examination. This situation sometimes occurs in long-term planning when both the timetable and the future development of demand are unknown.

The idea of this thesis is to provide some general knowledge about operative properties of single-track lines. Infrastructure and timetable as well as delays thus appear as variables and the examination becomes rather descriptive. To achieve this it is necessary to disregard congested situations. Congestion appears for special combinations of infrastructure, timetable and delay patterns and is, as the literature already states, of special interest. However, congestion is an undesirable and troublesome state that appears when systems are overloaded. In order to understand the effects of congestion it is necessary to start with more simple non-congested situations.

Even if the infrastructure and timetable are treated as variables they have to be limited in some way. A single-track implies several constraints on the timetable. Therefore the proposed “timetable-free” analysis is more natural on single-track systems that have inherent limitations.
3 Methodology

Railways are complex technical systems. They are often considered to be static, stiff and inflexible. As long as only constant, non-dynamic parts, such as the infrastructure, are considered, the system is quite easy to understand. Reality is different. The variance in different parameters makes the system difficult. Some examples:

- **The timetable** creates a well defined structure. However, the capacity is utilised differently every day since the actual timetable *varies* from one day to the other due to delays, extra trains and cancelled trains.

- **The available capacity**, which is a very important condition for the timetable, varies over time. Failures, construction work, accidents and delays all make the available capacity *vary* over time.

- **Vehicle properties**, relative to those assumed during timetable construction, *vary*. Important examples are freight trains whose train mass often differs from the timetabled train mass, change of vehicle types without corresponding change of the timetable, partial vehicle failures, weather conditions that affect adhesion etc.

- **The railway system is used and operated by humans.** Human behaviour *varies* naturally from one time to another. Train crews, dispatchers and passengers all contribute to this variance.

All these variances make the railway system complex and very interesting to analyse. To find general relations in this noise of superposed variances is not easy. Moreover, strong interactions between factors can be expected, that make the task of analysis even more difficult.

There is no obvious choice of method of analysis. Two alternative methods appear:

1. Analysis of a specific situation in detail.
2. General analysis given simplifying assumptions and delimitations.

The first method results in deep knowledge of a specific traffic situation. By delimiting the examination to a specific situation it is possible to use a commercial railway operation simulation tool. This type of simulation requires detailed information about infrastructure, timetable, dispatching and disturbances, which in fact implies a great many hidden assumptions.

A simulation can be compared to a spot test. Most parameters have to be set by hand since the available simulation programs for railway operation do not handle automatic parameter variation. Therefore, only a few variants may be analysed, which in practice requires some kind of simplifying assumptions!
The program codes of the available simulation programs are not open. Therefore, it is in many cases difficult to interpret the results when it is not obvious how the programs work.

The second method means that only quite simple traffic situations are analysed. Given simplifying assumptions, mathematical models become possible. These may be performed to handle several variables simultaneously.

This thesis aims to capture fundamental knowledge that is general rather than specific. For this reason it is natural to start with simple cases, seek simple relations and learn about important principles. After that it is possible to increase the complexity and perform detailed studies.

Simplifying assumptions and delimitations are therefore accepted at this basic level and a mathematical model that is adjusted to the actual issues is a natural method. Compared to simulation a more analytical model has several advantages:

- The researcher controls all assumptions and parameters. Modelling of dispatching and vehicle movements are two important examples.
- The model may be adjusted in detail so that a systematic examination of important variables is possible. Infrastructure configuration and punctuality are examples.
- The model may be constructed without specific assumptions about timetables, which means that the infrastructure can be more unbiased examined.
- The model may be adjusted for calculation of new concept variables. One important example is timetable flexibility.
- The results become easier to grasp. This is particularly true if independence assumptions are used.

A simple mathematical model is well suited for systematic analyses of a great many simple fictive cases. A central part of this project has therefore been the construction of a mathematical model that gives general results for simple operative cases.

### 3.1 Modelling single-track operation

The operation of single-track lines is well suited for analytical modelling since the traffic is well-defined by the crossing stations. The crossing situations, where trains travelling in opposite directions meet, are probably the most important part of single-track operation since they cause time losses and delay propagation. These time losses are hereafter referred to as crossing time.

In order to examine the influence of crossings a model, named SAMFOST, has been constructed. The model stands on two fundamental assumptions:
1. Two crossing trains are independent before crossing.

2. Different crossing situations are independent of each other. This implies that following trains are always far enough apart not to interfere with each other when they cross a train in the opposite direction.

These assumptions have high validity in non-congested situations, i.e. when the combination of infrastructure configuration, timetabled frequency of services and punctuality is such that following trains only occasionally come so close that independence is broken.

SAMFOST performs a stepwise analysis where each step means that a new set of parameters have to be assigned:

Step 1: The combination of infrastructure and vehicle data together with passenger stop data, give a timetable-free characteristic of the line.

Step 2: The addition of delay data and timetabled crossing point gives the full distribution of the crossing time.

The most important advantage of the two fundamental assumptions is that the first model step becomes independent of further assumptions about timetable and punctuality. This makes it possible to define the so-called crossing time function, which is a timetable- and punctuality-free description of the infrastructure properties. The crossing time function shows the time needed to perform a crossing as a function of the theoretical crossing point that would be realised on a double-tracked line.

In the second step the punctuality is taken into account and hereby the model clearly shows how punctuality influences the operative result. Using this stepwise modelling
a deeper understanding of the combination of infrastructure design, timetable and punctuality becomes possible.

In order to check the validity of the model a crossing time function was derived from data from the simulation of a reference case. The differences between SAMFOST and the simulated results (RailSys) turned out to be very small.

After choosing mathematical modelling as method of analysis some questions about the methodology remains, i.e. how the developed model shall be used in order to answer the specific questions that is part of the project objectives.

Unlike most previous research, the infrastructure is here treated as a variable. In order to let the infrastructure vary, the exact configuration of existing railway lines has to be abandoned. A natural way to do this is to examine fictive railway lines with extensive symmetries.

In these infrastructure designs important parameters such as station design and inter-station distances follow defined standards. The symmetries make the results easier to interpret. Within each infrastructure variant the inter-station distances are equal, all stations are identical, the same vehicle and punctuality are used for both directions etc.

To complete the examination of the three specific objective measures a factorial experiment is performed. Varying more variables, in this case six, simultaneously helps to increase the generality of the results. Moreover such an experiment reveals hidden interaction effects that connect the variables to each other.
4 Results

4.1 Influences of station length (partial double-track) and inter-station distance on delays and delay propagation on single-track lines with regional rail traffic

This paper shows the advantages of simple analytical models. The crossing time function is introduced and merely by looking at the function the characteristics of a single-track line become clear.

The crossing time is defined as the extra time needed to perform a crossing on a single-track compared to a double-track where crossings do not imply any extra time consumption. The crossing time function shows how the crossing time varies along the line with local minima at the locations of crossing stations and local maxima in-between them. Figure 4-1 shows an example.

Figure 4-1 Crossing time function for a line with nine, equally spaced, crossing stations.

Without knowing anything about the timetable the crossing time function tells us how the crossing possibilities vary along the studied line. Figure 4-1 shows the crossing time function for a completely symmetric line with 9 identical crossing stations (denoted B1, B2, F1-6 and C), equal inter-station distances and no passenger stops.

Each crossing station results in a time interval with low crossing time. This is the lowest achievable crossing time given the actual combination of infrastructure design and vehicle performance. Compared to a double-track, that has a constant zero crossing time function, the increase in travel time, that has to be paid for a crossing, is rather high.
The crossing time function can be used to estimate the time supplement that has to be added in the timetable for different locations of the timetable crossing point. However, this estimation will only be correct as long as the trains arrive with a zero delay difference, i.e. when they arrive in the planned time relation to each other.

As soon as the delay difference is non-zero the realised crossing time will take another, often higher, value than the planned crossing time. The actual crossing time will thus consist of a deterministic part that is given by infrastructure and vehicle parameters and a stochastic part that depends on the actual delay difference as well as infrastructure and vehicle factors. The latter part varies with the delay difference and may be referred to as delay propagation.

The second part of the paper deals with the effect of partial double-tracks and decreased inter-station distances, respectively. As shown in figure 4-1 some conclusions may be drawn without knowledge about the location of the timetable crossing point or the punctuality. By introducing a distribution for the delay difference the complete crossing situation for a given timetable crossing point may be illustrated and examined. Figure 4-2 shows an example.

![Figure 4-2 Crossing time function for different inter-station distances and probability density function for delay difference (dashed).](image)

Figure 4-2 shows an example of a probability density function and crossing time functions for three different inter-station distances: 15 km (bold solid), 9 km (thin dashed) and 3 km (thin solid). The combination of a crossing time function and a probability density function makes the situation clear and intuitive. As indicated, a decrease in inter-station distance results in a higher frequency and lower maximum crossing time. The minimum crossing time is not affected since all stations have exactly the same features!
Combining different inter-station distances with three levels of arrival punctuality gives the following results:

- Shorter inter-station distances have a general effect. The effect on mean crossing times is small but rather independent of the arrival punctuality. This is also true for the crossing time variance. Altogether this means that shorter inter-station distance is a suitable measure when robustness is more important than travel times.

Please note that this feature is a consequence of the assumption that all stations are identical in design. If a passenger stop is introduced at one or some of the stations this symmetry is lost and the system properties are dramatically changed. This is treated in the following papers.

The other measure that is examined in the paper is partial double-tracks (increased station length).

![Figure 4-3 Crossing time function for a line with partial double-track and probability density function for delay difference (dashed).](image)

In the example shown in figure 4-3 the crossing time function corresponds to a line with a 13 km partial double-track and four ordinary crossing stations on each side. The location of the probability density function shows the position of the timetable crossing point, i.e. the point having the highest probability density.

When an ordinary crossing station is replaced by a partial double-track, the lowest crossing time decreases. The resulting lowest level only depends on the speed restriction at the entrance and exit points. Within this time interval no signal interference between the crossing trains occurs, which is always the case in the
corresponding situation on an ordinary crossing station. An extension of the double-track also results in a wider time interval having this lowest crossing time.

The effect of a partial double-track is thus two-fold: the crossing time is locally reduced and this reduction has a greater extension compared to an ordinary crossing station. Both features are important when the line is operated and the actual delay difference varies stochastically according to some distribution.

Combining the two functions it is possible to calculate the crossing time distribution for a given timetable crossing point. For different infrastructure designs, giving different crossing time functions, and/or different punctuality levels, resulting in different density functions, it is possible to compare the effect of different parameters.

In the paper different lengths of the partial double-track are combined with three punctuality levels and the result may be summarised thus:

- Partial double-tracks have only a local effect on the crossing time function. Therefore the effect of a partial double-track is highly dependent on the arrival punctuality of the trains. Low punctuality means that the mean crossing time does not decrease as much as for higher punctuality. Even more important, however, is the fact that the crossing time variance is substantially higher when arrival punctuality is low.

- The speed restriction at entrance and exit points delimits the effect of partial double-tracks. A less restrictive speed would result in higher crossing time variance but the decrease in mean crossing time would be significant.

Some of these features change when a passenger stop is introduced on the partial double-track. This is treated in the following papers.

4.2 SAMFOST – a Timetable-free Way of Analysing Single-track Railway Lines

This paper deals with the conditions for independence that are assumed in the SAMFOST model, validation of the model against the simulation tool RailSys and some important applications such as combined crossing and passenger stop. Finally, an existing Swedish railway line is used to exemplify the usefulness of the model.

Two important independence assumptions open for a stepwise systematic approach like SAMFOST. The crossing time is determined by a great many factors such as infrastructure, vehicle, timetable crossing point, driver behaviour, delays, surrounding trains etc. In SAMFOST the infrastructure, vehicles and driver behaviour are regarded as deterministic, the delays are regarded as stochastic and surrounding trains are excluded from the model.

Disregarding trains other than the two crossing ones and assuming deterministic behaviour of other factors means that the crossing time only depends on the arrival delay difference of the two crossing trains. This makes it possible to calculate a crossing time function for each combination of infrastructure, vehicle and pattern of
passenger stops. In a case with several interacting trains a given delay difference would result in different crossing times depending on the positions and delays of surrounding trains.

It is therefore only possible to disregard congestion effects as long as situations where more than two trains interfere are rare. How often congested situations occur depends on infrastructure design, vehicle parameters and punctuality. For a given combination of infrastructure, vehicle and stopping pattern, it is therefore possible to calculate a congestion-free capacity for different levels of punctuality. In this thesis a situation where interference between following trains occurs less often than 1/200 is regarded as congestion-free. This level is somewhat conservative and so all types of congestion effects may be ignored. In practical use a higher value may be accepted.

Figure 4-4 shows how the congestion-free capacity falls when punctuality deteriorates and that the combination of infrastructure design, vehicle parameters and passenger stop pattern is of great importance. The upper curve represents a situation with short inter-station distances, high acceleration vehicles and no passenger stop while the lower curve represents longer inter-station distances, lower acceleration vehicles and passenger stop at the timetabled crossing station. In all cases congestion is accepted in 0.5 % of the crossing situations. SAMFOST can thus be applied with high validity for all points under each curve.

This diagram of congestion-free capacity is an important result in itself since it gives an idea of the correlation between capacity and punctuality. When a line is operated under non-congestion conditions the traffic is likely to be stable and controllable.
Non-congested operation should be an aim in timetabling in order to avoid severe punctuality problems in the operation.

Given the two assumptions of independence the SAMFOST model has been validated through comparison with the simulation tool RailSys [11]. Random samples of deceleration courses, acceleration courses and combinations of both were taken and run in both models. Only small differences, less than one second per run, occurred.

The dispatching and signalling functions were also validated. Two important dispatching situations occur on single-track lines: choice of crossing station and track disposition at the chosen station. When a crossing station has been chosen and track disposition decided, signalling functions are crucial for the trains’ interaction during the crossing course.

The validation shows that SAMFOST chooses crossing station properly. Unfortunately, it was not possible to validate the track disposition since RailSys does not model this in a proper way. The actual crossing course, including the important signalling functions, was validated with good results. In a large sample of delay differences (runs), the difference in crossing time between the models never exceeded 1.5 seconds. A similar validation was performed to test the effects of a combined crossing and passenger stop, which is a more complicated dispatching situation. This validation also turned out well.

Altogether the validation shows that SAMFOST models crossing situations very similarly to RailSys, as long as the fundamental assumptions about independence hold.

Having validated the model and analyzed its delimitations it can be used for real analyses. One situation of special importance is the combination of crossing and passenger stop.

If a crossing is planned at a station where the trains have a regular passenger stop the time for deceleration and acceleration, as well as most of the waiting time, is useful time that does not become crossing time. Therefore, the combination of crossing and passenger stop implies time efficient crossings. This can be seen in the crossing time function in figure 4-5.
Figure 4-5 Crossing time function for a line with 9 crossing stations (solid) and passenger stop at mid-station. Different locations of timetable crossing point imply different positions of the probability density function for the delay difference (dashed).

Most often at Swedish stations the platforms are located at one end of the station which results in a slightly asymmetric crossing time function. Referring to the crossing time function (solid line) in figure 4-5, the effects of a passenger stop may be summarised as follows.

- Time efficient crossings are made possible due to the stop. A time interval occurs, whose length depends on the timetabled dwell time. Speed restrictions at the farther end of the station imply a small crossing time.

- The time interval within which the stop station is used becomes wider than would be the case without a stop.

- The maxima surrounding the stop station become higher due to accelerations and decelerations that increase the run time on the single-track sections surrounding the stop station.

A passenger stop means that the amplitude of the crossing time function increases significantly. This means that the variance and sensitivity of punctuality also increase compared to a line without any passenger stop.
Figure 4-6 shows that the mean crossing time (solid curves) is highly dependent on punctuality. The uppermost curve corresponds to low punctuality, Exp(350s)-distributed arrivals, whereas the lowermost curve corresponds to higher punctuality, Exp(150s). This dependence is a result of the described asymmetry that implies that the timetable crossing station cannot be replaced by any other station without greatly increasing the crossing time.

Although the level of the crossing time is significantly lower compared to a case without any passenger stop, the punctuality dependence implies that it is difficult to find a feasible timetable since the amount of supplements is correlated to punctuality.

The standard deviation (dashed curves) shows only a slightly stronger dependence on punctuality compared to the case without any passenger stop. However, the level of the deviation is much higher, making it more difficult to construct a robust timetable.

The analyses described so far assume that the timetable crossing point is located symmetrically on a crossing station. Different kinds of timetable dependencies often make it impossible to obtain the optimal timetable crossing point for every crossing. It is therefore of interest to examine different locations of the timetable crossing point. This means that the probability density function is located at different locations along the crossing time function, as exemplified in figure 4-5.

The timetable crossing point is here defined as the theoretical point where the crossing would take place at arrival delay difference equal to zero if the line was double-track. Assuming high punctuality, i.e. Exp(150s)-distributed arrival delays in both directions, and letting the density function shift stepwise along the crossing time function, the mean crossing time and its standard deviation for each location...
result in figure 4-7. This figure shows results for seven different inter-station distances, i.e. seven different crossing time functions.

Figure 4-7 Mean crossing time (solid) and standard deviation (dashed) as a function of the location of timetable crossing point for different inter-station distances.

Figure 4-7 clearly shows the differences between infrastructure designs (curves) as well as differences between locations (horizontal axis). Uppermost curves represent 15 km inter-station distance while lowermost curves represent 3 km inter-station distance (step length 2 km/curve). The shifting inter-station distance is reflected in different frequency and different crossing time levels.

Every minimum corresponds to a crossing station. At the middle crossing station all trains have a passenger stop which results in an absolute minimum for the mean crossing time.

Looking at the standard deviation it becomes clear that a shorter inter-station distance gives lower crossing time variance while a passenger stop results in higher crossing time variance.

This type of mean crossing time functions and standard deviation curves is also an important methodological result of this research since the method displays the impact of punctuality on crossing time and timetable construction.

SAMFOST has also been used to analyse an existing single-track railway line, the Svealand line from Södertälje to Eskilstuna southwest of Stockholm in central Sweden. This analysis shows the effect of real asymmetries with unequal inter-station distances and different punctuality for different traffic directions.
Figure 4-8 Crossing time function (solid thin). Mean (solid bold) and standard deviation (dashed) for crossing time. Two different levels of arrival punctuality are shown.

Figure 4-8 shows the characteristics of the Svealand line. The solid thin line is the crossing time function revealing four crossing stations, of which one is combined with a passenger stop (Nkv) and three are not, one partial double-track and two flanking double-tracks.

The figure also shows mean crossing time functions and standard deviation curves for two different levels of punctuality: existing punctuality and increased punctuality. The existing punctuality results in a mean function that is quite insensitive to changes in the crossing time function, which results in a rather low amplitude and small differences between alternative timetable crossing points.

If punctuality is improved the mean crossing time function follows the crossing time function much better and the amplitude is much higher. This follows the fact that the mean time function converges towards the crossing time function when punctuality increases. When punctuality is absolute and no delays occur, the two functions coincide.

Another important result of this analysis is that the timetable crossing points on double-tracks near the border point to a single-track need special attention. This is the fact to the left of Eskilstuna (Et) and to the right of Södertälje (Söö) where double-track sections help to decrease, but not completely eliminate, the crossing time. The crossing time variance in particular is high on these sections.
4.3 Crossing Times on Single-track Railway Lines – Dependencies of Different Infrastructure and Traffic Factors

In the third paper the concept of timetable flexibility is defined and exemplified. SAMFOST is then used to examine the impact of six different infrastructure and traffic factors on the crossing time at the optimal timetable crossing point as well as on the timetable flexibility.

The mean crossing time function and standard deviation curves that were derived in the second paper may be used to define timetable flexibility. When a new railway line is projected or a new timetable is to be constructed it is of great importance to assemble knowledge about potential crossing points’ features. One way of doing this is to calculate the mean crossing time function and standard deviation curve and use these to classify and compare alternative timetable crossing points.

The paper presents a first attempt to choose accepted timetable crossing points out of all possible points along the studied line. Two requirements for timetable crossing points are natural: as low a mean crossing time as possible and as low a crossing time variance as possible. In this first, and very simple, test only the mean crossing time function is used.

Figure 4-9 Mean crossing time function (solid) and standard deviation curve (dashed) for a line with five crossing stations. A tolerance line (horizontal at approx. 200 s) marks the accepted timetable crossing points located below it.

A tolerance level is defined 5 seconds above the highest located (local) minimum of the mean crossing time function. All timetable crossing points that have a lower mean crossing time than this tolerance level are regarded as accepted timetable crossing points. Other kinds of choice criteria, also including the standard deviation
might be a good alternative, but for this first study only the simplest rule has been applied.

Figure 4-9 shows an example. The mean crossing time function (solid) originates from a line with five crossing stations. Trains stop at the mid-station which makes the mean function fall whereas the standard deviation curve rises. Following the simple rule every timetable crossing point that has a mean crossing time that is lower than the horizontal tolerance line is regarded as accepted.

The choice rule works very well as long as the mean crossing time function does not vary irregularly. From practical timetabling it is known that all crossing stations are accepted timetable crossing points. It is therefore reasonable to accept all local minima, although a more restrictive rule would only accept points near to the absolute minimum.

This reasoning gives that the simple rule of choice may be applied in sections “A” (outside outermost dashed vertical lines) and “C” in figure 4-9, whereas sections marked “B” still remain to be decided on. This could be done by also introducing the standard deviation. However, to keep this first attempt to define timetable flexibility as simple as possible, the simple rule has also been applied in “B-sections”. This will probably overestimate the effect of a passenger stop or other properties that decrease the crossing time locally.

All accepted timetable crossing points contribute to the timetable flexibility and the group of accepted points may be subject to different kinds of evaluation to describe the flexibility. When doing this analysis, it is very important to keep in mind that punctuality affects the mean crossing time function. The paper briefly discusses possibilities to make the effects of punctuality strong or weak.

In the paper three different measures for timetable flexibility are defined. All of them are quite natural when a mean crossing time function is studied:

- Share of accepted timetable crossing points.
- Spread in position (horizontal axis), i.e. position variance for accepted timetable crossing points.
- Spread in crossing time (vertical axis), i.e. crossing time variance for accepted timetable crossing points.

Together these measures tell us a great deal about the possibility to find different timetable solutions with accepted crossing properties. The first two measures are fundamental whereas the third is a sort of check for the imperfect rule of choice.

In the second part of the paper a factorial experiment is presented. A model like SAMFOST is useful for multivariate analyses since the combined effect of several variables may be examined systematically, thereby making it possible to find and examine the effect of not only single variables, but also groups of variables that cause different kinds of interaction effects.
One simple way to perform a multivariate analysis is a factorial experiment at two levels. In such an experiment a number of factors are chosen. Each factor may take one of two values, denoted high and low level. The “experiment” is then run for all combinations, $2^n$ in total, and for each run the values of the response variables of interest are saved. By evaluating the saved values, conclusions may be drawn about both main and interaction effects.

Six factors were chosen, forming a $2^6$-experiment, and levels were assigned as follows.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Description</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Track length at station / on double-track where timetable crossing point is located</td>
<td>0.75 km 10 km</td>
</tr>
<tr>
<td>B</td>
<td>Inter-station distance</td>
<td>15 km 7.5 km</td>
</tr>
<tr>
<td>C</td>
<td>Passenger stop</td>
<td>No stop 60 s</td>
</tr>
<tr>
<td>D</td>
<td>Speed restriction at points at station where timetable crossing point is located</td>
<td>100 km/h 160 km/h</td>
</tr>
<tr>
<td>E</td>
<td>Vehicle type</td>
<td>X2 X50</td>
</tr>
<tr>
<td>F</td>
<td>Arrival punctuality (mean arrival delay)</td>
<td>200 s 100 s</td>
</tr>
</tbody>
</table>

Four factors, A, B, C and F have been examined earlier in different ways. To widen the examination two more factors were added. According to previous results speed restriction at end points on partial double-track (D) might be a variable of interest. To test the effects of lower acceleration the vehicle type (E) was also varied.

The levels were chosen to be reasonable and then the experiment was run with respect to five different response variables:

- **Centred timetable**: mean and standard deviation for crossing time.
- **Timetable flexibility**: simple flexibility, position spread and crossing time spread for accepted timetable crossing points.

After 64 runs the six main effects and 58 interaction effects were estimated for each response variable.

The results for a centred timetable, i.e. timetable crossing point fixed to the mid-station (or partial double-track), are shown in figure 4-10 and 4-11. Each main effect is here defined as the difference between the mean value over experiments with the factor on its high level (32 experiments) and the mean value over experiments with the factor on its low level (32 experiments).

In a similar way each two-way interaction effect is calculated as a difference between the effect of changing the first factor from low to high level when the second factor
is low and the effect of changing the first factor from low to high level when the second factor is high.

In the figures all effects are shown by absolute value and each bar represents the effect of a change from low to high level. The sign above each bar indicates whether the effect decreases (−) or increases (+) the response variable.

Figure 4.10 shows that passenger stop (C) and partial double-track (A) are the most important factors for the mean crossing time. However, a strong interaction between these factors (AC) tells us that the effects are not additive. The lowest main effect is that of inter-station distance (B) that affects the mean crossing time least.

An important interaction occurs between partial double-track (A) and speed restriction at points (D). In this case the AD-interaction is almost as high as the D-effect itself. This is explained by the fact that a high speed at the points is only useful when combined with a partial double-track. With an ordinary crossing station, having a track length of only 750 m, higher speed than 100 km/h gives only a minor effect.
The results from these factorial experiments may also be presented in tables including the signs of each effect. The following tables, omitted in the paper but shown here in order to clarify the results, show the six main effects and the 15 two-way interaction effects. Please note that all effects are related to the overall mean value presented in the first column denoted “µ”. This mean value corresponds to a mean level of all factors. Due to linearity assumptions of the model all tabled effects take half the value compared to figures 4-9 and 4-10 showing the entire difference between low and high level.

<table>
<thead>
<tr>
<th></th>
<th>µ</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean [s]</td>
<td>103</td>
<td>-30</td>
<td>-7</td>
<td>-44</td>
<td>-11</td>
<td>-11</td>
<td>-14</td>
</tr>
<tr>
<td>Std [s]</td>
<td>53</td>
<td>-8</td>
<td>-11</td>
<td>8</td>
<td>5</td>
<td>-3</td>
<td>-11</td>
</tr>
</tbody>
</table>

Knowing the level of each factor the expected mean crossing time can be calculated as a sum according to the following equation. Please note that the sign of each factor must be included. Each parenthesis shall therefore be replaced by either +1 or -1.

\[
\theta_{\text{kin}} = \mu(\pm) A(\pm) B(\pm) C(\pm) D(\pm) E(\pm) F
\]

The indices \(i, j, k, l, m, n\) denote the level signs of the six factors, i.e. + or -, \(\mu\) denotes the overall mean value taken over all 64 runs and capital letters A-F, AB-EF denote estimations of effects that can be read in the tables above. In order to emphasise that all values are estimations, each effect letter is marked with a horizontal line on the top. The following example shows how the formula works.

Assume that we have a line with a partial double-track, an inter-station distance of 15 km, a passenger stop, 160 km/h at the points of the partial double-track, X2-vehicles and punctuality at 200 s. This means that the signs will be \(+ - + + - -\). If only main and two-way interaction effects that have an absolute effect \(\geq 5\) seconds are considered, the resulting mean crossing time can be estimated as:

\[
\theta_{-+++} = \mu + A - B + C - D - E - F - AB + AC + AD - CE =
\]

\[
= 103 - 30 + 7 - 44 - 11 + 11 + 14 - 5 + 19 - 10 - 8 = 46 \text{ s}
\]

This example clearly shows how a partial double-track and a passenger stop decrease the mean crossing time, whereas a longer inter-station distance, a vehicle with weak acceleration, and low punctuality increase the mean crossing time. Also note how the strong AC-interaction reduces the effect, “increasing” the crossing time by 19 seconds since A- and C effects are not additive.
If the standard deviation is analysed with the same method a somewhat different result appears, see figure 4-11. First, the inter-station distance (B) turns out to be the most important single factor which means that a short inter-station distance results in a low crossing time variance. However, the combination of passenger stop (C) and partial double-track (A) is very important for the variance. When a passenger stop is introduced at an ordinary crossing station it results in a substantial increase in crossing time variance. This increase can be eliminated if the passenger stop is combined with a partial double-track. This fact is reflected in a strong interaction between these factors.

Figures 4-12 and 4-13 show how the six factors affect timetable flexibility. All effects are shown by absolute value and each bar represents the effect of a change from low to high level. The sign above each bar indicates whether the effect decreases (−) or increases (+) the response variable.

**Figure 4-12** Main (dark bars) and interaction (light bars) effects on simple flexibility.

**Figure 4-13** Main (dark bars) and interaction (light bars) effects on position spread of accepted timetable crossing points.
The inter-station distance (B) is the factor that has the greatest impact on the simple flexibility, i.e. the share of the line that has accepted mean crossing time. This is natural since a decrease in inter-station distance is a general measure that works along the whole line and alters the mean crossing time function entirely.

Also, a partial double-track increases the simple flexibility, but in this case the new timetable crossing points make a time interval along the double-track. Please observe the strong interaction effect AB that means that these effects are not additive.

Looking at the second ratio of timetable flexibility, it is clear that a shorter inter-station distance increases the spread of timetable crossing points whereas the effect of a partial double-track is very low.

It is remarkable that punctuality in fact greatly affects timetable flexibility. Given the definitions of the two flexibility ratios a higher punctuality actually means less timetable flexibility, which means that it becomes more difficult to find a feasible timetable when punctuality is high.

Also the flexibility ratios may be calculated with the sign formula presented above using the tabled results below.

<table>
<thead>
<tr>
<th></th>
<th>μ</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple flex. [s]</td>
<td>79%</td>
<td>11%</td>
<td>17%</td>
<td>8%</td>
<td>0%</td>
<td>-3%</td>
<td>-6%</td>
</tr>
<tr>
<td>Position std [s]</td>
<td>477</td>
<td>-5</td>
<td>38</td>
<td>12</td>
<td>-1</td>
<td>-8</td>
<td>-13</td>
</tr>
<tr>
<td></td>
<td>AB</td>
<td>AC</td>
<td>AD</td>
<td>AE</td>
<td>AF</td>
<td>BC</td>
<td>BD</td>
</tr>
<tr>
<td>Simple flex. [s]</td>
<td>-9%</td>
<td>-3%</td>
<td>0%</td>
<td>1%</td>
<td>2%</td>
<td>-6%</td>
<td>0%</td>
</tr>
<tr>
<td>Position std [s]</td>
<td>2</td>
<td>13</td>
<td>1</td>
<td>0</td>
<td>-4</td>
<td>-18</td>
<td>0</td>
</tr>
</tbody>
</table>

The factorial experiment shows that several combinations must be examined if some sort of “general” view is to be obtained. For many combinations the results from the experiment cases confirm previous conclusions and results. One of the clearest conclusions seems to be that shorter inter-station distances are a simple measure that acts on the crossing time variance and timetable flexibility, but not on the mean crossing time. A partial double-track, on the other hand, is a much more complex measure (several interaction effects indicate this) whose effects are more difficult to explain. Also, a passenger stop turned out to be a complex measure with a strong negative effect on the increase in crossing time variance. Irrespective of this, the important decrease in mean crossing time is a valuable property, in particular when punctuality is high.
5 Discussion of the main contributions of the thesis

The contributions of this thesis may be divided into five areas. A simplified method of analysis has been introduced that gives clear, well-structured results, the importance of punctuality is emphasised, and the infrastructure is treated as a variable which makes it possible to gain knowledge about alternative configurations. All analyses are, in principle, free from timetable assumptions and a concept of timetable flexibility is presented. The proposed model can be useful in infrastructure planning as well as timetable construction.

5.1 Simple models

Despite a well-defined infrastructure configuration and timetable, railway operation is complex. Different kinds of variances contribute to this complexity. Most of the methods of analysis at hand require assumptions about these variances.

An important contribution of this work is that a few fundamental assumptions open for simple models. These models are simple to use and the results are understandable and easy to display. By using these models the results become more general and systematic.

Its ease of use, together with the systematic analysis the model allows, is an important advantage in infrastructure planning where the results may be used to design more detailed studies with other techniques. In timetabling the simple models may contribute when standards and rules-of-thumb are to be decided.

It could be argued that the fundamental assumptions make the model too rough. This is probably true in cases where good predictions of future demand and operation exist. However, this is in fact seldom the case. Instead, reality shows that assumptions used to make predictions during the construction of railway systems often differ from the actual traffic that is operated when the line is in use. It is often of great importance to know somewhat more about the system properties.

This fact indicates that a more general knowledge of the operation conditions is valuable. By using less detailed overview perspective it is possible to study how the system would react if operation conditions change.

5.2 Punctuality dependencies

Almost the whole Swedish railway network suffers from low punctuality. On several lines the punctuality problems delimit the capacity since the number of secondary delays accepted is limited. Still, punctuality is usually used only as a follow-up measure. Although the secondary delays predominate, punctuality is seldom systematically used as feed-back in timetable construction. Even less often is the infrastructure design evaluated with respect to actual punctuality.

In this work punctuality is one of the most important parameters. An important contribution is that the impact of punctuality is shown explicitly. Together, the crossing time function and the mean crossing time function clearly show that
punctuality matters. The crossing time function shows the situation with full punctuality where the effect of infrastructure configuration is clearly seen, while the mean crossing time function, for a low punctuality case, shows a situation that is apparently non-sensitive to the infrastructure configuration.

So the effects of infrastructure measures are in fact highly dependent on punctuality. This again shows the need for a systematic variation of important parameters when railway operation is analysed.

The work also shows examples of the impact of punctuality on capacity. The calculation of congestion-free capacity, where punctuality is combined with infrastructure parameters, clearly states that capacity is extremely dependent on punctuality, as long as the occurrence of congested situations has to be limited. For the Swedish railway this is an important contribution since lack of systematic feedback and knowledge of punctuality-capacity relations mean that several lines are over-utilised compared to the required punctuality levels.

### 5.3 Infrastructure as a variable

In most other studies the infrastructure is treated as a constant, or a variable that can take on only a few predefined values. This thesis shows examples of how the infrastructure, in a very systematic way, can be treated as a variable. This means that fictive model lines are examined, instead of existing infrastructure designs that are highly asymmetric and unique.

This is an important contribution in itself, since almost all other studies focus on a given infrastructure. In order to learn more about railway operation one has to allow the infrastructure to vary within a rather wide area. In some sense this fact means that a new dimension of the analysis is accepted.

One result of this is that not only one length of partial double-track or inter-station distance is examined. It is therefore possible to draw conclusions about marginal effects of double-track extensions and station distance shortenings. It is also possible to make a clear definition of partial double-track, display patterns that result from station properties that are repeated at every station etc.

The general knowledge that is gained by analysing symmetric cases is very useful when existing lines are to be improved. Also in this case it is useful to treat the infrastructure as a variable that is changed systematically starting from the existing configuration. Results from symmetric cases here tell us what kind of improvements that are of interest to examine.

### 5.4 Timetable free analyses and timetable flexibility

One of the most important contributions of this thesis is the idea of timetable-free analyses. In many cases railway operation analyses are based on some kind of timetable assumption that implies that only one, or a few, operation variants (timetables) are tested.
However, the timetable develops much faster than the infrastructure and sooner or later the system is operated differently than it was originally planned for. There are several Swedish examples of this and so timetable-free methods would be an important complement to existing methods of operation analysis.

By evaluating several hypothetic timetables the analysis becomes much more general and the knowledge of the system deeper. The train separation that occurs on single-track lines makes it possible to perform this kind of analyses.

A natural development of the timetable-free concept is the evaluation of timetable flexibility. This idea ought to be an important contribution, in particular when asymmetries from a real line are included in the analysis. It is useful to analyse timetable flexibility both during timetable construction and in infrastructure planning. In the latter case it is a way to check sensitivity to timetable changes relative the dimensioning “planning timetable”.

The main contribution regarding timetable flexibility is the idea and conceptual thoughts rather than the proposed definitions that need to be developed further. Some kind of flexibility measure is probably required when infrastructure measures, such as partial double-tracks and decreased inter-station distances are to be completely analysed.

### 5.5 Tool for infrastructure and timetable planning

SAMFOST is a valuable tool in infrastructure planning as well as timetable construction. This is exemplified by the evaluation of different inter-station distances and lengths of partial double-track. A fundamental result from these evaluations is that the operation of line sections with only ordinary crossing stations and no passenger stops are insensitive to punctuality. It is, however, difficult to achieve large decreases in crossing time by introducing only crossing stations. Instead, this measure mainly affects the crossing time variance. Shortened inter-station distances therefore mean increased reliability rather than increased time efficiency in crossings.

Partial double-tracks, on the other hand, result in shorter crossing times. This, however, is achieved at the cost of increased sensitivity to punctuality. Partial double-track is therefore an appropriate measure in some situations but not generally. The same holds for a combined crossing and passenger stop. In this case it is important that the time supplement added in the timetable is adjusted to the level of punctuality.

These results are important to the Swedish Rail Administration, who are planning for new single-track lines and upgrades of existing lines. The clear results of partial double-tracks and passenger stops tied to crossings also confirm the experience of sensitivity to punctuality that is characteristic for such lines.

The following figures show a real example where SAMFOST has already been used to evaluate alternative improvements of a Swedish line, the Svealand line southwest of Stockholm.
Figure 5-1 shows the characteristics of the existing Svealand line. The line is constructed for 60-minute traffic which implies a crossing every 30 minutes. In order to make 30-minute traffic with time efficient crossings, the infrastructure has to be improved.

Two different improvement strategies are shown in the following figures. In figure 5-2 the existing partial double-track is extended and a new partial double-track is introduced 15 minutes run time (one direction) away from the first one. This strategy maintains the main features of the line with a highly varying crossing time function.

An alternative improvement is shown in figure 5-3. In this case the existing partial double-track is extended to the right so that it reaches the existing double-track system. On the rest of the line three new crossing stations are introduced. This results in a less fluctuating crossing time function.
Figure 5-2 Improvement strategy: New and extended partial double-track. Crossing time function (solid thin), mean crossing time (solid bold) and standard deviation for crossing time (dashed).

Figure 5-3 Improvement strategy: New crossing stations and “eliminated” partial double-track. Crossing time function (solid thin), mean crossing time (solid bold) and standard deviation for crossing time (dashed).
6 Final remarks and future work

This thesis deals with several fundamental properties of single-track railways. The SAMFOST model is a first step to describe these properties. Having examined only simple operational cases it is natural to go on and model also more complex situations.

An obvious case of special interest is the effect of series of crossings that interfere. To model such cases it is natural to consider time supplements and their effects on operation. Not only the amount of supplement but also the distribution along the line is important.

No internal sources of disturbance are modelled in SAMFOST, i.e. the trains are only given initial delays. One way to make the model more realistic is to include delays that occur during the modelled train courses such as dwell time extensions etc. It is also possible to include driver behaviour in the model, which would make the trains behave more stochastically. This factor may also represent shifting adhesion situations etc.

Another situation, which is closely related to series of crossings, is operation during congestion. When more than two trains affect a crossing a second order of crossing time, which is caused indirectly by the position of other trains, is added. Analysing such situations would probably give valuable knowledge about capacity of single-track lines.

This thesis concerns only homogenous passenger traffic. In reality, single-tracks are most often operated with mixed traffic. A natural continuation would therefore be to examine cases with mixed traffic, freight trains etc.

All these extensions mean that more variables are included. This in turn indicates that a more systematic, multivariate analysis may be useful. Therefore, an important part of the future work is to develop methods to examine and describe the effect of several variables.

Another possible continuation is a further development of the timetable flexibility concept. This could be done either on the basis of the existing SAMFOST model, or on a further developed model that includes the features discussed above. Further knowledge of timetable flexibility may be very valuable in the work of timetable construction as well as in the infrastructure planning.

The work has explicitly shown the importance of punctuality. Since punctuality is a great problem in railway operation, deeper studies of how punctuality affects timetable flexibility, capacity and infrastructure design are valuable in the whole railway sector.

Two principles of infrastructure design are examined in this thesis. The analysis shows that there are several other infrastructure strategies that could be of interest. Therefore, a natural continuation of the work is to examine more infrastructure configurations. One example of interest might be shorter inter-station distances locally around stations with passenger stops. Such a configuration would help to
reduce the mean crossing time as well as the standard deviation for crossings planned to the passenger stop station.

The infrastructure measures examined in this thesis affect the operation properties quite differently. Construction costs also differ. Therefore it is of interest to compare different measures. A complete comparison requires evaluation of public economy. This includes traffic factors such as travel times, frequency of service, punctuality as well as infrastructure factors such as construction costs and operation costs. This type of comparison is a natural continuation of the evaluation presented in this thesis.

All these possible fields of further studies aim to increase knowledge of a complex technical system that is surprisingly difficult to model. Nevertheless, more knowledge of railway operation is very important if the railway is to contribute to the development of society.
7 References


