Proceedings of the 35th European Safety and Reliability & the 33rd Society for Risk Analysis Europe Conference
Edited by Eirik Bjorheim Abrahamsen, Terje Aven, Frederic Bouder, Roger Flage, Marja Ylönen

doi: 10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P2837-cd

Modelling impact of defects and faults on railway punctuality

Jørn Vatn

Department of Mechanical and Industrial Engineering, NTNU - Norwegian University of Science and Technology, Norway. E-mail: jorn.vatn@ntnu.no

Josep Rodríguez Estrella

Department of Mechanical and Industrial Engineering, NTNU - Norwegian University of Science and Technology, Norway. E-mail: josep.rodriguez99@gmail.com

To prioritise maintenance and renewal in the railway infrastructure it is required to have a good understanding of the impact defects and failure have on punctuality. The paper presents and compare two different assessments methods. The first approach is based on a network model where each train journey is modelled in terms of speed, acceleration, opportunities for crossing etc. The case study is performed on a Norwegian single track line, hence the limited places for crossings will be important in the model. For such a simulation model it is then possible to induce a fault or a degradation which either represent a full stop, or speed reduction. Depending on "where and when" the failure or defect occur, the consequences may vary. The second approach is based on analysis of statistics. In Norway there is a system here a so-called "event ID" is linked to each train that arrives late to a station. This means that the train operations centre is assigning a specific condition as a reason for the delay. This provides lists of "who to blame" for the punctuality. Fault and defects in the infrastructure is part of this picture, but in addition comes delays due to failures of the rolling stock, other train being delayed causing delayed departure etc.

Keywords: Railway Punctuality, Fault and defects, Maintenance and renewal, Modelling

1. Introduction

1.1. *Background*

Railway punctuality is a main concern in most European countries and in a study by the European Commission - Mobility and Transport (European Commission, 2023) some overall figures of the situation are provided. Figure 1 shows the evolution of punctuality of regional and local services during the years 2015 and 2018 to 2020. In Switzerland punctuality is calculated based on delays larger or equal to 3 minutes whereas the remaining countries use delays larger or equal to 5 minutes as a basis for punctuality calculations.

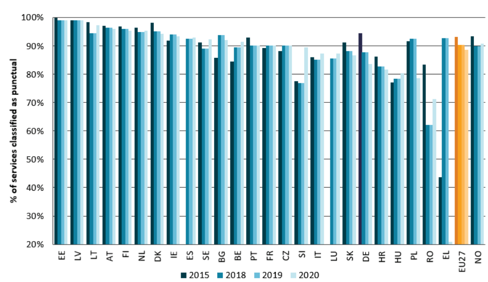

Fig. 1. Punctuality statistics - Norway to the far right

Although Norway is not that bad compared to other European countries, customer satisfaction is rather low and the Government has launched a new strategy for both road and rail transportation where then main focus is to maintain existing infrastructure rather than building new. The strategy comes with increased maintenance budget which is not too early. However, an important question is how to spend the budget in a way which gives the highest improvement on punctuality.

In maintenance modelling focus is often on establishing the relation between the unavailability of a system and the maintenance strategy. Multiplying unavailability with cost per hour combined with the maintenance model gives a very good starting point for maintenance optimization. In railways this is not that straight forward because the relation between the unavailability of a line segment and the punctuality is hard to establish.

1.2. *Objective and novelity*

The objective of this paper is therefore to establish punctuality models that can be used to support prioritization of maintenance. In particular a scenario based model where train operation is simulated will be developed. Such a model shall be able to simulate failures and defects in the infrastructure and how they impact the circulation in the period of time until the failures are repaired, but also cascading effects until normality is achieved.

To validate the model the result from the simulations will be compared to statistics obtained from the Norwegian Railway Administration (Bane NOR).

The main contribution of this work is the development of a scenario model to assess the punctuality impact of events like failures, speed reduction and environmental related events. Such a model can then be used to plan and optimize maintenance and renewals.

1.3. *Related work*

1.3.1. *Analytical models*

Goverde (2005) presents in his PhD thesis, *Punctuality of Railway Operations and Timetable Stability Analysis* a new way to analyse railway timetable stability using max-plus algebra. His objectives where to:

(1) Develop a data mining tool to retrieve accurate and reliable realizations of train movements from data records of the signalling and safety system.
(2) Perform a detailed punctuality analysis, with the emphasis on fitting probability distributions of characteristic operations performance indicators and quantifying disruptive train dependencies.
(3) Develop an analytical approach to evaluate and quantify critical network dependencies on capacity utilization and timetable stability.

The work demonstrates the practical application of these methods to railway systems, as it is applied to the Dutch railway timetable. Thus, this project combines analytical methods with empirical data analysis to achieve more accurate and precise railway timetables. However the work is not focusing on how the punctuality model could be used to model failures and how to use these models for maintenance decision support.

1.3.2. *Data-driven approaches*

Jiang et al. (2019) combines the interpretability of logistic regression models with the robustness and accuracy of Random Forest models to create a combined model which was applied to predict punctuality. The data consists of relative timetable deviation of train movement for all stations, as well as punctuality observations at destination stations. The data was recorded for both passenger and freight trains in Sweden between year 2017 and 2018.

In Sweden, punctuality is measured as a weighed reliability percentage, categorized depending on the lengths of the delay. In the data processing, the statistics for relative deviation and punctuality are merged on every passenger and freight train. The Swedish Administration has a goal that 95% of all trains arrive within the 5 minutes and 59 seconds limit.

In the data of study is included: train characteristics (train number, type of train, operator. . .), infrastructure features (rail class), type of error (categorized as infrastructure, operators, traffic control, accidents or is a consequence of previous error), and time information with the length of relative delay, and the actual time for arrival and departure.

1.3.3. *Simulation programs*

Looking in the market, two simulation programs were tested to develop the punctuality model of this project: OpenTrack and AnyLogic.

OpenTrack is a user-friendly railroad network simulation program developed at the *Swiss Federal Institute of Technology's Institute for Transportation Planning and Systems*. Simulates rail system operations based on user defined train, infrastructure, and timetable databases. OpenTrack (2024).

Also, it is a microscopic synchronous simulation model, and it gives the user full of flexibility for defining dispatching logic as well as operational variables.

OpenTrack has three main elements:

(1) **Input data:** Divided into three modules: Rolling stock, infrastructure, and timetable. The input information is stored in the program's database.
(2) **Simulation:** As a dynamic rail simulation program, it depends on the state of the system at each step, and also, the desired scheduled defined previously.
(3) **Output:** Many formats are available to represent the results: graphs, tables or images.

Currently, many railway operators and administrations use OpenTrack. Some of them are: **First Rail** (UK), **Ferrocarils de la Generalitat de Catalunya, FGC** (Spain), **SNCF** (France) or **Operail** (Estonia), among others.

Although OpenTrack is very strong in simulating the train circulation it is not straight forward to model deviations and defects for the purpose of this study, This would have required an interface where it would be possible to "manipulate" the infrastructure model which was not possible within the scope of the present study.

1.3.4. *AnyLogic*

AnyLogic provides a single platform for all the business' dynamic simulation modelling needs. It can simulate any system or process related to business or research, from manufacturing lines and logistics to marketing campaigns and social changes.

It is a multi-method modelling software, leading to use with **system dynamics, discrete event modelling, and agent-based modelling**.

AnyLogic has clients worldwide. Big companies in consultancy, such as Deloitte and PwC, also logistics as DHL or FedEx or even in the automotive industry, like Volvo, Volkswagen or Ford, support this software and its wide range of use. Compared to OpenTrack AnyLogic was consider easier to use as basis for our punctuality modelling, but also here the flexibility to implement defects in a realistic manner was limited.

1.3.5. *Factors affecting punctuality*

Other aspect

The main objective of this work is developp punctuality prediction models. Such models are important for optimizaiton and decision support. For example Sánchez Navarro (2021) use punctuly models to minimize the average waiting time at stations, optimizing both the schedule and the train type selection according to its service.

In order to build punctuality models it is important to understand the contributing factors to puncutality. Some relevant studies are presented by Olsson and Økland (2020); Olsson and Haugland (2004).

2. The Network Simulation Model

2.1. *Introduction*

The objective of the network simulation model is to simulate the operations of trains and calculate the delays for each train at each station. The model implemented in Python contains the following elements:

- A network representation where each station is an object with line segments are connecting the stations.
- The trains, i.e., each trains that operate a line one day
- A traffic management / train operations model defining rules for crossings
- A disturbance model used to specify failures and defects and other conditions that causes disturbance in the train operations

The line to study is **Dovrebanen F6**, operated by Bane NOR. Taking into account that during the week are different timetables, a regular working day with focus on passenger trains was considered leaving for further studies the addition of freight trains and make a more complex understanding of the line.

Figure 2 shows the map of the aforementioned line with the stops along the section Trondheim S $\leftrightarrow$ Oslo S.

A summary of the model is:

- **Name of the line**: Dovrebanen F6.

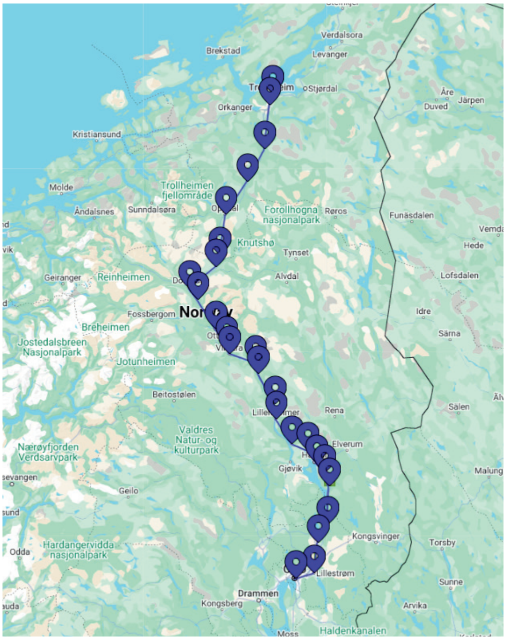

Fig. 2. Dovrebanen F6 train route

- **Type of train**: Passengers only.
- **Type of day**: Regular working day (Monday to Friday).
- **Stations**: 22 conventional stations

In addition there are 41 "virtual stations", that is sections that allows crossings without enabling passengers to get on and off the train. These are very important as most of the line has only one track for both directions.

2.2. *Track Model*

The track model allows to specify the stations with speed profiles along the track between stations. The speed profile is obtained from the computerized maintenance management system used by Bane NOR. It contains a description of the track with its objects, failure and maintenance reports and the preventive maintenance program, and maximum speed for subsection with associated lengths (km) between stations.

The travelling time between stations depends on acceleration and deceleration performance as well as limitation in the allowed maximum speed for each train. In the prototype model developed a simplification is made by using the time table to calculate the average speed each train can run for each line segment.

From the time table it is also possible to find the number of minutes the trains stops at each station for crossing and for take on and leave off passengers.

2.3. *Train model*

In the train model each train is represented with properties. In the current model only maximum average speed it can run between the stations is used, and in future development of the models also acceleration, deceleration and maximum speed will be specified.

There are 2 directions (Trondheim S $\rightarrow$ Oslo S and Oslo S $\rightarrow$ Trondheim S), with 6 trains on each.

First departure from Trondheim S is at 5:55 h and the last one at 23:17 h. On the other side, first train from Oslo S, leaves at 8:02 h, and last of the day at 22:50 h. Two night trains (6 and 12) operate regularly, arriving the following morning to their destination.

The average trip duration is 6:56 h. The fastest train to do this route is the **train 5 (to Oslo S, 15:26 h)** with a scheduled time of 6:36 h. Night trains use longer time to increase comfort for the passengers.

Given the waiting time, arrival and departure times and length of the section, it is possible to calculate the average speed, useful to put time to the non-passengers stations and the following modelling of the line.

Taking into account the waiting time at each station (depends on the train), the distance between the stations, the departure time, as well as the non-stop stations, is enough to represent the route of a train.

2.4. *Train operations*

The train operations module is used to run the trains and ensure that they can cross at stations where crossings are allowed. A train cannot leave a station earlier than it's scheduled departure time for stations taking up or leaving off passengers. For stations where a specific train is not scheduled for taking up or leaving off passengers three is not such limitation. Also in the train operations module there are rules that ensures that a train cannot enter a line segment occupied by another train. This applies for the part of the route where there is only one track. For the double track part of the route the track is represented by parallel segments.

Train movement logic

This block is crucial for the model development. In the Python code each station has a variable *occupied_by* where a station is **free or occupied by the current train**. A station can also be **occupied** by other train (occupying):

First, get time from the last and next departure of trains. Then three more options appear:

- **Current train arrives earlier**, meaning the **occupying** train should wait until the current train arrives to the following station.
- **Trains cross each other at the station**. The scheduled departure time for the train arriving earlier is after the arrival of the other train.
- **Occupying train arrived earlier**, so the current train should wait until the **occupying** train leaves the track.

A complete presentation of the logic for train operations and the corresponding Python code can be found in Estrella (2025).

2.5. *Disruptions*

The previous modules ensure that we can "execute" the time table for a normal day. However, since the objective of the modelling is to assess the punctuality implication of disruptions, the model has also implemented a disruption module. The following situations/scenarios could be specified with corresponding details related to the location, durations etc.:

(i) Reduced speed of a train.
(ii) Reduced maximum speed in a section.
(iii) Closed section.
(iv) Closed track.
(v) Closed track and reduced max speed.

In the **Train Operations** module it will be checked if a section is affected by one of the scenarios specified. For example if there is a speed reduction on a line segment for some hours, trains cannot run faster than the specified speed.

2.6. *Example*

Estrella (2025) presents example results for all the disruption scenarios listed above. In the following we discuss one of these examples, i.e., the "Closed track and reduced maximum speed". This example is a combination of two simple disruptions: **Closed track** and **reduced maximum speed**. These are the conditions:

- *Speed variation*: Full stop ‖ Up to 60 km/h.
- *Time slot*: From 7 am until 10 am ‖ All the day.
- *Station(s) affected*: Oppdal ↔ Dombas ‖ Dombas ↔ Hamar.
- *Train(s) involved*: All the trains that are between these stations during the affected time slot ‖ All the trains.

Figure 3 shows the results when these conditions are applied to the network. The labels used are: i) *ATA*: Indicates the real time of arrival at the station, ii) *ETD*: Scheduled time for the departure. Ideally, should be the same as ATD, iii) *ATD*: Real time of departure at the station. It includes the waiting time and any delays that may occur. iv) *route*: A dashed line to easily keep track of the journey. and iv) *OTT*: When used, it represents the route of the train under normal conditions (without any disruption or delay).

As we can see this scenario represents a a collapse in the network and was the worst case scenario in the study by Estrella (2025) .

3. Statistical model

3.1. *Introduction*

In the network model it is possible to assess the effect of a given disruption by specifying it in the model and "execute" the operation for that day. To validate the result a statistical model has been developed. In the pilot study a simplified version of the model were tested, but in the following we also present a more comprehensive approach to be implemented at a later stage. The statistical model

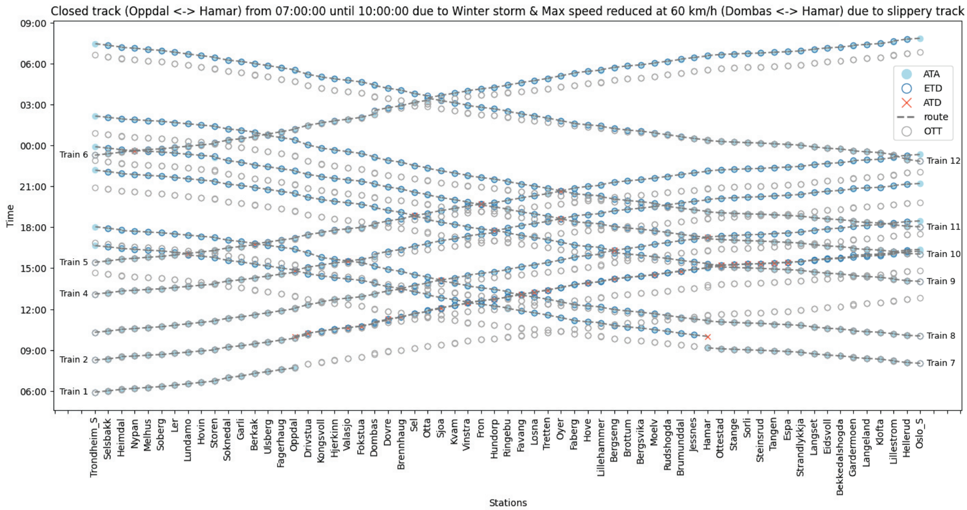

Fig. 3. Closed track due to winter storm and reduced max speed due to slippery track

will utilize information from two systems in Bane NOR, i.e., the TIOS and the BaneData system

3.2. *TIOS*

In TIOS all passenger trains in Norway are registered. If a train is more than four minutes late, according to the set timetable, this is registered with a code for specifying the cause of the delay. A total of sixteen different codes are divided into the general reason codes: infrastructure, traffic management, train company and external conditions Olsson and Økland (2020). TIOS holds information about both relative and absolute delay. While the relative delay provides information about how much delay has occurred since the previous station, the absolute delay provides information about the total delay of a train.

3.3. *BaneData*

BaneData is a register with information about elements that are part of the railway infrastructure, as well as events related to these objects, and is used for maintenance management of the infrastructure. Each event registration includes information about when the fault occurred, the duration of the fault, fault correction and the implementation of maintenance.

In the HL (from the Norwegian word *Hendelseslogg*, meaning event log), all deviations from the normal operating situation on the railway must be registered. Each event is assigned to an event number (HL-ID), and this can be used to link information between TIOS and BaneData.

In TIOS the event number in the event log is linked to a delay and a cause, while in BaneData the event number is registered to the element that has failed and led to the delay Olsson and Økland (2020). In this way, you can get a broader picture of the delay cause than if you only look at the overall cause categories for delay that are used in TIOS.

3.4. *Preliminary statistical results*

Table 1 depicts the causal codes with their explanations.

Figure 4 shows the distribution of delays according to its causal code.

Track failures are dominating when it comes to number of delays which points to the importance of more maintenance.

However the magnitude of the delay is also important and 5 shows that code 3 - Electricity problems causes the largest delay, i.e., one hour in average.

From statistics as presented by Figure 4 we can obtain contribution to delays by the various

Table 1. Codes used in TIOS

Code	Description
1	The track
2	The signaling system and switches
3	The electricity
4	The train radio and communication
5	Planned works
6	Another train had problems
7	Delayed by another train. Usually crossings
81	Problems with the train
82	Late to the origin station
83	Missing people to operate the train
84	Long time at platform
85	Driving too slow
91	Late from Sweden
92	Extreme weather conditions
93	Accidents and derailments
94	Unwanted things

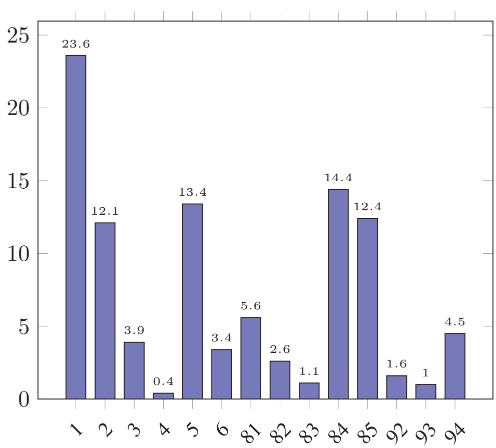

Fig. 4. % of delays by causal code

causal codes. On a general level this can be used to prioritize measures to improve punctuality, e..g, track maintenance. However, such overall statistics cannot be used to assess the consequences of for example a full stop on the line for a given time period as was shown in the network model example.

The original objective was therefore to combine data from TIOS with data from BaneData through the HL-ID key. Within the time-frame of the study this has not been possible, so a simplified

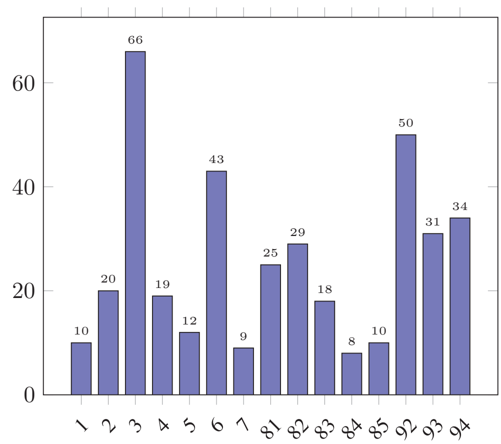

Fig. 5. Average delay time

approach was applied:

Selected delays from the TIOS system were structured by manually processing several trains affected by what seems to correspond to one particular event. Then each of these events are described by location, which time period was affected, and the actual implication for traffic, e.g., speed reduction. This was done to compensate for the lack of information ideally to be found by the HL-ID key.

4. Comparison between the two models

Several scenarios from the TIOS database for Dovrebanen were tested to evaluate the railway network model.

This first case represents the timetable of February 3, 2022. On this day, there were two main causes of delay: i) *Ref. 2 — Track surface coating including salt coating* affecting train 12, and ii) *Ref. 81 — Locomotive fault* affecting train 7.

Other minor faults such as delay for crossing or slow driving have not been considered due to their low contribution to the general timetable, and also, to simplify the model.

Table 2 shows impact of the scenario on the affected trains and Table 3 shows the deviation between the historical data and the result obtained by the network model which is quite good.

The second case represents the timetable of May 31, 2022. On this day, there was a main

Table 2. Description of case 1

Train	Distruption	Station(s)
7	Speed reduction of 15%	Oyer ↔ Hjerkinn
12	Full stop of 40 min	Dombås

Table 3. Description of case 2

Train	Delay	Data vs mode
7	0:23 h (5.8%)	0:03 h
12	0:38 h (8.1%)	0:00 h

cause: *93 — Collision* affecting trains 2, 3, and 4 (not the same collision).

Table 4. Description of case 2

Train	Distruption	Station(s)
2	Full stop of 50 min	Støren
3	Full stop of 55 min	Fagerhaug
3	Speed reduction of 25%	Faberg ↔ Oslo S
4	Full stop of 60 min	Lundamo

Table tab:case3 shows impact of the scenario on the affected trains and Table **??** shows the deviation between the models which also are quite good for this case.

Table 5. Description of case 2

Train	Delay	Data vs model
2	0:53 h (13.5%)	0:07 h
3	0:25 h (5.8%)	0:02 h
4	1:10 h (17.4%)	0:00 h

5. Discussion and further work

A network model for assessing the punctuality impact of defects and failures in the railway infrastructure is presented. Such a model is required for prioritization of maintenance resources to ensure that the maintenance budgets are used to give the highest punctuality improvement. The model has been tested against a statistical model to validate the output. In the current version of the model simplifications has been introduced and it is required to improve the model. Important improvements will be to add a more realistic speed model for each train. Further not only long distance passenger trains should be modelled, but also local trains around Trondheim and Oslo should be specified in addition to freight trains.

The validation of the model is based on a very simple statistical model. The aim is to develop this model by using the HL-ID to link punctuality events more explicit to the maintenance database describing the defects and failures.

References

Estrella, J. R. (2025). A comparative analysis of network simulation and statistical approaches on a norwegian railway line. Master thesis, Norwegian University of Science and Technology.

European Commission (2023). Rail Market Monitoring (RMMS).

Goverde, R. M. P. (2005). *Punctuality of Railway Operations and Timetable Stability Analysis*. Ph. D. thesis, Delft University of Technology.

Jiang, S., C. Persson, and J. Akesson (2019, 10). Punctuality prediction: combined probability approach and random forest modelling with railway delay statistics in sweden. pp. 2797–2802.

Olsson, N. O. and H. Haugland (2004). Influencing factors on train punctuality—results from some norwegian studies. *Transport Policy 11*(4), 387–397.

Olsson, N. O. E. and A. Økland (2020). Factors affecting punctuality in rail freight transportation. *Research in Transportation Economics 88*, 100824.

OpenTrack (2024). OpenTrack Railway Technology. OpenTrack Railway Technology Website. [Online]. Available: https://www.opentrack.ch/.

Sánchez Navarro, P. (2021). Diseño de un modelo de optimización para la determinación de horarios en líneas ferroviarias de tránsito rápido con flota heterogénea y demanda variable.