Impact of railway disruption predictions and rescheduling on passenger delays

Ghaemi, Nadjla; Zilko, Aurelius; Yan, Fei; Cats, Oded; Kurowicka, Dorota; Goverde, Rob

DOI
10.1016/j.jrtpm.2018.02.002

Publication date
2018

Document Version
Final published version

Published in
Journal of Rail Transport Planning and Management

Journal of Rail Transport Planning & Management 8 (2018) 103–122

Contents lists available at ScienceDirect

Journal of Rail Transport Planning & Management

journal homepage: www.elsevier.com/locate/jrtpm

Impact of railway disruption predictions and rescheduling on passenger delays

Nadjla Ghaemi[a,*], Aurelius A. Zilko[b], Fei Yan[a], Oded Cats[a], Dorota Kurowicka[b], Rob M.P. Goverde[a]

[a] *Department of Transport and Planning, Delft University of Technology, The Netherlands*
[b] *Delft Institute of Applied Mathematics, Delft University of Technology, The Netherlands*

ARTICLE INFO

Keywords:
Railway disruption
Prediction
Dependence model
Short-turning
Passenger assignment

ABSTRACT

Disruptions such as rolling stock breakdown, signal failures, and accidents are recurrent events during daily railway operation. Such events disrupt the deployment of resources and cause delay to passengers. Obtaining a reliable disruption length estimation can potentially reduce the negative impact caused by the disruption. Different factors such as the location, cause of disruption, traffic density, etc. can determine the disruption length. The uncertainty inherent to the variability of each factor and the unavailability of sufficient data results in a wide distribution of disruption lengths from which a certain value should be selected as the length prediction. The rescheduling measure considered in this research is short-turning the trains that are heading to the disrupted area. To investigate the impact of the disruption length estimates on the rescheduling strategy and the resulting passengers delays, this research presents a framework consisting of three models: a disruption length model, short-turning model and passenger assignment model. The framework is applied to a part of the Dutch railway network. The results show the effects of short (optimistic) and long (pessimistic) estimates on the number of affected passengers, generalized travel time and number of passengers rerouting and transferring.

1. Introduction

Railway operations are repeatedly disrupted by events such as technical and mechanical failures of infrastructure and rolling stock, traffic accidents and malicious attacks. Railway timetables are usually designed to compensate for some delays by including time allowances. However in case of long disruptions and infrastructure unavailability, these time allowances are ineffective and a new timetable should be designed with adjusted train services. Before making any decision regarding the traffic, the relevant information about the disruption should be collected. Upon occurrence of a disruption, there is usually a high level of uncertainty regarding the situation. It takes some time before the exact location and the nature of the disruption is known to the traffic controllers. Once the cause of disruption and the location is known, the traffic controllers are able to proceed with handling the disrupted traffic. An essential piece of information that has a crucial role in their decisions regarding the traffic is the predicted disruption length. Since any change in the timetable is costly, if the predicted length is shorter than a specific threshold then they might decide not to implement major changes to the timetable. In case the length is larger than the threshold, the prediction is used for rescheduling the timetable. Moreover the disruption length and the changes of the timetable should be communicated to the passengers so they can make informed decisions about their trips when they are disturbed. Thus a reliable disruption length prediction is key

* Corresponding author.
E-mail address: N.Ghaemi@tudelft.nl (N. Ghaemi).

https://doi.org/10.1016/j.jrtpm.2018.02.002
Received 22 September 2017; Received in revised form 12 February 2018; Accepted 12 February 2018
Available online 03 May 2018

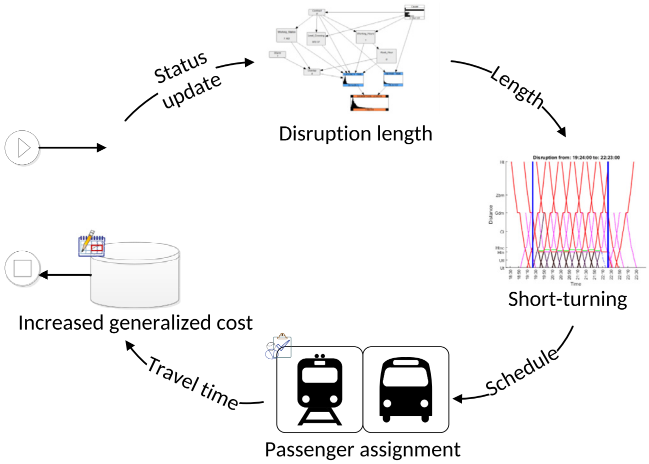


Fig. 1. The interaction between the models.

information in designing a new timetable that can limit the negative impact on the passengers.

The length of a disruption depends on many factors like the type of failure, the component that failed, the repair needed, and the spare parts that a repair crew has with them. However, data on most of these factors are not available, which leaves much uncertainty in any prediction of the disruption length. Zilko et al. (2016) developed a probabilistic Bayesian Network model to predict the disruption length using conditioning on information that becomes available about a disruption. However, most factors included are indirect, like accessibility of the disruption location. Hence, the obtained distributions of disruption lengths still exhibit a wide range. Then the next question is how to get a point estimate from this distribution that can be used by traffic controllers as a prediction of the disruption length and take actions accordingly. Simply taking the mean (expected value) may not be effective. On one hand, a point estimate could be optimistic and underestimate the disruption by which traffic controllers have to rethink their decisions and passengers may have taken the wrong decisions about their travel compared with the optimal choice if they would have known the exact disruption duration. On the other hand, a point estimate could also be pessimistic and overestimate the disruption which might lead to unnecessary train cancellations and long travel times for passengers.

In this paper a framework is proposed to investigate the effects of different choices of predictions on the rescheduling solution and consequently the passenger delay. The contribution of the framework is the unique integration of three components (see Fig. 1):

- Estimating the disruption length as a point estimate from a conditional distribution based on available data.
- Rescheduling the timetable given the estimated disruption length.
- Measuring the passenger delays based on the computed adjusted schedule.

The framework provides the possibility of reproducing different disruption scenarios, where the information about the disruption is gradually updated and accordingly the timetable is rescheduled and finally the impact on the passengers is measured. This allows testing the consequences of different disruption lengths and their over- or underestimation with the corresponding mitigation measures on passenger flows and travel time losses.

In the remaining of the paper, the process of handling disruption and the relevant literature is presented in Section 2. The three components are described in more details in Section 3. The modeling framework is then demonstrated using an application to part of the Dutch railway network in Section 4. In this Section the impacts of the optimistic and pessimistic estimates are modeled and assessed. Section 5 concludes with practical implications and directions for future studies.

2. Disruption management

The decisions regarding the rescheduling of resources need to be carefully communicated between the railway infrastructure manager, the train operators and other involved actors to ensure the feasibility of the plan. To facilitate the challenging task of the traffic controllers in such cases, many countries use contingency plans designed specifically for disruption scenarios (Chu and Oetting (2013)). In The Netherlands these plans are manually designed by expert traffic controllers and are specific for each location and disruption case regardless of disruption length.

The proposed solution in the contingency plans is based on the timetable (basic hourly pattern) and the remaining capacity of the disrupted location. The solution instructs the traffic controllers how to deal with the disrupted traffic by determining cancelled services, short-turned or rerouted services and services that are allowed to operate as in the original timetable. Short-turnings are particularly beneficial for isolating the disrupted area, while maintaining services on both sides of the disruption. This implies short-turning the arriving trains at a station before the disruption (on both sides) and continue service in the opposite direction. In case of short-turning, the stations where the short-turning should occur as well as the platforms and departure times also need to be

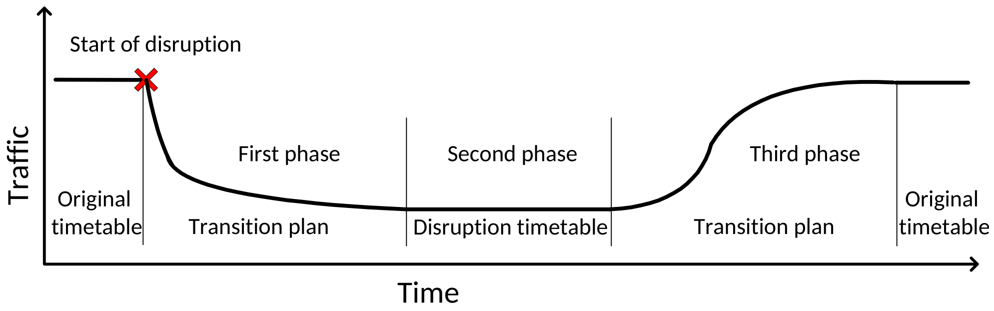


Fig. 2. The service level during disruptions.

determined. By means of simulating the short-turning Coor (1997) concluded that this measure is most efficient in case of large disruptions.

The traffic level during a disruption can be conceptualized as a process that resembles a bathtub (Ghaemi et al. (2017b)). As shown in Fig. 2 some services need to be cancelled due to the disruption. This reduction in train traffic starts immediately after the disruption occurs. Three phases can be identified within the disruption period. In the first phase the traffic controllers are facing lots of uncertainty regarding the disruption location, cause and most importantly the estimation of disruption length. In Dutch railway practice, rough estimates exist for the length of different kinds of disruptions. These estimates are used to inform the passengers about the expected disruption length. Once the location is known, the traffic controllers retrieve the relevant contingency plan designed for that specific location. In case there is a need for repairing the disrupted infra, the repairmen are sent to the field to deal with the cause of disruption. The repairmen estimate the required time for resolving the problem and report it to the traffic controllers. If the cause of disruption is resolved earlier than the informed disruption length, the operation is not resumed before the communicated time. In case the disruption takes longer than the initially estimated length, the passengers are updated with a new disruption length. Throughout this paper the estimates that are longer than the actual disruption length are referred to as pessimistic and those that are shorter than the real length are referred to as optimistic estimates.

Once the communicated length has passed and the cause of the disruption has been removed, the traffic can resume and recover back to the original level. The first and third phases are called transition phases where the operation has a transition from the original timetable to the disruption timetable and vice versa. The contingency plan corresponds to the second phase of the bathtub model where there is a steady though decreased level of traffic. Since there is no reliable disruption length estimation, the contingency plans do not provide any insight regarding the third phase of the transition from the disruption timetable to the original timetable. In practice, the effects of different disruption length estimations on the rescheduled timetable during the three phases of disruption, and consequently the affected passengers, are unknown.

While the bathtub model is widely known and used to conceptualize traffic states during disruptions, only limited research efforts have been devoted to analyzing and modeling railway disruption management. Ghaemi et al. (2017b) provide a review of rescheduling models for disruptions and conclude that only a few studies considered all three phases. Examples of such models were developed by Veelenturf et al. (2016) and Nakamura et al. (2011). However, the disruption length is assumed to be known in advance and passenger delays are not taken into consideration. Meng and Zhou (2011), Yang et al. (2013) and Yang et al. (2014) model the third phase by taking into account the uncertainty of the disruption length. Besides the lack of a timetable for the first and second phase, their approaches do not explicitly model the influencing factors on the disruption length. In particular, Yang et al. (2013) and Yang et al. (2014) model the disruption length as a fuzzy variable that reflects the estimation by expert judgment. Hirai et al. (2009) and Zhan et al. (2015) focus on the first phase where trains need to stop before the disruption area. However both approaches disregard the uncertainty regarding the disruption length and the consequences for passenger delay. De-Los-Santos et al. (2012) and Cats (2016) introduce indexes for measuring network robustness by measuring the effects of disruptions in terms of changes in passengers travel times. Yet the defined indexes are not suitable for real-time application where the disruption length is not yet known and might get updated frequently. Canca et al. (2016) propose a short-turning model to accommodate extra demand induced by a disruption at the tactical level, but the disruption length is not taken into account. Zhan et al. (2016) and Nielsen et al. (2012) incorporate the uncertainty of disruption length for rescheduling through a rolling horizon framework. Their approaches do not include the impact factors on the disruption length. Moreover, the impact of the rescheduled timetable on the passengers is disregarded. Kumazawa et al. (2008) develop a rescheduling model considering passenger inconvenience. They do not provide any information regarding the length of the disruption. Cats and Jenelius (2014) analyze the impacts of disruptions on passenger welfare using a non-equilibrium passenger loading model. They quantify the value of real-time information provision in case of disruption.

As mentioned above, the contingency plans do not provide any instructions regarding the transition phases and the proposed solution is given independently of the disruption length estimation. The changes of the disruption lengths are discussed by Takeuchi and Tomii (2005). In reality, the disruption length is very uncertain. Having a reliable disruption length prediction is instrumental in devising the rescheduling measures and thus for achieving a smooth and fast transition to the original timetable in the third phase of the bathtub model. To tackle this problem, Zilko et al. (2016) represent the disruption length as a probabilistic model. Several determinants of disruption length are considered from which the joint distribution between disruption length and these factors is constructed with a Copula Bayesian Network. Having the joint distribution enables the traffic controller to obtain a conditional distribution of disruption length when a disruption occurs, by conditioning the model on the observed values of the influencing factors.

A disruption length prediction is derived from this conditional distribution. Having a probability distribution enables the traffic controller to choose different values of prediction corresponding to different quantiles of the distribution. If the controller is optimistic about the disruption length, a lower quantile of the distribution can be chosen. Alternatively, a higher quantile of the distribution can be chosen. Without a reliable length estimation, there is no support for the traffic controllers for the third transition phase.

3. Framework

The framework in a nutshell consists of three models; disruption length model, short-turning model and dynamic passenger assignment model (see Fig. 1). The interaction between these models is designed in such a way that the real-time uncertainty of the disruption is captured. During disruption the information about the problem becomes available to the traffic controllers gradually. To model this gradual information update, an iterative approach is designed. With each information update, a disruption length is predicted and then used for rescheduling. The new schedule is then used for evaluating the impact of the disruption on the passengers. Different causes of disruption would lead to different disruption length distributions. The short-turning model does not consider severe weather conditions that would have altered the entire timetable. Each iteration starts with an information update. The iterative process terminates once the disruption is over. Different performance indicators are defined to evaluate the impact of different disruption length predictions on passenger-related metrics. The three components are explained in detail in sub-sections 3.1 to 3.3 and the interaction between these components and the passenger-related key performance indicators are described in subsection 3.4.

3.1. Disruption length model

We use the probabilistic distribution length model of Zilko et al. (2016) to compute conditional probability distribution of the disruption length. The disruption length is divided into two sequential stages: the *latency time* and the *repair time*. The latency time is the length of time the mechanics need to get to the disrupted site while the repair time is the length of time they need to repair the problem.

The joint distribution between the latency time and repair time with the influencing factors is constructed using a Copula Bayesian Network. As a prototype, a Copula Bayesian Network model has been constructed for disruptions caused by track circuit failures in the Netherlands. These disruptions are affected by eight influencing factors: (1) contract type, (2) distance to the nearest mechanics' workshop, (3) distance to the nearest level crossing, (4) whether or not the disruption is during the mechanics' contractual working time, (5) whether or not the temperature is above 25°C, (6) whether or not the disruption occurs during rush hour, (7) whether or not there is another disruption going on at the same time, and (8) the cause of disruption (Zilko et al. (2016)). To give an example, factor (5) is considered because the track circuits as part of the rail infrastructure are prone to failure for high temperatures.

The Copula Bayesian Network uses a Bayesian network to represent the dependence between the variables. A Bayesian network is a directed acyclic graph consisting of nodes and arcs, representing the variables and flow of influence between the variables, respectively. Fig. 3 (a) presents the track circuit disruption length model. The eleven nodes in the structure correspond to the ten variables in the model and the variable “Disruption Length” which is the sum of the latency and repair times. The distribution of each variable is shown in the box of each node. From these histograms it can be seen whether the variable is continuous or discrete. In the header of each box, the variable name is shown. The footers show the mean value and (after ±) the standard deviation of each distribution. The blue boxes represent the latency and repair time which are influenced by the other variables. Influences are represented by arcs. Eventually, the variable disruption length shown in the orange box is the sum of the latency and repair times. The arcs represent the flow of influence between the variables. The absence of an arc between two nodes indicates (conditional) independence between the variables the two nodes represent.

The joint distribution of the ten variables is constructed using *copula*. A copula is an n-dimensional joint distribution in the unit hypercube of n uniform random variables. It is a popular tool to model the dependence between variables (see, e.g. Nelsen (2006) and Joe (2014)). The theorem of Sklar states that any cumulative distribution function $(X_1, ..., X_n)$, denoted as $F_{1,...,n}$, can be rewritten in terms of the corresponding copula C as

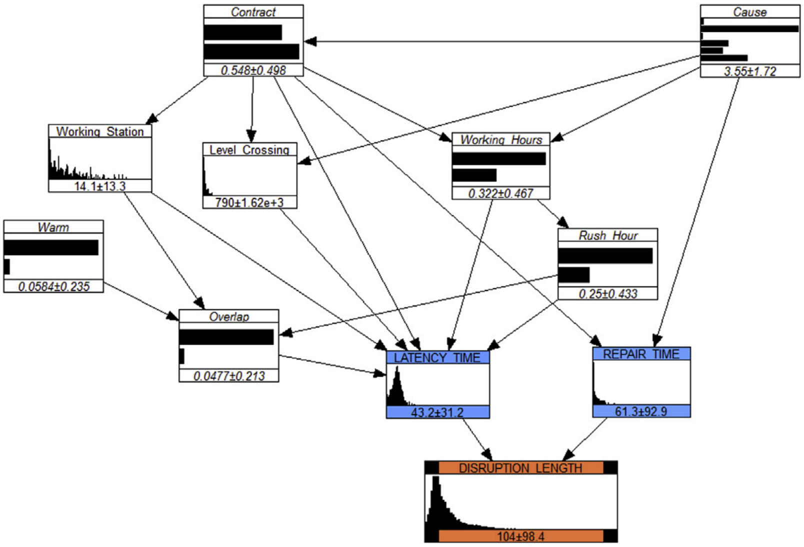
$$F_{1,...,n}(x_1, ..., x_n) = C(F_1(x_1), ..., F_n(x_n)) \tag{1}$$

where $F_i(X_i)$ denotes the marginal distribution of the i-th variable. There are many different copula families. This approach used the multivariate Normal, or Gaussian, copula C_Σ to construct the TC disruption length model. This copula is defined as

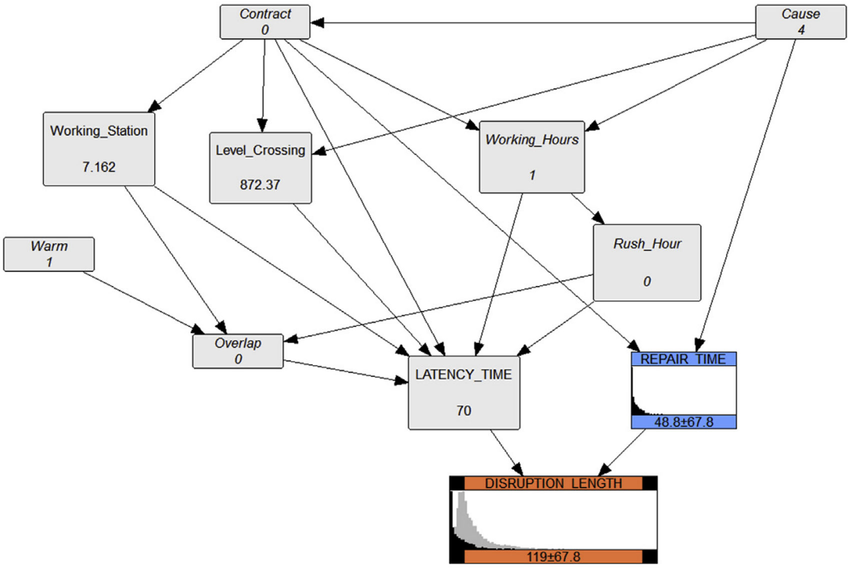
$$C_\Sigma(u_1, ..., u_n) = \Phi_\Sigma(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_n)) \tag{2}$$

where Φ^{-1} denotes the inverse cumulative distribution of a univariate standard normal distribution and Φ_Σ denotes the cumulative joint distribution of a multivariate normal distribution with a mean value of zero and correlation matrix Σ. The parameter Σ of the track circuit disruption length model corresponds to the arcs in the Bayesian network structure in Fig. 3. This copula is of interest because it allows conditionalization to be computed rapidly, a very useful feature in the real-time decision making environment of the traffic control centers.

The copula parameter Σ is computed using the maximum likelihood approach to disruption data from the database provided by



(a) The unconditional track circuit Bayesian network



(b) A conditional track circuit Bayesian network

Fig. 3. The track circuit Bayesian network.

the rail infrastructure manager. The constructed model, as presented in Fig. 3, was validated using empirical data with the model being able to obtain a good conditional distribution of disruption length (Zilko et al. (2016)). Fig. 3(a) presents the unconditional Bayesian network, i.e., the track circuit Bayesian network model when no information is available. Fig. 3(b) presents a conditional Bayesian network when information about the influencing factors is available. Notice that the distribution of disruption length changes. The distribution now has a different shape with updated mean and standard deviation values presented in the footer of the box.

However, conditionalization on the variable Cause can only be performed after the mechanics diagnose the problem and find the cause. This time is called the "diagnosis time". Unfortunately, the diagnosis time is not available in the data and is actually included in the definition of repair time. The data does not provide any information to allow decoupling the diagnosis time from the actual repair time. In practice, usually the mechanics are given 15 min to diagnose the problem after they arrive at the site. Therefore, in this model we assumed that diagnosis time always takes 15 min and the cause is always found in this time.

This modeling component is adopted in this paper and integrated into the real-time prediction and mitigation framework. Its

Table 1
The notation of the input used in the short-turning model.

S	The set of all stations s.
L	The set of all lines l.
V	The set of all scheduled services.
$V_{l,n} \subset V$	The n^{th} set of ordered services in line l.
$v_{l,n}^{i} \subset V_{l,n}$	The i^{th} service in set $V_{l,n}$.
$P_{s,v_{l,n}^{i}}$	The set of platform tracks for service $v_{l,n}^{i}$ in station s.
$S_{v_{l,n}^{i}} \subset S$	The departure and arrival stations $\left(s^{d}_{v_{l,n}^{i}}, s^{a}_{v_{l,n}^{i}}\right)$ of $v_{l,n}^{i}$.
$S_1 \subset S$	The first surrounding station (a) around the disrupted area x.
$S_2 \subset S$	The secondary surrounding station (a') around stations in S_1.
$S^{l} \subset S$	The set of possible short-turning stations for line l.
$d_{l,n}^{i}$	The scheduled departure time of service $v_{l,n}^{i}$.
$a_{l,n}^{i}$	The scheduled arrival time of service $v_{l,n}^{i}$.
$\theta^{min}_{v_{l,n}^{i}}$	The minimum short-turning time needed for service $v_{l,n}^{i}$.
$\tau^{run}_{v_{l,n}^{i}}$	The running time of service $v_{l,n}^{i}$ between two stations.
$\tau^{dwell}_{v_{l,n}^{i},v_{l,n}^{i+1}}$	The dwell time between service $v_{l,n}^{i}$ and service $v_{l,n}^{i+1}$ in the station.
$\tau^{h}_{v_{l,n}^{i},v_{w,z}^{u}}$	The minimum headway time between two train services $v_{l,n}^{i}$ and $v_{w,z}^{u}$.
$\omega^{c}_{v_{l,n}^{i}}$	The penalty for cancelling service $v_{l,n}^{i}$.
$\omega^{z}_{v_{l,n}^{i}}$	The penalty for delaying the arrival of service $v_{l,n}^{i}$.
$Out_{v_{l,n}^{i}}$	The possible scheduled departures for the arriving service $v_{l,n}^{i}$.
$In_{v_{l,m}^{j}}$	The possible arriving services for the scheduled departure $v_{l,m}^{j}$.
$L_{Dist} \subset L$	The set of lines that are affected by the disruption.
$V'_{s,l} \subset V$	The arriving services from line l at station s.
$V'''_{s,l} \subset V$	The scheduled departures in the opposite direction from line l at station s.
M	A large constant.

outcome is the conditional distribution of disruption length, i.e., the distribution of disruption length given actual information of the influencing factors. Moreover, the computation of the conditional distribution is done very fast, and the model has been validated to the Dutch railway disruption data.

The next section describes the short-turning model. The disruption period is used to compute the new schedule. Thus a quantile needs to be selected from the conditional distribution of disruption length that will be used as a point estimate for the disruption length to the short-turning model.

3.2. Short-turning model

The short-turning model is designed to cope with the disruption cases of complete blockages where no train can use the infrastructure during the disruption period. In such cases, all the trains running towards the disruption location should short-turn before the blockage or be cancelled. The short-turning model is the Mixed Integer Linear Program introduced in Ghaemi et al. (2016) that computes the optimal short-turning time and station for each approaching train service. The rescheduling measures include short-turning, partial cancellation, and platform track assignment while the original train orders are maintained. The notation is introduced in Table 1.

We briefly discuss the optimal short-turning problem while focusing on the recovery plan. In this formulation a service is a trip between a departure and arrival (either with or without dwell time). Thus, a train line that may have multiple stops, consists of an ordered set of services performed by a train. Each service is denoted as $v_{l,n}^{i}$ where i indicates the sequence order of the service within the operational line, l is the line number that determines the stops and n determines the time of operation.

The short-turning model is an assignment model that allocates the arriving trains to scheduled departures in the opposite direction. In the example illustrated in Figs. 4–6 there is a disruption between stations a and b. In this example service $v_{l,n}^{i}$, can either short-turn in station a' to serve $v_{l,m}^{j}$, $v_{l,o}^{j}$, or $v_{l,r}^{j}$ or it can continue as service $v_{l,n}^{i+1}$ and short-turn to serve $v_{l,m}^{j-1}$, $v_{l,o}^{j-1}$ or $v_{l,r}^{j-1}$ in station a. Obviously in case $v_{l,n}^{i+1}$ short-turns to serve $v_{l,m}^{j-1}$, this departure would be delayed. The reason is that the arrival time of train $v_{l,n}^{i+1}$ is scheduled after the departure of service $v_{l,m}^{j-1}$ which is shown in grey. The short-turning possibilities are shown by the red arcs for the arriving service $v_{l,n}^{i}$. The output of the model is the short-turning pattern which refers to the selected arc that determines the departure time correlated to an arriving train.

If the short-turning of service $v_{l,n}^{i}$ occurs in an earlier station a' to serve $v_{l,m}^{j}$ as shown in Fig. 5 then the service $v_{l,n}^{i+1}$, all of the associated short-turning patterns in station a, and $v_{l,m}^{j-1}$ should be cancelled. The passengers travelling between these stations will be affected by these cancellations. Notwithstanding, early short-turning can reduce the delay propagation to the opposite stations

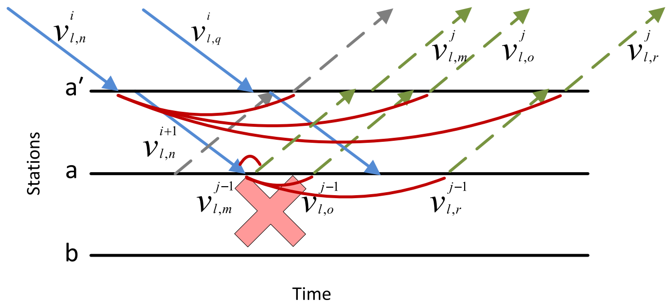


Fig. 4. The possible short-turning patterns.

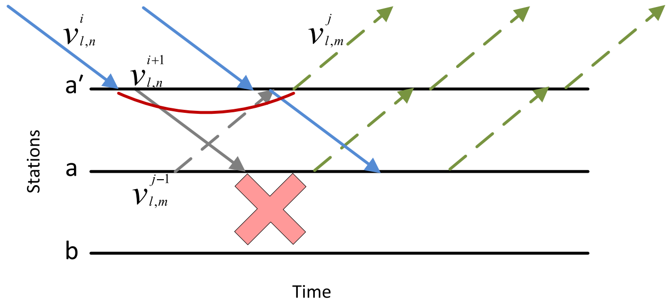


Fig. 5. The cancellation resulting from short-turning in station a'.

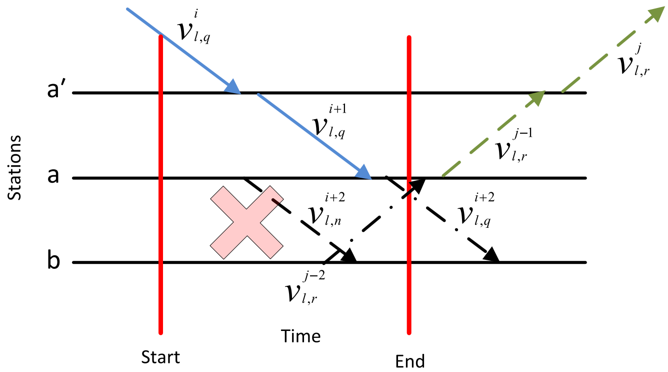


Fig. 6. The transition services shown by dash-dotted arrows ($v_{l,q}^{i+2}$ in one direction and $v_{l,r}^{j-2}$ in the other direction).

(Ghaemi et al. (2016)).

In the preprocessing phase, we define the services that operate in the disruption area during the disruption period. Those services with both planned departure and arrival within the disruption area and period are cancelled. If there is a service towards the disrupted location that departed before the start of the disruption, then it is assumed that this train service could continue its journey. In case the departure of the approaching services towards the disrupted area is within the disruption period and the arrival is close to the end of disruption, then it might be better to wait until the disruption is over and then depart on the original route. The main decision for the final service is whether to continue short-turning or wait until the disruption is over and continue on the original route. This decision concerns those trains that arrive close to the end of the disruption period. If there are more arriving services than the number of scheduled departures in the opposite direction within the disruption period, then the extra train service (due to the periodicity of the timetable, there is usually one extra train service) should wait until the disruption is over before using the recently resolved blockage. If there are more scheduled departures in the opposite direction, then the final arriving train can either short-turn or wait until the disruption is over and continue in the same direction. Based on this decision, there would be a cancellation either for the scheduled departure in the opposite direction, or the scheduled departure in the same direction. The short-turning model computes the disruption timetable for the second phase of disruption and the transitions. With a disruption length prediction we are able to plan the recovery phase where trains are able to continue their original routes and start using the track after the blockage ends.

A set of candidate transition services that might be cancelled needs to be defined. As shown in Fig. 6, service $v_{l,n}^{i+2}$ is cancelled since its departure and arrival is within the disruption period. The transition services ($v_{l,q}^{i+2}$ and $v_{l,r}^{j-2}$ shown by dash-dotted arrows) are planned to depart before the end of disruption and arrive after the disruption. These services are either cancelled or wait until the disruption is over and continue on their original route. In other words, for transition services the possibility for operating on the original route with a possible delay is considered.

In this example, in case service $v_{l,r}^{j-2}$ is cancelled, the following service $v_{l,r}^{j-1}$ is either performed by a short-turning in station a or should also be cancelled. But if the service $v_{l,r}^{j-2}$ is not cancelled, it should depart after the end of disruption and this would introduce a delay to all of the following services performed by this train. For notation simplification, the approaching train services are shown as v' and those services that are scheduled to depart in the opposite direction are denoted by v'''. The model includes four continuous variables; $t^a_{v^i_{l,n}}$ (arrival time of $v^i_{l,n}$), $t^d_{v^i_{l,n}}$ (departure time of $v^i_{l,n}$), $d^a_{v^i_{l,n}}$ (arrival delay of $v^i_{l,n}$), and $d^d_{v^i_{l,n}}$ (departure delay of $v^i_{l,n}$). In addition, there are three binary variables; $c^i_{l,n}$ (cancellation of service $v^i_{l,n}$), $\lambda_{v',v'''}$ (the short-turning pair (v', v''')), and $p_{v^i_{l,n},q}$ (the assignment of platform track q to $v^i_{l,n}$).

The main objective of the short-turning model (3) is to minimize the departure and arrival delay ($d^d_{v^i_{l,n}}$, $d^a_{v^i_{l,n}}$) and the number of cancelled services. The penalties assigned to departure or arrival delay and a cancelled service $v^i_{l,n}$ are denoted by $\omega^{d^d}_{v^i_{l,n}}$, $\omega^{d^a}_{v^i_{l,n}}$ and $\omega^c_{v^i_{l,n}}$, respectively. The objective function is formulated as

$$\min \sum_{v^i_{l,n}\in V} \left(\omega^{d^d}_{v^i_{l,n}} \cdot d^d_{v^i_{l,n}} + \omega^{d^a}_{v^i_{l,n}} \cdot d^a_{v^i_{l,n}} + \omega^c_{v^i_{l,n}} \cdot c_{v^i_{l,n}}\right). \tag{3}$$

In addition to the operational constraints that guarantee the feasibility of the operation such as running times, departures, etc. short-turning constraints are also considered. For the complete MILP model we refer to Ghaemi et al. (2018). Ghaemi et al. (2017a) extend the developed short-turning MILP with a microscopic rescheduling model. To describe the essence of the rescheduling model, the short-turning constraints are explained here. The short-turning constraints ensure a feasible short-turning plan.

$$\sum_{(v',v''')\in Out_{v^i_{l,n}} \cup Out_{v^{i+1}_{l,n}}} \lambda_{v',v'''} = c^{i+2}_{l,n}, \qquad \forall\, v^i_{l,n} \in V'_{s,l},\ s \in S_2,\ l \in L_{Dist}, \tag{4}$$

$$\sum_{(v',v''')\in In_{v^j_{l,m}} \cup In_{v^{j-1}_{l,m}}} \lambda_{v',v'''} + c^j_{l,m} = c^{j-2}_{l,m}, \qquad \forall\, v^j_{l,m} \in V'''_{s,l},\ s \in S_2,\ l \in L_{Dist}, \tag{5}$$

$$\sum_{(v',v''')\in Out_{v^i_{l,n}}} \lambda_{v',v'''} = c^{i+1}_{l,n}, \qquad \forall\, v^i_{l,n} \in V'_{s,l},\ s \in S_2,\ l \in L_{Dist}, \tag{6}$$

$$\sum_{(v',v''')\in In_{v^j_{l,m}}} \lambda_{v',v'''} + c^j_{l,m} = c^{j-1}_{l,m}, \qquad \forall\, v^j_{l,m} \in V'''_{s,l},\ s \in S_2,\ l \in L_{Dist}, \tag{7}$$

$$\sum_{(v',v''')\in Out_{v^{i+1}_{l,n}}} \lambda_{v',v'''} + c^{i+1}_{l,n} \le 1, \qquad \forall\, v^i_{l,n} \in V'_{s,l},\ s \in S_2,\ l \in L_{Dist}, \tag{8}$$

$$\left(1 - \lambda_{v',v'''}\right)\cdot M + t^d_{v'''} \ge t^a_{v'} + \theta^{min}_{v'}, \qquad \forall\, (v', v''') \in Out_{v'}, \tag{9}$$

$$\sum_{q\in P_{s,v^i_{l,n}}} p_{v^i_{l,n},q} + c^i_{l,n} = 1 \qquad \forall\, v^i_{l,n} \in V'_{s,l} \cup V'''_{s,l},\ s \in S_1 \cup S_2,\ l \in L_{Dist}, \tag{10}$$

Constraints (4) make sure that every approaching service should either short-turn at a station prior to the disruption area, which leads to cancellation of one service (in the disrupted area) or continue, that only happens after the end of disruption. For every scheduled departure in the opposite direction there should be only one matching approaching service in case of short-turning. This relation is ensured by constraints (5). In case of early short-turning, there will be two cancelled services that are guaranteed by constraints (6) and (7). If service $v^{i+1}_{l,n}$ is cancelled due to an early short-turning, no short-turning couple can be selected for service $v^{i+1}_{l,n}$. This relation is represented by constraints (8). To ensure the minimum short-turning, constraints (9) are considered. If the service $v^i_{l,n}$ is not cancelled, a platform should be assigned to it. Constraints (10) represent this assignment.

3.3. Dynamic passenger assignment model under disruptions

A dynamic passenger assignment model is developed to represent how passengers are distributed over the network in the event of a disruption. The dynamic passenger assignment model allows assessing the impact of alternative scenarios on passengers by calculating passengers' total travel delay, transfer times and the number of transfers compared to the scheduled timetable. Based on the given rescheduled timetable, the dynamic passenger assignment model generates alternative travel routes for each pair of origin and destination (OD) and assigns passengers to selected routes from the corresponding alternative routes. Data concerning the number of daily passengers between all pairs of OD stations based on smart card transactions was obtained from the Netherlands Railways (*Nederlandse Spoorwegen*/NS). The daily distribution of passenger demand was specified based on data made available by NS which manifests the conventional morning and afternoon peaks. Passengers face different route choice conditions during the course of the

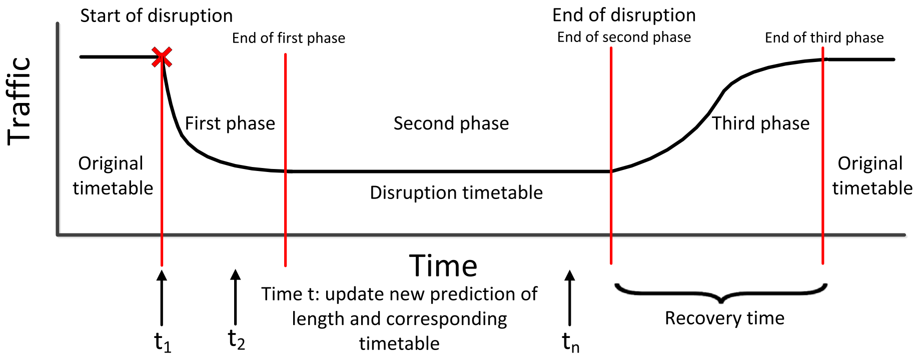


Fig. 7. Transition points in the process of dynamic passenger loading.

disruption. In particular, three phases, as illustrated in Fig. 7. Under normal operations route choice is based on the planned timetable. When a disruption occurs, a rescheduled timetable is generated based on the predicted disruption length and passengers are informed of the new departure times after which they will choose their route based on the prevailing conditions. Hence, passengers choose from a new set of alternative routes based on the rescheduled timetable. Furthermore, passengers who have already boarded a train might need to reroute as a consequence of the disruption and the rescheduling. Additional updates to the disruption length predictions may result in additional rescheduling of train services and consequently the rerouting of passengers. Finally, when the disruption has ended, delays might still occur as the service recovers back to the original timetable. For this reason, the simulation time that is denoted by t_{sim} ends after a certain simulation period denoted by t_{ult}. Here the simulation refers to the probabilistic and dynamic passenger route choice model based on the new schedule. Note that the simulation period is longer than the disruption period.

The dynamic passenger assignment model consists of three main steps. In the first step, alternative routes for each origin-destination (OD) pair are generated, followed by a probabilistic route choice model based on the framework of discrete random utility models. The latter determines the share of passengers that are assigned to each route. In the final step, the passenger travel time is measured. The overall workflow is depicted in Fig. 8.

The dynamic passenger assignment model under a disruption is conducted as shown in Fig. 8 by performing the following sequence of steps:

- **Step 1:** Alternative route generation: Given a passenger OD demand matrix and scheduled timetable, find alternative routes for each pair of stations and calculate passenger's total in-vehicle time and transfer time for each route. This is an initialization phase.
- **Step 2:** Passenger route choice: Simulate passenger generation and train movements. Progress simulation clock from the beginning of the disruption. Use a logit model to obtain the passenger proportion of each route.
- **Step 3:** Network load and passenger travel experience: Obtain the passenger load on each train route segment and calculate the generalised cost for each passenger route departing on each minute.
- **Step 4:** Update: In case the simulation period is not over and the schedule is updated, then repeat from step 2.
- **Step 5:** Transition and normal operations: Repeat step 4 until the last prediction is made, then simulate the model until t_{ult} in order to have a fair comparison among prediction scenarios. Passenger loading results are assessed by calculating the following outputs: passenger total generalised travel time, total passenger nominal travel time, the number of passengers, average passenger transfer time, the number of average transfer, average in-vehicle time, average waiting time, average train load and link load.

The core three steps are described in the following subsections.

3.3.1. *Alternative route generation*

Given a timetable, either the original or the outcome of the rescheduling model, the alternative route generation module computes a set of alternatives from which individuals travelling between a given pair of OD will choose from. Note that for the ODs that contain the blocked section, bus services are considered during the disruption. However the bus services are not considered during the non-disrupted situation, as they are considerably slower than the train services. The choice-set is generated by iteratively searching for routes with an increasing number of transfers. A forward search algorithm is applied where transfer alternatives further

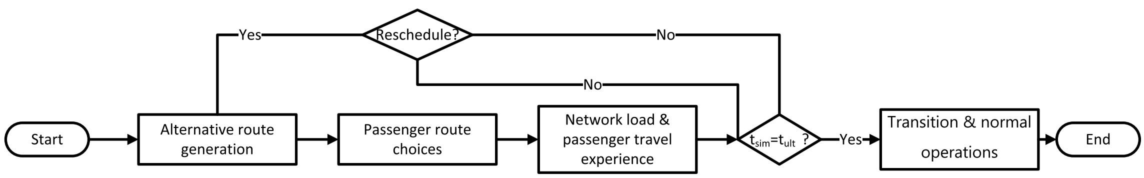


Fig. 8. Passenger loading model under disruption.

downstream are examined by considering all scheduled train trips and their corresponding stopping pattern and scheduled arrival and departure times. For indirect alternatives, the transfer time must be within a user-defined acceptable range, $[\gamma_{trans}^{min}, \gamma_{trans}^{max}]$, satisfying the minimum transfer time – to ensure a sufficient time between train arrival and the next train departure, and a maximum transfer time – to avoid excessively long transfer times. In addition, indirect alternatives that induce a detour that exceeds a user-defined ratio of γ_{detour}^{max} are removed from the choice-set as well as alternatives that are dominated by other alternatives when considering the number of transfers, in-vehicle time, transfer time and service level (i.e. intercity vs. regional). For each of the alternatives obtained in this process, the following attributes are stored along with the route itinerary: total travel time, number of transfers, total in-vehicle time and transfer time. These attributes are then used in the following choice step.

3.3.2. Passenger route choices

A multinomial logit (MNL) choice model is applied for route choice, to calculate the share of passengers travelling along each route alternative, shown as

$$P_{ijk} = \frac{exp(-\theta \tilde{t}_{ijk})}{\sum_{k \in R_{ij}} exp(-\theta \tilde{t}_{ijk})} \tag{11}$$

P_{ijk} is the share of passengers that choose route k when travelling between i to j and R_{ij} is the set of alternative routes from i to j. Route k consists of an ordered set of legs denoted by a sequence of stations, $k = (s_{k,1}, s_{k,2}, ..., s_{k,|k|})$ and $s_{k,m} \in S$ where S is the set of stations in the network, $\tilde{t}_{ijk}$ is the total generalized cost of route k for a given OD pair (ij), and θ is the logit scale factor for route choice.

3.3.3. Network load and passenger travel experience

The passenger travel experience is measured by the generalized travel time that consists of waiting time, in-vehicle time, transfer time and other fixed penalties as

$$\tilde{t}_{ijk} = \beta^{w} \cdot t_{s_{k,1}}^{w} + \sum_{v=1}^{|k|-1} \beta^{in} \cdot t_{s_{k,v}, s_{k,v+1}}^{in} + \sum_{q=2}^{|k|-1} \beta^{tr-time} \cdot t_{k,q}^{tr} + \beta^{tr} \cdot N^{tr} + \beta^{re} \cdot N^{re} \tag{12}$$

t_{s}^{w}, t_{s_1,s_2}^{in} and t_{s}^{tr} are the initial waiting time, in-vehicle time and transfer time, respectively, and β^{w}, β^{in} and $\beta^{tr-time}$ are the corresponding weights. $N^{tr} = |k| - 1$ and N^{re} stand for the number of transfer and rerouting decisions, and β^{tr} and β^{re} are penalty terms for each transfer and rerouting respectively, represented in time equivalent units. The well-known "independence from irrelevant alternatives" (IIA) property of the MNL model is partially counteracted by the filtering rules which result in a choice-set comprising of distinct alternatives, where the most correlated paths are either eliminated due to dominance rules or merged into hyper-paths as described in Cats et al. (2016). By calculating the choice model probabilities, the dynamic passenger assignment on each diachronic time-dependent network link can be calculated by

$$f_{ijk} = d_{ij} \cdot P_{ijk} \tag{13}$$

where d_{ij} is the number of passengers from i to j. The actual nominal travel time can be obtained as

$$t_{ijk} = t_{s_{k,1}}^{w} + \sum_{v=1}^{|k|-1} t_{s_{k,v}, s_{k,v+1}}^{in} + \sum_{q=2}^{|k|-1} t_{k,q}^{tr} \tag{14}$$

Analysing the impacts in terms of changes in total generalized travel time allow the summation of several passenger-related costs (e.g. transfer penalties, delays) using a compensatory function that is equivalent to total passenger welfare loss (Cats and Jenelius (2014)). Furthermore, using passenger value of time, impacts can be monetarized to assess the value of total time losses.".

3.4. Interaction between the models

The three models interact in a dynamic fashion, i.e., interaction occurs every time new information becomes available. New information can be input concerning the observed influencing factors in the disruption length model or when we learn that the previous disruption length prediction was too short. Fig. 9 shows an example of a disruption.

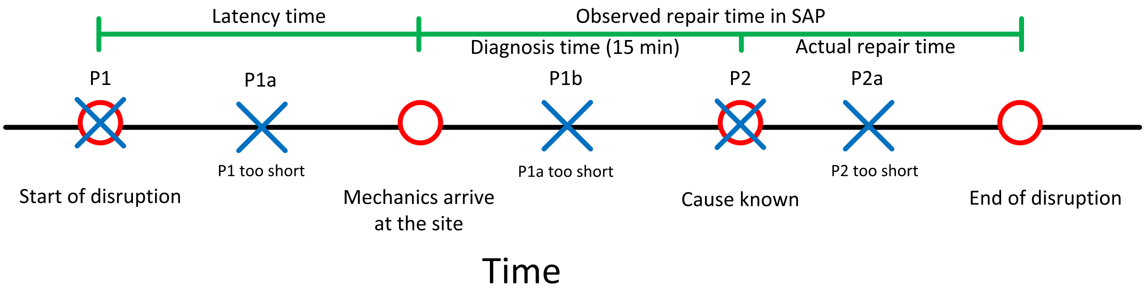


Fig. 9. The time diagram of a railway disruption.

The crosses in this time diagram illustrate when during the disruption period the interaction might take place. When a disruption occurs, only the influencing factors of the latency time are typically known. The unconditional Bayesian network is conditionalized based on this information. Predictions made from this conditional Bayesian network model are called the P1 predictions where, mainly, the latency time is predicted using the additional sources of information, potentially yielding more accurate predictions. P1 suggests that the disruption ends within a certain time period marked by P1a. The start and end time of the disruption period that are predicted by P1 are communicated to the short-turning model. Based on the disruption location, the relevant timetable and the disruption period, a disruption timetable is computed and passed on to the dynamic passenger assignment model.

When the predicted length P1 has elapsed, it might be realized that the disruption is not over yet. If the prediction is too short, the disruption is still unresolved even after the predicted disruption ends. This situation occurs when the prediction is too "optimistic", i.e. the chosen quantile of the conditional distribution of disruption length is too low for the case. If this happens, the prediction is updated by approximating a new conditional distribution of disruption length on the information that the disruption length is longer than the last prediction. This is done via sampling the original conditional distribution on the quantiles higher than the prediction. In this paper, these "revised" predictions are denoted alphabetically in orderly fashion. For instance, a P1 prediction is updated to P1a, P1b, and so on. With each prediction update the cycle shown in Fig. 1 repeats and the mentioned statistics are computed and stored.

Fifteen minutes after the arrival of the mechanics to the disruption site, they report the diagnosis about the cause of disruption. Knowing the cause of disruption, the Bayesian network is further conditionalized. The new conditional Bayesian network is used to produce the P2 predictions. Similarly with each prediction update, the short-turning model computes the disruption timetable and the passenger assignment model computes and stores the total number of affected passengers, generalized travel time, and total number of reroutings and transfers.

To measure and compare the impact of alternative disruption management scenarios the following key performance indicators (KPIs) are computed and stored:

1. The total number of passengers being affected during the disruption period.
2. The total experienced generalized travel time corresponding to equation (12) of all passengers considered in the experiment.
3. The total number of passenger reroutings and transfers.

4. Experiments

The framework that is applied on a part of the Dutch railway network is depicted in Fig. 10. The disruption length model is constructed using the computationally-efficient software called UNINET, which was developed at Delft University of Technology and is available at www.lighttwist.net/wp/uninet. The short-turning model is implemented in MATLAB, YALMIP (Löfberg (2012)) is used to construct the MILP, and Gurobi is used as the solver. The dynamic passenger assignment model is constructed in MATLAB. In the experiment, we consider a complete blockage in the railway segment between stations Utrecht and Houten. The blockage is caused by a track circuit failure.

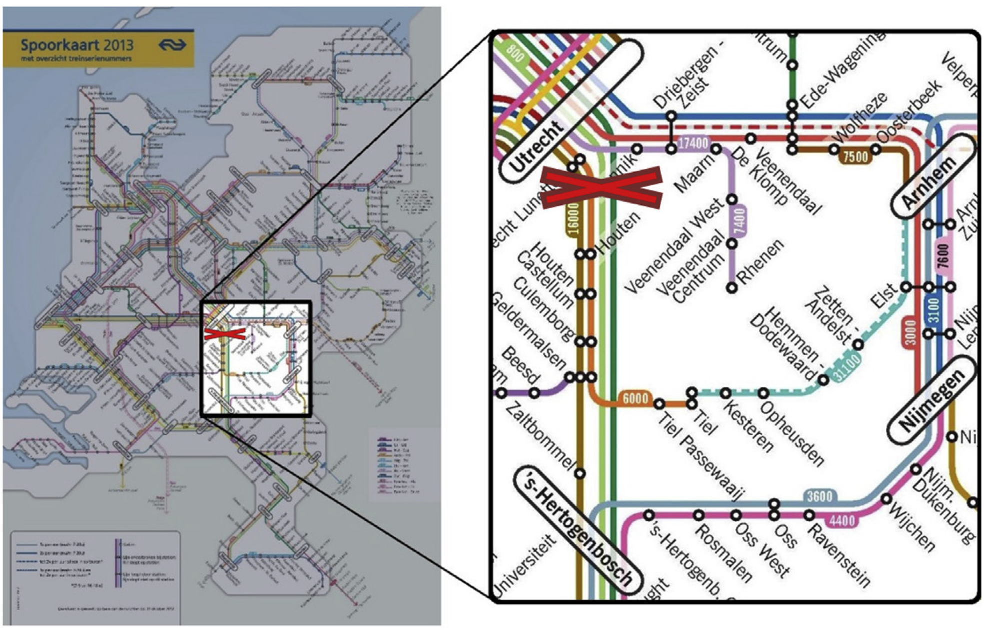


Fig. 10. The area the dynamic passenger assignment model considers in the experiment. The illustration is adapted from NS′ main train service map. Source: The Dutch railways (NS).

Two local train lines are considered by the short-turning model: line 16000 between Utrecht (Ut) and 's Hertogenbosch (Ht) and line 6000 between Ut and Tiel (Tl). Due to the disruption, these trains have the possibility to short-turn either at station Geldermalsen (Gdm) or the latest at station Houten (Htn).

To study the effect of different choices of quantiles of the conditional distribution of predicted disruption length, different disruption length predictions are examined. For each prediction, when a successive prediction is made, the new prediction is chosen to be the value of the new conditional distribution corresponding to the same quantile as in the previous prediction. For instance, when a prediction using the quantile 50% (median) is updated, the new prediction is also taken to be the median of the updated conditional distribution. In principle, this does not necessarily have to be the case. This choice is made to narrow down the space of possible "combinations" of prediction scenarios.

Note that the P1 predictions are only valid until new information regarding the cause is available from which the updated P2 predictions can be made. This means that the P1 predictions and the entailed disruption timetables are only used until at most 15 min after the mechanics arrive at the disrupted site.

The P2 predictions are updated until the actual disruption ends. When this happens, no new timetable is computed and the last P2 timetable is run until the P2 predicted end of disruption. This choice is made to penalize a prediction that is too "pessimistic", i.e., a prediction that surpasses the realized disruption length. In contrast, choosing a low quantile is undesirable because it is likely to be too optimistic and thus results in many "revised" predictions. This is not attractive from the passengers' point of view, who will presumably perceive the communicated information as unreliable. Additionally, having many "revised" predictions is not practical from a logistical point of view. In practice, every time a new prediction is made, aside from the train traffic, the traffic controllers must also revise the rolling stock and crew assignments. Therefore, train operators are not inclined to choose the lower quantile predictions. Therefore, in this experiment, we only consider the 25% quantile as the representative of these scenarios for comparison. The remaining quantiles that are considered in the experiments are 50%, 75%, 85%, 90%, as well as the mean.

The short-turning model computes the disruption timetable for lines 16000 and 6000 through stations Ut, Htn, Houten Castellum (Htnc), Culemborg (Cl), Gdm, Tiel Passewaaij (Tpsw), Tl, Zaltbommel (Zbm) and Ht. It is assumed that the new schedule is immediately available and communicated to the passengers (e.g. journey planner information is populated with the new timetable). For the rescheduling, the timetable of 2016 is considered. In the short-turning model, the main parameters are the penalties for arrival, departure delays and cancelled services. Since the interval of the services in Houten is either 16 or 14 min in both directions, and with each cancelled service the travellers need to wait around a quarter of an hour for the next train, the cancellation penalty ($\omega^{c}_{v^i_{l,n}}$) is set to 1000 s. Arrival and departure delays ($\omega^{d^a}_{v^i_{l,n}}$ and $\omega^{d^d}_{v^i_{l,n}}$) are equally penalized by 1. Moreover the minimum short-turning time is assumed to be 7 min (420 s). For the choice of minimum short-turning time we refer to the study by Chu and Oetting (2013). A norm of 3 min is considered for the minimum headway.

Given a disruption timetable, the dynamic passenger assignment model computes the passenger traffic. Due to data availability limitations, the dynamic passenger assignment model considers in this case study only the passenger trips for which both origin and destination are within the case study area, i.e. the loop shown in Fig. 10. The route choice model parameters are specified as follows: $\beta^{w} = 2$, $\beta^{in} = 1$, $\beta^{tr-time} = 2$, $\beta^{tr} = 5$, $\beta^{re} = 10$. For more details on the specification of the travel cost weights, see Cats et al. (2016).

The total demand in the case study network is considered inelastic, assuming that temporary service reductions do not have implications on other trip decisions such as mode and destination choices, which may or may not be feasible responses in a real situation. Empirical evidence from unplanned network disruptions shows that people are relatively reluctant to change travel mode (see Zhu et al. (2010) and references therein).

Passengers travelling between Utrecht Centraal and Houten stations have two alternatives: to detour via Arnhem and Nijmegen or to take the public bus service between the two stations. The travel time by bus between Utrecht Centraal station and Houten is about 35 min while a regular train service would have taken only 9 min. On the other hand, the detour via Arnhem and Nijmegen is also not very attractive due to the tremendous detour it induces. For passengers travelling to Houten from Utrecht, this detour takes almost 2 h. To fairly compare the different choices of predictions, we monitor the train traffic and passengers flow for a fixed period of six hours in all scenarios. Two actual disruption cases are chosen and examined. As already mentioned, different variables can impact the disruption length. Here we are particularly interested in examining the impact of time. Thus the consequence of the same disruption occurring on a different day or time-of-day is investigated. Since considering different parameters would have hamper the comparison and the ability to perform a comparison and draw conclusions, the same parameters were used in all cases. Due to the long computation time needed for the dynamic passenger assignment distribution model we examine only four case studies; weekday afternoon, weekday evening, Saturday evening and Saturday afternoon which are explained in detail in the following sections.

4.1. Case study 1: weekday afternoon

The first case study is based on an incident which occurred on Thursday, 10 July 2014. The incident started at 14:22 and had the information listed in Table 2. Moreover, the real observed latency and repair time are 70 and 73 min, respectively. This means the total disruption length is 70 + 73 = 143 minutes.

Table 3 presents the P1 predictions which are presented in terms of the length (in minutes) and the predicted end of disruption in time. Notice that in the scenario corresponding to the 25% quantile, in total there are 8 predictions that are generated throughout the disruption. The predictions in Table 3 are used by the short-turning model to produce the disruption timetable whose cyclic characteristics are shown in Table 4. This Table contains the number of cancelled services ($\#C$), the number of delayed services ($\#D$), the number of short-turned services ($\#ST$), and the total train delay in minute (Del).

Table 2
The initial information of the disruption case 1.

Contract type	OPC
Working station distance	7.1620
Level Crossing distance	872.372
Working Time	yes
Warm	yes
Rush Hour	no
Overlapping disruption	no
Cause	a setting problem caused by heat

Table 3
The P1 predictions for Case Study 1.

Qtl (%)	P1		P1a		P1b	
	Length	Time	Length	Time	Length	Time
25	49	15:11	72	15:34	97	15:59
50	81	15:43	144	16:46		
75	143	16:45				
85	205	17:47				
90	254	18:36				
Mean	118	16:20				

Table 4
The results of the short-turning model for P1 in Case Study 1.

Qtl (%)	P1				P1a				P1b			
	#C	#D	#ST	Del.	#C	#D	#ST	Del.	#C	#D	#ST	Del.
25	12	2	3	6	20	0	5	0	28	0	7	0
50	20	9	5	24	40	0	10	0				
75	40	0	10	0								
85	56	0	14	0								
90	68	0	17	0								
Mean	32	0	8	0								

Table 5
The P2 predictions for Case Study 1.

Qtl (%)	P2		P2a		P2b		P2c		P2d	
	Length	Time	Length	Time	Length	Time	Length	Time	Length	Time
25	85	15:47	102	16:04	118	16:20	135	16:37	150	16:52
50	95	15:57	134	16:36	179	17:21				
75	133	16:35	240	18:22						
85	160	17:02								
90	194	17:36								
Mean	119	16:21	188	17:30						

Table 6
The results of the short-turning model for P2 in Case Study 1.

Qtl (%)	P2				P2a				P2b			
	#C	#D	#ST	Del.	#C	#D	#ST	Del.	#C	#D	#ST	Del.
25	24	0	6	0	28	0	7	0	32	0	8	0
50	24	9	6	24	36	0	9	0	48	0	12	0
75	36	0	9	0	64	0	16	0				
85	44	0	11	0								
90	52	0	13	0								
Mean	32	0	8	0	52	0	13	0				

Table 7
The remaining results of the short-turning model for P2 for Case Study 1.

Qtl (%)	P2c				P2d			
	#C	#D	#ST	Del.	#C	#D	#ST	Del.
25	36	0	9	0	40	0	10	0

Table 8
The impact of different predictions to the passengers in Case Study 1.

Qtl (%)	Excess (minute)	# Affected Passengers	Inc. Orig (%)	Inc. Bench (%)	# Rerouting		# Transfers
					1	2	
25	7	8948	17.71	0.26	595	4	3597
50	36	11469	20.53	2.67	466	1	4015
75	97	16560	24.33	5.90	247	0	4671
85	17	9791	17.56	0.13	5	0	3775
90	51	12854	20.16	2.35	0	0	4126
Mean	45	12299	20.33	2.49	282	0	4136
Real	0	8374	17.4026	0	0	0	3773

At 15:32, the mechanics arrive at the site. After 15 min of diagnosis time, the cause of a track circuit failure is identified and at 15:47 the P1 predictions are updated to the P2 predictions. These P2 predictions are presented in Table 5. The results of the short-turning model are shown in Tables 6 and 7.

Each timetable is used when the prediction is still "valid", i.e., it has not changed. For instance, the disruption timetable generated with the P1a prediction of the 50% quantile is used only for four minutes between 15:43 and 15:47 (15 min after their arrival at the disrupted location). At 15:47, the prediction is updated to P2 and a new disruption timetable is constructed.

These disruption timetables are used by the dynamic passenger assignment model to compute passenger route choice and the resulting passenger distribution over train services for six hours between 14:22 and 20:22. In this period, there are 22163 passengers who are travelling in the case study area. We measure the impact on the passengers for each choice of quantile. The results are presented in Table 8.

The last row of Table 8 shows the benchmark case with the true disruption length. In this case, the true end time of disruption is already known when the disruption starts at 14:22[1]. Each scenario is compared to this case to measure the increase in the impact of the prediction on the passengers with respect to the ideal situation.

The second column of Table 8 shows the difference (in minute) between the true end of disruption and the last P2 prediction when the blocked railway section between Utrecht and Houten is opened for train operation. The third column presents the total number of passengers travelling during the blockage of the section. The fourth and fifth column provide the increase (in %) in the total generalized travel time with respect to the normal situation without disruption and the benchmark, respectively. In the sixth and seventh column, the total number of passengers who have to reroute once or twice in each scenario is provided. The number of transfers performed by the passengers can be found in the last column.

The benchmark case represents the best possible situation. In this case, fewer passengers are affected and no passengers have to be rerouted since the initially provisioned information is accurate. Unsurprisingly, the increase in the generalized travel time with respect to the no disruption situation is also the lowest.

In general, the longer the difference between the true end of disruption and the last P2 prediction is, the more passengers are affected. However, this does not necessarily translate to a higher total generalized travel time. Notice that the increase in the generalized travel time is higher in the 25%-quantile scenario than in the 85%-quantile scenario even though the difference between the prediction and the realized value is only 7 min in the former and 17 min in the latter. The eight predictions in the 25%-quantile scenario cause many passengers to reroute due to the frequent updates of the disruption timetable. Consequently, the total generalized travel time is penalized severely. In the 85%-quantile scenario, much fewer passengers need to reroute due to the pessimistic prediction.

Notice that the P2 predicted end time of the disruption of the 75%-quantile scenario (at 16:35, see Table 5) is 10 min shorter than the actual end time. Because of this slightly too optimistic prediction, the predicted end time is updated to P2a which is at 18:22. This new prediction is dramatically too long and consequently disrupts the late afternoon peak demand period. As a result, this scenario is the worst as indicated by the number of affected passengers and the increase of the total generalized travel time.

[1] This is the best possible situation but, of course, is not realistic.

Table 9
The P1 predictions for Case Study 1A.

Qtl (%)	P1		P1a		P1b		P1c	
	Length	Time	Length	Time	Length	Time	Length	Time
25	54	20:18	77	20:41	102	21:06	127	21:31
50	86	20:50	149	21:53				
75	149	21:53						
85	209	22:53						
90	258	23:42						
Mean	122	21:26						

Table 10
The results of the short-turning model for P1 in Case Study 1A.

Qtl (%)	P1				P1a				P1b				P1c			
	#C	#D	#ST	Del.	#C	#D	#ST	Del.	#C	#D	#ST	Del.	#C	#D	#ST	Del.
25	19	0	3	0	23	2	4	6	31	0	6	0	39	0	8	0
50	27	0	5	0	43	2	9	2								
75	43	2	9	2												
85	55	5	12	39												
90	73	2	15	8												
Mean	35	9	7	15												

Table 11
The P2 predictions for Case Study 1A.

Qtl (%)	P2		P2a		P2b		P2c		P2d	
	Length	Time	Length	Time	Length	Time	Length	Time	Length	Time
25	104	21:08	121	21:25	137	21:41	154	21:58	169	22:13
50	114	21:18	153	21:57	198	22:42				
75	152	21:56	259	23:43						
85	179	22:23								
90	213	22:57								
Mean	138	21:42	207	22:51						

Table 12
The results of the short-turning model for P2 in Case Study 1A.

Qtl (%)	P2				P2a				P2b			
	#C	#D	#ST	Del.	#C	#D	#ST	Del.	#C	#D	#ST	Del.
25	31	0	6	0	35	2	7	6	39	2	8	6
50	35	0	7	0	43	9	9	24	55	2	12	8
75	43	9	9	15	73	2	15	10				
85	47	4	10	38								
90	55	12	12	69								
Mean	39	9	8	15	55	5	12	31				
Real	47	0	10	0								

4.2. *Case study 1A: weekday evening*

In order to investigate the effect of the disruption's time of occurrence on the choice of predictions, an artificial disruption is considered in this case study. The exact same disruption as in Case Study 1 is assumed to occur later on the same day, at 19:24. The realizations of the latency and repair time of this artificial disruption are taken from the values of the computed conditional distribution of latency and repair time which correspond to the same quantiles as the realizations of Case Study 1. In this case, the latency and repair time are 89 and 73 min, respectively. The total length is 162 min and the disruption ends at 22:06.

The P1 predictions and the short-turning results are presented in Tables 9 and 10.

The P2 predictions are made at 21:08, 15 min after the mechanics' actual arrival time at 20:53. These predictions are presented in Table 11. The corresponding results of the short-turning model are represented in Tables 12 and 13.

Table 13
The remaining results of the short-turning model for P2 in Case Study 1A.

Qtl (%)	P2c				P2d			
	#C	#D	#ST	Del.	#C	#D	#ST	Del.
25	43	9	9	33	47	2	10	10

As before, the predictions are used by the short-turning model to produce the disruption timetables which are used by the dynamic passenger assignment model to attain the distribution of passengers traffic. Between 19:24 and 01:24, 7102 passengers are travelling in the case study area. Notice that there are fewer passengers in this set-up than the previous one due to the different time of the day under consideration. Table 14 summarizes the impact on passengers for each quantile.

In comparison to Case Study 1, the increase in the total generalized travel time with respect to the normal situation is higher. This is because the disruption is longer than in the previous case study due to the longer latency time.

The benchmark case still represents the best situation with fewer passengers being affected. The total generalized travel time is the lowest in this scenario and no passengers are rerouted.

The longer the last P2 prediction is, the more passengers are affected by the disruption. For this reason, the 75%-quantile scenario yields the highest total generalized travel time. Notice that as in Case Study 1, the P2 prediction of this scenario is 10 min shorter than the actual end time of the disruption. Consequently, the prediction is updated to P2a, which is then too long.

The nine predictions in the 25% quantile scenario cause many passengers to reroute. In this scenario, the simulation shows that there are a total of 1487 rerouting activities, out of which 438 had to reroute once and 195 had to reroute twice in addition to a considerable number of passengers who needed to reroute more than twice. As a result, the total generalized travel time of this scenario is the second largest due to the heavy penalty associated with rerouting.

4.3. Case study 2: Saturday evening

In this case study, we consider another real track circuit disruption at the same location which occurred on Saturday, 18 October 2014 and started at 19:24. The incident had the information shown in Table 15. The observed latency and repair time are 47 and 88 min, respectively, with a total length of 135 min.

The P1 predictions and the short-turning results for this case study are presented in Tables 16 and 17. Fifteen minutes after the mechanics' actual arrival time at 20:11, the P2 predictions are made. Tables 18 and 19 presents these predictions and the short-turning model. Table 20 summarizes the impact on passengers for each choice of quantile.

The benchmark case represents the best possible situation. With the least number of affected passengers with no need for re-routing, the total generalized travel time is the lowest of all scenarios.

Notice that the final predicted end time of disruption of the 50%-quantile and 75%-quantile scenario are the same. Consequently, the same number of passengers are affected in both cases. However, the 50%-quantile scenario has three predictions while the 75%-quantile scenario has only two. Consequently, many passengers need to be rerouted and more transfers need to be performed in the former. This results with higher total generalized travel time in the case of the 50%-quantile scenario. The pessimistic 90%-quantile scenario disturbs the most number of passengers because of a P2 prediction that is too long. Consequently, the total generalized travel time is the largest, making this scenario the worst performing one in this case study.

4.4. Case study 2A: Saturday afternoon

Similarly to Case Study 1A, the incident in Case Study 2 is also considered to occur at a different time of the day. In this case study, we assume this hypothetical incident to occur on the same day at 14:22. In this case, because the incident occurs during the weekend, the prediction lengths do not change from Case Study 2; only the time-dependent passenger demand generation process is adjusted.

Table 14
The impact of different predictions to the passengers in Case Study 1A.

Qtl (%)	Excess	# Affected	Inc. Orig	Inc. Bench	# Rerouting		# Transfers
	(minute)	Passengers	(%)	(%)	1	2	
25	7	4966	32.68	7.47	438	195	1353
50	36	5563	31.09	6.18	479	8	1430
75	97	6563	35.12	9.45	83	0	1512
85	17	5180	29.46	4.86	52	0	1379
90	51	5843	32.15	7.05	0	0	1461
Mean	45	5733	27.16	3.00	91	0	1306
Real	0	4812	23.4516	0	0	0	1291

Table 15
The initial information of the disruption case 2.

Contract type	OPC
Working station distance	7.1620
Level Crossing distance	872.372
Working Time	no
Warm	no
Rush Hour	no
Overlapping disruption	no
Cause	a cable problem

Table 16
The P1 predictions for Case Study 2.

Qtl (%)	P1		P1a	
	Length	Time	Length	Time
25	54	20:18	77	20:41
50	86	20:50		
75	149	21:53		
85	209	22:53		
90	258	23:42		
Mean	122	21:26		

Table 17
The results of short-turning model for P1 for Case Study 2.

Qtl (%)	P1				P1a			
	#C	#D	#ST	Del.	#C	#D	#ST	Del.
25	19	0	3	0	23	2	4	6
50	27	0	5	0				
75	43	2	9	2				
85	55	5	12	39				
90	73	2	15	8				
Mean	35	9	7	15				

Table 18
The P2 predictions for Case Study 2.

Qtl (%)	P2		P2a		P2b		P2c	
	Length	Time	Length	Time	Length	Time	Length	Time
25	67	20:31	94	20:58	120	21:24	151	21:55
50	104	21:08	173	22:17				
75	173	22:17						
85	237	23:21						
90	280	00:04						
Mean	142	21:46						

Table 19
The results of short-turning model for P2 for Case Study 2.

Qtl (%)	P2				P2a				P2b				P2c			
	#C	#D	#ST	Del.	#C	#D	#ST	Del.	#C	#D	#ST	Del.	#C	#D	#ST	Del.
25	23	0	4	0	27	9	5	33	35	2	7	4	43	2	9	6
50	31	0	6	0	47	2	10	18								
75	47	2	10	18												
85	65	7	13	102												
90	77	8	16	79												
Mean	43	0	9	0												
Real	39	2	8	2												

Table 20
The impact of different predictions to the passengers in Case Study 2.

Qtl	Excess	# Affected	Inc. Orig	Inc. Bench	# Rerouting		# Transfer
(%)	(minute)	Passengers	(%)	(%)	1	2	
25	16	4562	14.43	0.64	426	32	1219
50	38	5052	17.11	2.99	252	32	1506
75	38	5052	16.11	2.11	0	0	1437
85	102	6246	19.30	4.92	0	0	1558
90	145	6792	21.43	6.79	0	0	1693
Mean	7	4351	14.89	1.04	2	0	1385
Real	0	4182	13.7099	0	0	0	1309

Table 21
The P1 predictions for Case Study 2A.

Qtl (%)	P1		P1a	
	Length	Time	Length	Time
25	54	15:16	77	15:39
50	86	15:48		
75	149	16:51		
85	209	17:51		
90	258	18:40		
Mean	122	16:24		

Table 22
The results of the short-turning model for P1 in Case Study 2A.

Qtl (%)	P1				P1a			
	#C	#D	#ST	Del.	#C	#D	#ST	Del.
25	16	0	4	0	20	2	5	2
50	24	0	6	0				
75	40	0	10	0				
85	56	0	14	0				
90	68	2	17	4				
Mean	32	2	8	4				

Table 23
The P2 predictions for Case Study 2A.

Qtl (%)	P2		P2a		P2b		P2c	
	Length	Time	Length	Time	Length	Time	Length	Time
25	67	15:29	94	15:56	120	16:22	151	16:53
50	104	16:06	173	17:15				
75	173	17:15						
85	237	18:19						
90	280	19:02						
Mean	142	16:44						

Tables 21 and 22 present the P1 predictions and the corresponding short-turning results. The P2 predictions and the short-turning results are presented in Tables 23 and 24.

The impact of different scenarios on the passengers is presented in Table 25. Notice that more passengers are affected by the disruption in comparison to Case Study 2. This is due to the disruption occurring during the day time when more passengers are travelling.

Like the previous three case studies, the scenario with the true disruption length is the best performing one in terms of the total generalized travel time. The least number of passengers are affected and none of them have to change their travel plans.

As in Case Study 2, the difference between the predicted end of disruption and the truth is 38 min in both the 50%-quantile and the 75%-quantile scenario. However, the increase in the generalized travel time is higher in the former case. This is due to the more frequent prediction updates so many passengers have to reroute and slightly more transfers are needed. Consequently, the total

Table 24
The results of short-turning model for P2 for Case Study 2A.

Qtl (%)	P2				P2a				P2b				P2c			
	#C	#D	#ST	Del.	#C	#D	#ST	Del.	#C	#D	#ST	Del.	#C	#D	#ST	Del.
25	20	0	5	0	24	9	6	15	32	0	8	0	40	2	10	2
50	28	0	7	0	48	0	12	0								
75	48	0	12	0												
85	64	0	16	0												
90	76	0	19	0												
Mean	36	9	9	33												
Real	36	0	9	0												

Table 25
The impact of different predictions to the passengers in Case Study 2A.

Qtl (%)	Excess	# Affected	Inc. Orig	Inc. Bench	# Rerouting		# Transfers
	(minute)	Passengers	(%)	(%)	1	2	
25	16	9031	12.31	1.22	723	120	3249
50	38	10928	15.87	4.43	491	46	3725
75	38	10928	14.60	3.28	0	0	3703
85	102	16359	19.95	8.11	0	0	4549
90	145	19044	23.25	11.08	0	0	5069
Mean	7	8294	11.10	0.13	4	0	3128
Real	0	7736	10.9567	0	0	0	3128

generalized travel time of this scenario is penalized more.

The very pessimistic 90% quantile scenario performs the worst in terms of the total generalized travel time. The dramatic difference between the predicted end of disruption and the truth means a lot of passengers are affected by the disruption. As a result, the total generalized travel time becomes very high.

5. Conclusions and future work

In railway operation, different disruption length predictions can lead to different consequences for the passengers. In this paper a framework is developed to investigate the impacts of disruption length prediction on the total passengers generalized travel time.

This framework integrates three models; a disruption length model using a Baysian network approach, a short-turning model formulated as a MILP, and a dynamic passenger assignment model using a multi-stage passenger load approach. The integration of these models has a unique iterative approach starting with a disruption length prediction, then rescheduling trains by short-turning them based on the predicted length, and finally measuring the passengers generalized travel time. The new iteration takes place upon receiving an update that leads to a new disruption prediction length, which results in an updated schedule and finally the new passenger generalized travel time is computed. These iterations continue until the disruption ends. In the dynamic passenger assignment model, the generalized travel time consists of the weighted total travel time of all passengers which takes into the account the waiting time, the in-vehicle time, the transfer time, the number of transfers, and the number of reroutings.

In a series of case studies, we have shown how different choices of predictions of disruption length will affect the passengers in terms of the generalized travel time. Our finding is that on one hand, when the prediction is too optimistic, many passengers have to be rerouted which increases the inconvenience and, hence, the total generalized travel cost. On the other hand, when the prediction is too pessimistic, more passengers are affected which results in a higher total generalized travel cost. The result from the case study corresponds to the current prediction practice which is using a pessimistic prediction that reduces the chance of further changes in the plan. When the prediction is just slightly shorter than the realized value, more passengers are affected. This is evident especially with pessimistic predictions which lead to keeping the track blocked much longer than necessary, significantly beyond the end of the disruption. The experiments provide thus insights on the impact of the predictions on passengers delays.

A hypothesis that can be investigated further is that "starting with pessimistic predictions and thereafter gradually switching to the optimistic predictions might result in a better overall prediction". We cannot conclude based on a few case studies which value of the conditional distribution of disruption length is the best for a prediction. To draw such a conclusion, many more case studies need to be performed. However, we show that different quantiles led to the smallest generalized travel time of passengers under different scenarios. The close-to-optimal solution can be found by testing a large number of points from the solution space, i.e., by generating many different realizations of disruption length from the conditional distribution. For each realization, the short-turning model and the dynamic passenger assignment model should then be run to compute the effect of different choices of disruption length predictions.

Moreover when a prediction had to be updated we took the same quantile of the conditional distributions. For future research direction the choice of quantile in the P2 prediction should be investigated to see whether there is a dependency on the quantile that realizes the latency time in the P1 prediction. The realization of the latency time might indicate how fast/slow the mechanics work on the specified incident which might be a useful information to produce a more accurate P2 prediction. Other possibilities could be considered as well.

The function of the total generalized travel time specified in this study can be extended in future studies. During peak demand periods and disruptions it is especially relevant to consider vehicle capacity constraints and the vehicle type. The bus service between Utrecht and Houten provides significantly less seat capacity than the train service. Moreover, different vehicle types provide different levels of comfort to the passengers. An Intercity train is generally more comfortable than a Sprinter train or a bus service. This can be accommodated in equation (12) by adding a weight corresponding to the vehicle type to the second term.

Choosing generalized travel time as the primary evaluation/selection criterion means the impact of the uncertainty in disruption length is only measured from the passengers' point of view. With this cost function, an optimistic prediction is not attractive only because it causes inconvenience to the passengers who would have to reroute. Note that we have not captured all costs associated with an optimistic prediction that the traffic control faces. However, from an operational point of view, having a lot of timetable updates is not practical due to the logistical issues that need to be carried out. For instance, the rolling stock and crew assignments need to be reorganized accordingly with every update. Future studies may incorporate additional aspects into the cost function in order to better reflect the aspects influencing real-time decision making in the context of disruption management.

Note that the effect of different predictions depends considerably on the location and the time of the incident. This is evident in the results of the case studies investigated in this paper. This is presumably even more pronounced in denser areas where there is a greater hierarchy among different locations, such as the Amsterdam area. Aside from the greater number of alternative rail services, passengers may switch to other modes of public transport, consisting of metro, tram and bus, which contribute to network redundancy.

Acknowledgement

This work is funded by ExploRail, a partnership programme of the Dutch Technology Foundation STW and ProRail, project no. 12257: "Smart information and decision support for railway operation control centres (SmartOCCR)". The authors would like to acknowledge Dirk Kes and André Duinmeijer from ProRail for the supervision and access to the data set.

References

Canca, D., Barrena, E., Laporte, G., Ortega, F.A., 2016. A short-turning policy for the management of demand disruptions in rapid transit systems. Ann. Oper. Res. 246 (1), 145–166.
Cats, O., 2016. The robustness value of public transport development plans. J. Transport Geogr. 51, 236–246.
Cats, O., Jenelius, E., 2014. Dynamic vulnerability analysis of public transport networks: mitigation effects of real-time information. Network. Spatial Econ. 14 (3), 435–463.
Cats, O., West, J., Eliasson, J., 2016. A dynamic stochastic model for evaluating congestion and crowding effects in transit systems. Transp. Res. Part B Methodol. 89, 43–57.
Chu, F., Oetting, A., 2013. Modeling capacity consumption considering disruption program characteristics and the transition phase to steady operations during disruptions. J. Rail Transp. Plan. Manag. 3, 54–67.
Coor, G.T., 1997. Analysis of the Short-turning Strategy on High-frequency transit lines. Tech. rep.. Massachusetts Institute of Technology.
De-Los-Santos, A., Laporte, G., Mesa, J.A., Perea, F., 2012. Evaluating passenger robustness in a rail transit network. Transport. Res. C Emerg. Technol. 20 (1), 34–46.
Ghaemi, N., Cats, O., Goverde, R.M.P., 2017a. A microscopic model for optimal train short-turnings during complete blockages. Transp. Res. Part B Methodol. 105, 423–437.
Ghaemi, N., Cats, O., Goverde, R.M.P., 2017b. Railway disruption management challenges and possible solution directions. Pub. Trans. 9 (1), 343–364.
Ghaemi, N., Cats, O., Goverde, R.M.P., 2018. Macroscopic multi-station short-turning model in case of complete railway blockages. Transport. Res. C Emerg. Technol accepted.
Ghaemi, N., Goverde, R.M.P., Cats, O., 2016. Railway disruption timetable: short-turnings in case of complete blockage. In: 2016 IEEE International Conference on Intelligent Rail Transportation (ICIRT), pp. 210–218.
Hirai, C., Kunimatsu, T., Tomii, N., Kondu, S., Takaba, M., 2009. A train stop deployment planning algorithm using a petri-net-based modelling approach. Q. Rep. RTRI 50 (1), 8–13.
Joe, H., 2014. Dependence Modelling with Copulas. CRC Press.
Kumazawa, K., Hara, K., Koseki, T., 2008. A novel train rescheduling algorithm for correcting disrupted train operations in a dense urban environment. Comput. Railw. XI, WIT Press 103, 565–574.
Löfberg, J., 2012. Automatic robust convex programming. Optim. Meth. Software 27 (1), 115–129.
Meng, L., Zhou, X., 2011. Robust single-track train dispatching model under a dynamic and stochastic environment: a scenario-based rolling horizon solution approach. Transport. Res. Part B 45 (7), 1080–1102.
Nakamura, T., Hirai, C., Nishioka, Y., 2011. A Practical Train Rescheduling Algorithm Using Three Predetermined Factors. Tech. rep.. Railway Technical Research Institute, West Japan Railway Company.
Nelsen, R.B., 2006. An Introduction to Copula. Springer, New York.
Nielsen, Lars Kjær, Kroon, Leo, Maróti, Gábor, 2012. A rolling horizon approach for disruption management of railway rolling stock. Eur. J. Oper. Res. 220 (2), 496–509. https://doi.org/10.1016/j.ejor.2012.01.037.
Takeuchi, Y., Tomii, N., 2005. Robustness indices for train rescheduling. In: Proceedings of the 1st International Seminar on Railway Operations Modeling and Analysis (RailDelft 2005), June, pp. 08–10.
Veelenturf, L.P., Kidd, M.P., Cacchiani, V., Kroon, L.G., Toth, P., 2016. A railway timetable rescheduling approach for handling large-scale disruptions. Transport. Sci. 50 (3), 841–862.
Yang, L., Zhou, X., Gao, Z., 2013. Rescheduling trains with scenario-based fuzzy recovery time representation on two-way double-track railways. Soft Comput. 17 (4), 605–616.
Yang, L., Zhou, X., Gao, Z., 2014. Credibility-based rescheduling model in a double-track railway network: a fuzzy reliable optimization approach. Omega 48, 75–93.
Zhan, S., Kroon, L.G., Veelenturf, L.P., Wagenaar, J.C., 2015. Real-time high-speed train rescheduling in case of a complete blockage. Transp. Res. Part B Methodol. 78, 182–201.
Zhan, S., Kroon, L.G., Zhao, J., Peng, Q., 2016. A rolling horizon approach to the high speed train rescheduling problem in case of a partial segment blockage. Transport. Res. E Logist. Transport. Rev. 95, 32–61.
Zhu, S., Levinson, D., Liu, H.X., Harder, K., 2010. The traffic and behavioral effects of the i-35w Mississippi river bridge collapse. Transport. Res. Pol. Pract. 44 (10), 771–784.
Zilko, A.A., Kurowicka, D., Goverde, R.M.P., 2016. Modeling railway disruption lengths with copula bayesian networks. Transport. Res. C Emerg. Technol. 68, 350–368.