Assessment of alternative line plans for severe winter conditions in the Netherlands

Trap, Maarten L.; Huisman, Dennis; Goverde, Rob M.P.

DOI
10.1007/s12469-016-0149-4

Publication date
2017

Document Version
Final published version

Published in
Public Transport

Citation (APA)
Trap, M. L., Huisman, D., & Goverde, R. M. P. (2017). Assessment of alternative line plans for severe winter conditions in the Netherlands. *Public Transport*, *9*(1-2), 55-71. https://doi.org/10.1007/s12469-016-0149-4

Public Transp (2017) 9:55–71
DOI 10.1007/s12469-016-0149-4

ORIGINAL PAPER

Assessment of alternative line plans for severe winter conditions in the Netherlands

Maarten L. Trap[1] · **Dennis Huisman** [1,2] · **Rob M. P. Goverde**[3]

Accepted: 22 November 2016

Abstract Winter weather has a major impact on railway operations in the Netherlands. To stay in control, the number of trains is reduced by 20% in a special "winter timetable". This results in a more controllable network, but an insufficient amount of transport capacity. Adapting the line plan can result in more transport capacity without losing controllability. This paper therefore focuses on the performance of a line plan under extreme weather conditions. We define several criteria to assess the performance of the line plan in terms of controllability and transport capacity. A case study has been conducted on the railway network in the Netherlands, which indicates that all alternatives are more controllable and yield more transport capacity than the current winter timetable.

Keywords Extreme weather · Line planning · Railway optimization · Controllability

1 Introduction

The winter of 2009/2010 was one of the most extreme winters in the Netherlands in decades. On several days, the railway operations got "out-of-control". In such a situation, the operational control organizations are no longer able to control the train traffic due to the many disruptions. The extreme weather circumstances resulted in

---

✉ Maarten L. Trap
maarten.trap@ns.nl

[1] Process quality and Innovation, Netherlands Railways, P.O. Box 2025, 3500 HA Utrecht, The Netherlands

[2] Econometric Institute and ECOPT, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

[3] Department of Transport and Planning, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, The Netherlands

broken trains and malfunctioning infrastructure, often at the same time and at multiple locations. Recovery from these disruptions is very difficult due to the intensive use of the Dutch railway infrastructure. The inter-dependencies between rolling stock and crew make that delays are easily propagating through the whole network. The normal approach to disruption management consists of rescheduling the timetable based on predefined contingency plans (Jespersen-Groth et al. 2009; Louwerse and Huisman 2014). Each contingency plan corresponds to a specific disruption scenario at a specific location like a fully or partially blocked track or station area. However, the actual situation always differs from a predefined disruption scenario so that traffic controllers have to adjust or combine plans to find a suitable solution which is an intensive task. Hence, multiple simultaneous disruptions quickly lead to out-of-control situations.

Since the first winter programme in 2010, different railway timetables have been deployed to increase the controllability and limit the impact of disruptions. The current alternative winter timetable in the Netherlands is the National Reduced Timetable (*Landelijke Uitgedunde Dienstregeling, LUD*). The LUD is largely based on the regular line plan with some mutations in line length and frequency. The LUD is therefore seen as a baseline, i.e. a *degraded* version of the original timetable. About 20% of the train trips is canceled throughout the day, effectively reducing the frequency of the train service nationwide to 2 trains/h. In the busy Randstad area this results in about 50% less trains. The LUD can be deployed within a relative short time frame, because it only requires mutations in the regular plan. As a result, the LUD has more or less the same pattern of arrival and departure times which makes that passengers and crew are easily familiarized with the alternative plan. As of this moment, the LUD has proven to be the most successful timetable regarding the controllability of the network. Consequently, the severity of winter-related problems has decreased and the preventive measures have converged to a stable solution.

Decreasing the frequency of train services results in extra allowances and less delay propagation in case of a disruption, but limits the transport capacity. Analysis in both theory and practice shows that the transport capacity of the LUD is not sufficient. Another alternative must be found to reduce crowding in the trains during extreme winter weather, while the controllability is conserved. Since the regular timetable is the foundation of the LUD, it might be useful to use a different foundation.

The main foundation of a timetable is the *line plan*. A more controllable timetable may thus be obtained by an alternative line plan, rather than just changing the frequencies of the existing lines. However, an alternative line plan requires a new detailed timetable, rolling stock assignment and crew planning. This paper focuses on the changes to the line plan and the implications for the assignment of rolling stock, as the latter greatly influences the transport capacity of the railway network. The objective of our study is to define criteria for a controllable line plan, and subsequently assess line plans for their controllability and transport capacity. This way, we aim for a line plan without overcrowded trains, while train controllers can appropriately respond to disruptions. To achieve this, multiple alternative line plans, along with a corresponding distribution of rolling stock, are evaluated to assess their controllability and transport capacity.

The scientific relevance of this study is bilateral. First, we use a different approach for *line planning* (i.e. creating a line plan). In practice, line planning choices are often made using an economical point of view. Minimizing costs is therefore a common recurring approach in literature (see e.g. Claessens et al. 1998; Goerigk et al. 2013; Schöbel 2012). Another common approach is to maximize the number of direct trips in the network (Bussieck 1998; Kaspi and Raviv 2013). For these approaches, line planning is often seen as a mathematical problem (the Line Planning Problem, LPP) with corresponding optimization models to create the line plan. In our study, the approach is to create a line plan where *controllability* and *transport capacity* are the key decision factors. We modified the traditional line planning approach to create such a specific line plan. Secondly, we evaluate both the controllability and transport capacity as being a result of the line plan. The assessment is made using the characteristics of the line plan itself, such that there is no need for a detailed timetable.

Section 2 describes how controllability is defined and evaluated with respect to the transport capacity. This method is subsequently applied to the railway network of the Netherlands Railways in Sect. 3. Section 4 gives conclusions.

2 Assessment of controllability and transport capacity

The goal of this study is to create a line plan and assess it for both its controllability and its transport capacity. Given the results, the line plan may be adapted to increase one or both aforementioned aspects. An iterative methodology has been used to do so, presented in Fig. 1. Each alternative line plan has its own underlying principle to initiate the planning process. The length and frequency of the lines determine the number of trains required to operate each line in the line plan. The composition of the trains depends on the number of passengers per train. This is calculated using an Origin–Destination Matrix (O–D Matrix) of the Dutch railway network. Using an allocation model, all passengers in the O–D Matrix are allocated to the trains resulting in a list of the travel demand per train composition. Based on the demand, the available rolling stock is assigned to the trains in order to calculate the transport capacity of the line plan. If the demand is larger than the capacity of a train, there is a capacity shortage. Adapting the line plan could result in less shortage, for instance by increasing the frequency. This is visualized by the feedback loop in the design process.

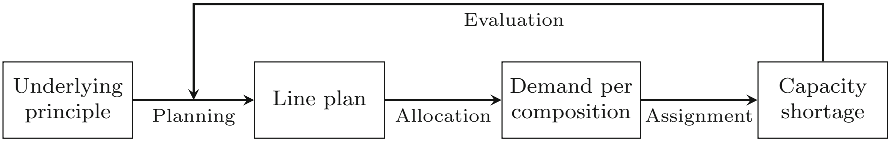

Fig. 1 Methodology to design an alternative line plan

There are many different possible line plans, all based on an underlying principle or objective. Many networks are optimized to reduce the costs and/or transfers, as is the regular NS railway network. Our methodology aims to create a line plan based on a controllable perspective.

The result of the line planning phase is a complete list of all lines in the line plan, consisting of their type (for instance InterCity or Regional train), frequency and the commercial stops per line. For all lines, the maximum train length can be determined by the stations the train serves. The shortest platform length of the served stations is the maximum train length.

2.1 Controllability of rail networks

Controllability is a common term in literature about rail networks, and is often used as synonym for *robustness, resilience and/or flexibility*. These terms are used indifferently and require further explanation. The general notion of a controllable system is the ability to adapt to a range of possible futures (Rosenhead 2013). In railway scheduling, the controllability is the ability to stick to the plan (i.e. the timetable) in case of a disturbance, as well as the ability to adapt the plan to reduce delays. Many authors therefore identify the difference between preventive and reactive methods to obtain a controllable plan (see e.g. Cacchiani et al. 2012; Klibi et al. 2010). Preventive methods are precautionary measures to absorb small disturbances while executing the plan. A typical way to do so is by adding time slacks in the timetable (Kroon et al. 2007). These time slacks can be created by adding buffers between trains (Andersson et al. 2013; Gestrelius et al. 2012), but also by adding time supplements to the running time of a train. Such a supplement will for instance allow trains to let a trip take longer than theoretically possible, such that a small delay can be compensated. Time supplements therefore enable the railway system to recover from a disturbance without intervention by a controller (Van Oort and Van Nes 2009a). These methods are all aiming to *prevent* the system from becoming disrupted and *limit* the impact of disturbances. Hence, these measures are planned in the timetable up front.

Reactive methods are *interventions* by agents, often a human train controller or controlling software, issuing measures to recover from larger disruptions. These measures range from small adjustments in the timetable with minimal effect on the train service, to large rescheduling with a major impact (Schaafsma 2001). This ability to control the network in case of a larger disruption is mostly referred to as resilience (Goverde and Hansen 2013). The degree of reactive controllability is depending on the capacity and the effectiveness of the controllers, but also on the complexity of the network itself. A less complex network can increase the controllability, as it reduces the time a controller needs to let the network recover from a disruption. The controllability is therefore depending on the line length, frequency and the density of the lines.

2.2 Assessment of controllability

Based on findings from literature in multiple fields and interviews with planners at NS, four criterion groups have been identified which define the controllability of a line plan. Most of these groups relate to the resilience of the railway network. Every group consists of one or multiple indicators, which enable measurement.

The first criterion is *line length*. The length of the lines in the line plan says something about the possible propagation of delays through the network (Van Oort and Van Nes 2009b). The shorter a line is, the less stops are being affected in case of a disruption. Longer lines result in more inter-dependencies between stations in the network, which makes that disruptions are more likely to propagate through the network. The impact of disruptions can be restrained by operating shorter lines. As indicator, we use the number of major stations a line serves. This is calculated by counting the number of major stations for every line, and taking the average. Shorter lines both increase the preventive and reactive aspect of controllability.

The second criterion is *traffic intensity*. A larger number of trains per edge per hour results in less time between trains and consequently less time to resolve problems before the next train arrives. If there are more lines sharing the same edge, a disruption on this edge would furthermore affect multiple lines. Traffic intensity is measured by determining the number of trains and lines on each edge. The following indicators are used:

- Average frequency per edge
- Number of edges with frequency >4
- Average line density per edge
- Number of edges with >2 lines

The values of these indicators are calculated by listing all edges on the main railway network and determining the number of lines that attend each edge. The average line density and the number of edges with more than two lines is directly derived from here. The frequency on every edge can be calculated in a similar way by taking the sum of the frequencies of the lines that serve the respective edge. A lower traffic intensity both increases the preventive aspect (less impact on a subsequent train) as the reactive aspect (less trains to control) of the controllability.

The third criterion is *control region attendance*. Railway lines often run through multiple control regions, i.e. the geographical regions where dispatchers are responsible for controlling the train traffic. In case of a disruption, these regions need to coordinate back and forth between other regions to control the problem. The less regions a train attends, the less regions are concerned with controlling the disruption. In the Netherlands the infrastructure manager is in charge of the train traffic, while the railway operator controls the rolling stock and the crew on the train. Therefore, both have their own geographical regions and both need to act in case of a disruption. Hence, we distinguish between traffic control regions of the infrastructure manager and the transport control regions of the railway operator. We look at the following indicators:

- Number of trains per transport control region
- Number of trains per traffic control region
- Average served transport control regions per line
- Average served traffic control regions per line

The values of these indicators are calculated by listing all lines and determining which transport and traffic control regions they serve. The average number of attended regions can be derived from here. The number of trains per region is calculated by taking the sum of the frequencies of all lines that serve the respective region. The number of attended control regions only affects the resilience.

The fourth and last criterion is *disruption risk*. Every train operation has a certain chance to cause an infrastructural disruption, but some have a higher probability than others. This could, for instance, be the crossing of a bridge, a special switch or a level crossing. In the Netherlands, some high-speed switches are notorious for their failure rate during winter weather. We therefore define two indicators regarding critical switches. These indicators illustrate the number of operational high-speed switches and the average number of switch movements per hour. The values are calculated by listing all edges with a high-speed switch and determining if traversing this edge triggers switch movement. If one switch is for instance controlling the junction between stations A, B and C as shown in Fig. 2, the switch is considered operational if both edges A–B and A–C are traversed. Switch operation is estimated using the frequency and assuming an equal pattern over the hour. If A–B is traversed once per hour and A–C twice/h, the assumed order during the hour is $\{A–B, A–C, A–C\}$. This implies two switch movements per hour. Locking these switches is a preventive measure.

Both the criterion group and the indicator have a weight related to their importance. These weights are discussed in Sect. 3. Once the values and weights of the indicators are known, a Multi-Criteria Analysis (MCA) is performed to determine the degree of controllability. We call this the *robustness index*, being a *weighted sum* of the above mentioned indicators of controllability. To make sure that all indicators are contributing on the same scale to the robustness index, the values of all indicators are standardized. The zero-alternative (A0) is used as a reference here, which means that the indicator values of each alternative are divided by the corresponding value of A0 and multiplied by 100 to create a new value that is relative to A0. For all indicator values holds: lower is better. This implies that a lower robustness index is preferred over a higher index as well, which makes that the robustness index of 100 is considered as the *upper bound*.

The weights are used to prioritize certain criteria and indicators over others, since not every aspect is of equal importance. Weights are determined using an Analytic Hierarchy Process (AHP), which makes it possible to systematically structure a

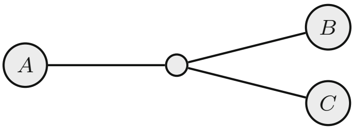

Fig. 2 Example of a junction where traversing both edges implies an operational switch

decision-making problem with multiple criteria (Saaty 1990). This is done by creating a hierarchy which divides the problem into different *levels*. The idea is to estimate how much more important one criterion is, compared to all other criteria. This yields a weight for all criteria, where the most important criterion gets the largest percentage. The sum of all weights is 100%.

2.3 Passenger allocation

In order to calculate the transport capacity of a line plan, the number of passengers per line and per train has to be determined. An O–D Matrix is used to estimate the travel demand D_{ij} from origin i to destination j. All passengers in the O–D Matrix must be allocated to the lines in the line plan. In this paper, the allocation of passengers is performed using an allocation model called TRANS (Warmerdam 2004), which determines the line(s) a passenger uses to travel from origin to destination. This is straightforward if there is only one possibility, but requires a discrete choice model once there are more travel options, especially when a transfer is required. First, TRANS allocates the passengers to the different lines in the line plan. Subsequently, the passengers are allocated to the trains on that line.

To allocate passengers to the lines, TRANS uses two phases. The first phase is the generation of all possible travel options. In similar studies, these options are also called itineraries. For every origin i to destination j (called O–D pair), TRANS generates a large set of possible travel options. Subsequently, TRANS determines which travel options are realistic by comparing two options with each other regarding travel time, transfers and frequency. The ticket price is not considered since it is assumed that a trip from i to j has the same price for all possible travel options. If one of the options is classified as "unrealistic", it is deleted from the set of options. This happens for instance if the difference in travel time between two options is greater than a certain threshold (20 min), while having the same number of transfers. This threshold and other parameters for the comparison have a default value based on research by NS. The result of the first phase is thus a set of travel options per O–D pair.

The second phase allocates the passengers to the travel options corresponding to the O–D pair using a discrete choice model. This is a mathematical function to predict the choice of a passenger based on the utility of the travel option (Akiva and Lerman 1985). The utility is calculated using travel time, the number of transfers and the time of transfers. The allocation is subsequently calculated using a Multinomial Logit (MNL) model based on utility maximization, included in TRANS (Dow and Endersby 2004). Normally, a stochastic error is added to the utility function to account for possible preferences that cannot be observed. Since TRANS does not account for this preference, it is assumed that travel options with the exact same utility will have an even amount of passengers. TRANS calculates the utility of the travel options for all O–D pairs in the O–D Matrix, such that the passenger load P^l_{ij} can be determined afterwards. This is the number of passengers on line l between stations i and j.

Next, the passengers are allocated to a specific train. The total number of trains per line is the quotient of the complete round-trip time and the scheduled headway.

TRANS distributes the passengers per line over the trains on that line, considering the time of day to account for peak hours. This yields the travel demand per train between all stations the train serves, hence the travel demand per *edge*. The *busiest edge* a train encounters is the edge with the largest demand. The train must at least have enough capacity to transport these passengers. TRANS calculates the busiest edges for all trains and also considers the time of the day. See Trap (2014) for details and an example.

The result of the allocation per train is a list of all trains required to operate the line plan, along with the maximum demand the train will encounter during the day. If all trains have enough capacity to at least accommodate this demand, there is sufficient transport capacity.

2.4 Rolling stock assignment

Once we know the required number of trains and their minimum capacity, the actual train units can be assigned to these trains. There is a fixed number of train units available which can be coupled to form a train *composition*, consisting of one or more train units of the same type. Each possible composition has its own length and capacity.

The assignment of compositions to the trains can be seen as an optimization problem with the objective to match the composition capacity with the number of passengers. In other words: the *shortage* of train capacity must be minimized. There is a shortage of capacity if not all passengers can be transported, for instance if the train is too short. The resulting shortage is expressed as the number of passengers that is unable to be transported in a decent way. An integer linear optimization model has been formulated to assign train compositions to every train on the network. This model is based on similar models presented by Abbink et al. (2004) and Fioole et al. (2006) and is adapted for the purpose of this study. Table 1 lists the sets, parameters and decision variables required for the model.

Table 1 Sets, parameters and decision variables for the rolling stock assignment model

Sets	W	Set of all trains
	C	Set of all possible train compositions
	S	Set of all train types
Parameters	D_w	Number of passengers on train w
	L_w	Maximum length of the composition for train w
	cap_c	Capacity of composition c
	l_c	Length of composition c
	$n_{c,s}$	Number of train units of type s in composition c
	N_s	Fleet size for train type s
Variables	z_w	Capacity shortage on train w
	$x_{w,c}$	$= \begin{cases} 1 & \text{if composition } c \text{ is assigned to train } w \\ 0 & \text{otherwise} \end{cases}$

The rolling stock assignment model can be formulated as follows:
Minimize:

$$Z = \sum_{w \in W} z_w \tag{1}$$

Subject to:

$$z_w \geq \sum_{c \in C} (D_w - cap_c) \cdot x_{w,c} \tag{2}$$

$$\sum_{c \in C} x_{w,c} = 1 \tag{3}$$

$$\sum_{c \in C} l_c \cdot x_{w,c} \leq L_w \tag{4}$$

$$\sum_{c \in C} \sum_{w \in W} n_{c,s} \cdot x_{w,c} \leq N_s \tag{5}$$

$$x_{w,c} = \{0, 1\} \tag{6}$$

$$z_w \geq 0 \tag{7}$$

The objective function (1) aims to minimize the total shortage of capacity over the complete network. Constraints (2) define the shortage per train if and only if the demand is larger than the capacity of the assigned composition. Constraints (3) ensure that every train is assigned exactly one composition. Constraints (4) limit the length of the assigned composition to the maximum allowed train length. The constraints in (5) limit the maximum number of assigned train units to the fleet size for each train type.

3 Case study

In this section, we present the results of our case study regarding the capacity shortages and the controllability of three alternative line plans for the Dutch railway network. First, we present the three alternatives and their underlying principles. Subsequently, we discuss the capacity shortages of each alternative as calculated using the methodology presented in Sect. 2. The results are compared with the current winter timetable, the LUD. Afterwards, the controllability of each alternative is determined using the robustness index. A sensitivity analysis on this robustness index is presented in Sect. 3.3.

The line plans in this case study are generated using a simple line planning tool, developed by NS for evaluation purposes. The line plans comply to the contractual agreements with the Dutch government, but are not tested for feasibility in a real timetable. From an infrastructural point of view, the line plans are feasible. Three variants of line plans have been used, the approach is based on the reasoning in Sect. 2.2, where we described the aspects of a controllable line plan:

- A1: An alternative with short lines, such that the number of major stations a line serves is constrained to a maximum of 4. This alternative aims to reduce the *impact* of disruptions on the network.
- A2: A control-based alternative, such that lines are limited to attend a maximum of 2 traffic control regions. This alternative aims to reduce *coordination* between control regions.
- A3: An infrastructure-based alternative where the operation of high-speed switches is evaded by locking the switch in one direction (either straight or diverged to left/right). This alternative aims to reduce the *risk* on disruptions at all.

The LUD is the current winter timetable and therefore used as reference, i.e. the zero-alternative (A0). In addition, we compare the transport capacity and the robustness index with the regular timetable.

3.1 Calculation of capacity shortage

The capacity shortage of each alternative has been calculated using the models presented in the previous section. The model has been implemented in AIMMS 3.14 using CPLEX 12.6. The used hardware is a Pentium i7 processor with 3.40 GHz and 16 GB RAM. Per alternative and line type, between 120 and 140 trains have been assigned a composition which gives a model with about 2900 decision variables and 600 constraints. Solving the model takes less than 0.1 s.

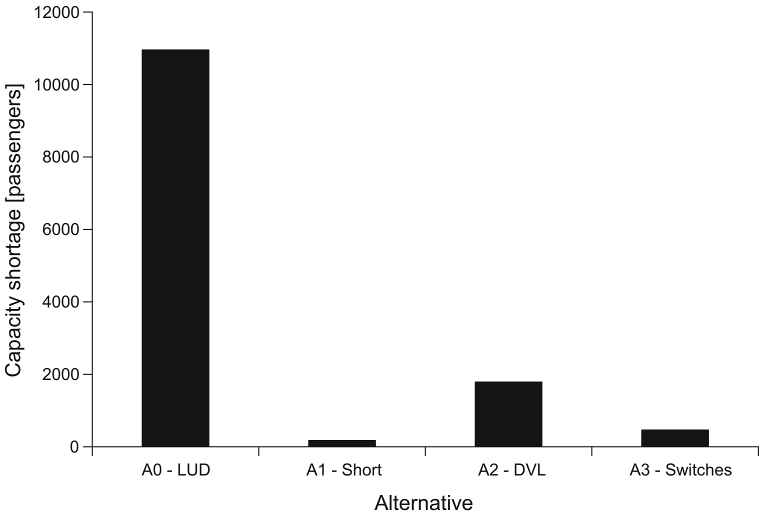

Fig. 3 Capacity shortages for all alternatives

Figure 3 gives an overview of the sum of the capacity shortage per alternative for all trains. This clearly shows the large capacity shortages during the LUD. If we analyze the results per train we note that some trains require more than 700 additional passenger places, which indisputably results in passengers left behind on the platform. All new alternatives are providing a considerably better transport capacity. A1 is the best of these.

3.2 Assessment of controllability

To assess the controllability of our alternatives, the indicators have been further specified. A list of major stations, control regions (called DVL for the infrastructure manager ProRail and RBC for the railway operator NS) and other details can be found in Trap (2014). Using the AHP, we have weighed our criteria and indicators. First, the four criteria have been evaluated. The importance of the groups (hence, their weight) has initially been estimated in consultation with an experienced transport controller. In a second stage, weights have been varied to verify the impact of the weights on the robustness index. This will be further elaborated on in Sect. 3.3. Line length is considered less important than all other criteria, since the line length cannot be expressed in a very structured way. Traffic intensity is considered the most important criterion. The busier the network is, the more dependencies there are between trains and lines. Since less trains will give more slack, this has been the most important reason to deploy the LUD for instance. The control region attendance and disruption risk are positioned in-between. The initial weights are shown in Table 2.

Secondly, all indicators *within* the criteria have been compared using the AHP. This is, again, initially done in accordance with a transport controller. The number of served major stations is the only indicator within its parent criteria group and therefore has a weight of 100%. Within the traffic intensity group, the number of edges with a frequency >4 is the most important indicator because this is more than the basic train service. A frequency of at most 4 trains/h is considered safe and can be controlled well in case of a disruption. Regarding the control region attendance, the number of trains in RBC regions is considered less important than the number of trains in DVL regions. This is because the traffic controllers are the first to respond in case of a disruption. The number of trains in DVL Amsterdam and DVL Utrecht is considered more important than in the regions Den Haag and Rotterdam. This is due to the size of these regions and the fact that the largest stations Utrecht Centraal and Amsterdam Centraal are located in these regions. All other DVL regions are not considered, since they have much less traffic to control. For the RBC regions, a

Table 2 Initial AHP weights for the criteria

Criterion	Weight (%)
Line length	9.7
Traffic intensity	36.5
Control region attendance	28.5
Disruption risk	25.3

similar reasoning holds. The operation ratio of the high-speed switches is furthermore considered more important than the number of operational switches itself. Multiplying the indicator weight by the weight of its parent (i.e. the criterion) yields the total weight. These initial weights are shown in Table 3. Section 3.3 describes the sensitivity of the robustness index to these weights. When we calculate the weighted sum of the indicators for all alternatives, we obtain the scores presented in Table 4.

The results indicate that A3 is the most robust alternative, followed by A2 and A1. All three alternatives are, according to these criteria, by far more robust than the zero-alternative. As a reference, we also calculated the robustness index of the regular line plan, being 118.6. This indicates that the LUD line plan is more robust than the regular line plan, which is in accordance with the expectation. The succeeding sections will elaborate further on the validity of these results and how the robustness index relates to the capacity shortage of all alternatives.

3.3 Sensitivity analysis

The initial weights used in the MCA are estimated using subjective judgement. To assess the impact of the weights on the calculated robustness index, a sensitivity analysis is performed. By changing the weights of the criteria and the indicators, the robustness index of the alternatives will change as well. The zero-alternative will always have the same index of ≈100.

Table 3 Initial AHP weights for the indicators

Indicator	Weight (%)
Major stations served	9.70
Frequency	3.43
Edges with Frequency >4	20.95
Line density	3.69
Edges with >2 lines	8.43
Trains in RBC Randstad Noord	1.25
Trains in RBC Randstad Zuid	1.25
Trains in RBC Utrecht	1.25
Attended transport control regions	1.48
Trains in DVL Amsterdam	4.96
Trains in DVL Den Haag	2.74
Trains in DVL Rotterdam	2.74
Trains in DVL Utrecht	4.96
Attended traffic control regions	7.92
High-speed switches in use	10.12
Switch operation ratio	15.18

Table 4 Results of the MCA with initial weights and standardized indicator values

Criterion	Indicator	Standardized values			
		A0	A1	A2	A3
Line length	Major stations served	100.00	73.41	56.01	77.51
Traffic intensity	Frequency	100.00	91.52	88.78	93.40
	Edges with Frequency >4	100.00	82.08	79.25	95.28
	Line density	100.00	71.81	69.95	70.32
	Edges with >2 lines	100.00	33.79	28.97	26.90
Control region attendance	Trains in RBC Randstad Noord	100.00	97.50	112.50	97.50
	Trains in RBC Randstad Zuid	100.00	118.75	106.25	96.88
	Trains in RBC Utrecht	100.00	102.08	100.00	104.17
	Attended transport control regions	100.00	82.71	76.28	88.76
	Trains in DVL Amsterdam	100.00	91.67	102.78	91.67
	Trains in DVL Den Haag	100.00	116.67	122.22	105.56
	Trains in DVL Rotterdam	100.00	121.43	100.00	85.71
	Trains in DVL Utrecht	100.00	102.78	88.89	105.56
	Attended traffic control regions	100.00	77.99	64.42	83.68
Disruption risk	High-speed switches in use	100.00	74.29	74.29	11.43
	Switch operation ratio	100.00	49.99	47.59	8.86
	Robustness index	100.06	75.61	70.66	64.75

Table 5 Scenarios for sensitivity analysis

Scenario	Description
S1	Default AHP weight as explained in Sect. 3.2
S2	Equal weight for all criteria
S3	Equal weight for all indicators within the same group
S4	Equal weight for both criteria and indicators within the same group (no weight)
S5	Line length is excluded from the analysis. Other criteria are reweighed
S6	Traffic intensity is excluded from the analysis. Other criteria are reweighed
S7	Attended control regions are excluded from the analysis. Other criteria are reweighed
S8	Disruption risk is excluded from the analysis. Other criteria are reweighed
S9	Line length is excluded from the analysis. Other criteria are of equal weight
S10	Traffic intensity is excluded from the analysis. Other criteria are of equal weight
S11	Attended control regions is excluded from the analysis. Other criteria are of equal weight
S12	Disruption risk is excluded from the analysis. Other criteria are of equal weight

The varying of weights is performed using different *scenarios*. Each scenario has a different distribution of weights, such that it is possible to focus on specific criteria or exclude indicators from contributing to the robustness index. Table 5 shows the scenarios that have been drafted.

All scenarios are based on the default scenario (S1), which means that unchanged weights are the same as in S1. Scenarios S2–S4 are used to determine the robustness index if the criteria and/or the corresponding indicators are weighted equally. Scenarios S5–S8 exclude one of the criteria from the analysis by changing its weight to 0% to assess the impact of the respective criterion on the robustness index. The other criteria are reweighed in order of importance using the AHP process. A third group of scenarios (S9–S12) excludes one of the criteria as well, while the other three groups are weighed equally.

Figure 4 presents the results of the sensitivity analysis. The lines in the figure indicate the robustness index for each alternative for each scenario. The robustness index of the regular line plan is added to indicate that the LUD is more robust than the regular line plan in all scenarios, which is in accordance to the expectations.

The chart in Fig. 4 also shows that the ranking order between the alternative line plans is very stable. In almost all scenarios, A3 has the lowest robustness index, followed by A2 and A1. In S8 and S12, however, A3 is *less robust* than both A1 and A2. In both scenarios, the criterion "disruption risk" is excluded from the analysis. We therefore conclude that the low value of the robustness index of A3 is mainly caused by this criterion. This is also visible in Table 4, as the scores of the indicators in this group are very low. Since A3 becomes the least robust alternative in S8 and S12, the usefulness of the number of operational switches and their operation ratio, or at least their weight, in the MCA is questionable.

Based on this sensitivity analysis, we can conclude that all alternative line plans are in any case more robust than the LUD, and that the robustness index is only slightly sensitive to the applied weights.

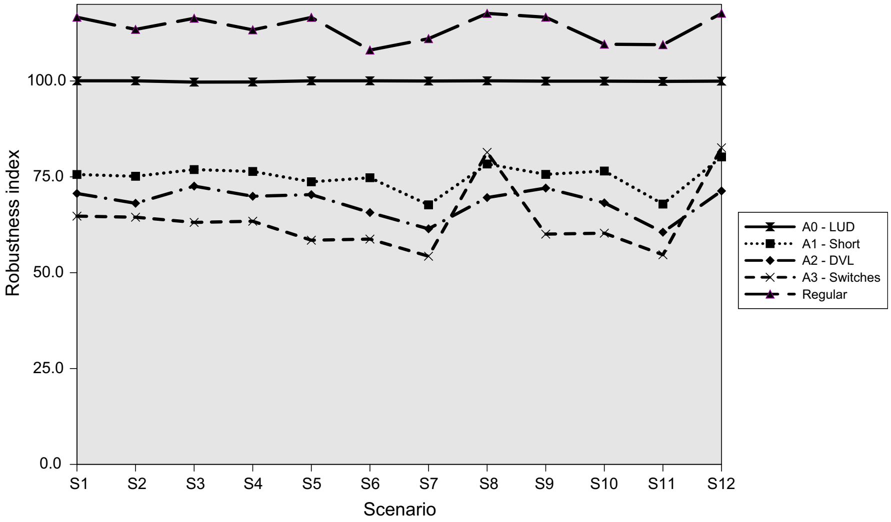

Fig. 4 Robustness index of all alternatives using different weight scenarios

3.4 Summary of the results

The LUD baseline and the alternatives all use the same pool of rolling stock. Figure 5 shows the relation between the robustness index and the capacity shortage. This clearly indicates that all three alternatives are better than the zero-alternative. The sensitivity analysis made clear that the robustness index of A3 is much depending on the weights in the MCA.

The following conclusions can be drawn based on the robustness index and the capacity shortage of the three alternatives:

- A1 and A3 have the least capacity shortage and no unacceptable shortage per composition. A1 is the best of these.
- A2 has a relatively large shortage and requires more rolling stock than in the operational fleet.
- Depending on the weight, A3 can be the best or the worst alternative regarding the robustness index, but is still more robust than A0
- A1 and A2 have a relatively stable robustness index.

Based on these statements, we can conclude that A1 and A3 are considerably better than A2. Moreover, all alternatives score much better than the current winter timetable. The difference between the alternatives and LUD is largely due to the train length. LUD has few very long lines which constrain the train length due to short platforms. Consequently, short trains run through the busy cities as their length is limited by a short platform in a peripheral town. For the alternatives, iterations have been made to reduce the capacity shortage to a minimum.

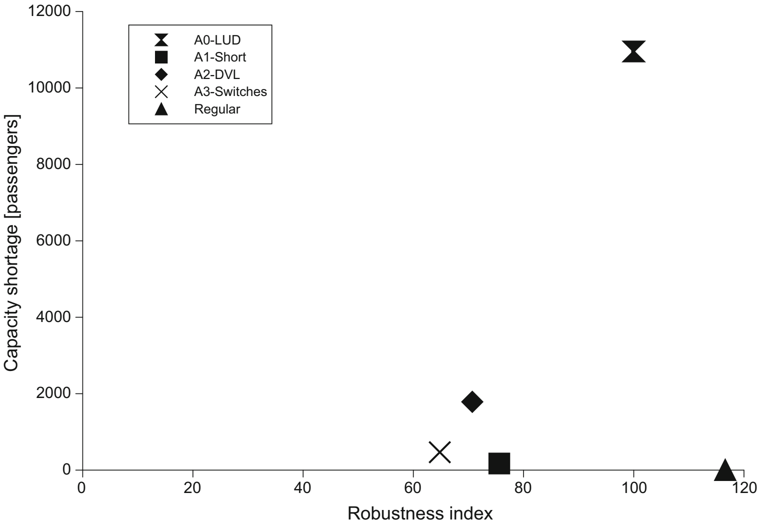

Fig. 5 Ranges of the robustness index over all alternatives and their capacity shortage

4 Conclusions

In this paper, we developed a methodology to design a robust line plan and compute its transport capacity. The controllability of a line plan is evaluated on several criteria, which are important on days with heavy weather conditions. To calculate the transport capacity, passengers from the O–D Matrix are allocated to the different lines and trains on the lines to estimate the travel demand per train. The difference between the demand and the train capacity determines the capacity shortage. We showed that there are several alternative line plans possible that score better than the currently operated winter timetable in robustness, resilience and transport capacity. The differences in transport capacity are largely due to the train length, which is often constrained by the maximum platform length.

A subsequent study could focus on the feasibility of the alternatives in a cyclic timetable. Many lines in the alternatives are operating on a frequency of three trains per hour, of which the impact should be analyzed. Furthermore, the infrastructure occupation has been simplified in this paper. A more comprehensive approach like the timetable compression method (Goverde et al. 2016) could make the results more realistic. The results could subsequently be used to further improve the performance of the alternatives.

Acknowledgements We would like to thank Timo van de Walle and Serge Hoogendoorn for their valuable input on an earlier draft of this paper. Finally, we are grateful to two anonymous referees for their suggestions to improve the manuscript.

References

Abbink E, Van Den Berg B, Kroon LG, Salomon M (2004) Allocation of railway rolling stock for passenger trains. Transp Sci 38(1):33–41. doi:10.1287/trsc.1030.0044

Akiva MEB, Lerman SR (1985) Discrete choice analysis: theory and application to predict travel demand, vol 9. MIT Press, Cambridge

Andersson E, Peterson A, Törnquist Krasemann J (2013) Introducing a new quantitative measure of railway timetable robustness based on critical points. Copenhagen, Denmark. http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-92712

Bussieck MR (1998) Optimal lines in public rail transport. Ph.D. Thesis, TU Braunschweig. http://d-nb.info/955665965/34

Cacchiani V, Caprara A, Galli L, Kroon LG, Maróti G, Toth P (2012) Railway rolling stock planning: robustness against large disruptions. Transp Sci 46(2):217–232. doi:10.1287/trsc.1110.0388

Claessens MT, Van Dijk NM, Zwaneveld PJ (1998) Cost optimal allocation of rail passenger lines. Eur J Oper Res 110(3):474–489. doi:10.1016/S0377-2217(97)00271-3

Dow JK, Endersby JW (2004) Multinomial probit and multinomial logit: a comparison of choice models for voting research. Electoral Studies 23(1):107–122. doi:10.1016/S0261-3794(03)00040-4

Fioole PJ, Kroon LG, Maróti G, Schrijver A (2006) A rolling stock circulation model for combining and splitting of passenger trains. Eur J Oper Res 174(2):1281–1297. doi:10.1016/j.ejor.2005.03.032

Gestrelius S, Aronsson M, Forsgren M, Dahlberg H (2012) On the delivery robustness of train timetables with respect to production replanning possibilities, Dubrovnik. http://soda.swedish-ict.se/5320/

Goerigk M, Schachtebeck M, Schöbel A (2013) Evaluating line concepts using travel times and robustness. Public Transp 5(3):267–284. doi:10.1007/s12469-013-0072-x

Goverde RM, Bešinović N, Binder A, Cacchiani V, Quaglietta E, Roberti R, Toth P (2016) A three-level framework for performance-based railway timetabling. Transp Res Part C Emerg Technol 67:62–83. doi:10.1016/j.trc.2016.02.004

Goverde RMP, Hansen IA (2013) Performance indicators for railway timetables, Beijing, pp 301–306. doi:10.1109/ICIRT.2013.6696312

Jespersen-Groth J, Potthoff D, Clausen J, Huisman D, Kroon LG, Maróti G, Nielsen MN (2009) Disruption management in passenger railway transportation. In: Ahuja RK, Möhring RH, Zaroliagis CD (eds) Robust and online large-scale optimization, Lecture Notes in Computer Science, vol 5868. Springer, Berlin, pp 399–421. http://link.springer.com/chapter/10.1007/978-3-642-05465-5_18

Kaspi M, Raviv T (2013) Service-oriented line planning and timetabling for passenger trains. Transp Sci 47(3):295–311. doi:10.1287/trsc.1120.0424

Klibi W, Martel A, Guitouni A (2010) The design of robust value-creating supply chain networks: a critical review. Eur J Oper Res 203(2):283–293. doi:10.1016/j.ejor.2009.06.011

Kroon LG, Dekker R, Vromans MJCM (2007) Cyclic railway timetabling: a stochastic optimization approach. In: Geraets F, Kroon LG, Schoebel A, Wagner D, Zaroliagis CD (eds) Algorithmic methods for railway optimization, Lecture Notes in Computer Science, vol 4359. Springer, Berlin, pp 41–66. http://link.springer.com/chapter/10.1007/978-3-540-74247-0_2

Louwerse I, Huisman D (2014) Adjusting a railway timetable in case of partial or complete blockades. Eur J Oper Res 235(3):583–593. doi:10.1016/j.ejor.2013.12.020

Rosenhead J (2013) Robustness analysis. In: Gass SI, Fu MC (eds) Encyclopedia of operations research and management science. Springer US, pp 1346–1347. http://link.springer.com/referenceworkentry/10.1007/978-1-4419-1153-7_200722

Saaty TL (1990) How to make a decision: the analytic hierarchy process. Eur J Oper Res 48(1):9–26. doi:10.1016/0377-2217(90)90057-I

Schaafsma A (2001) Dynamisch Railverkeersmanagement besturingsconcept voor railverkeer op basis van het Lagenmodel Verkeer en Vervoer. Ph.D. Thesis, Delft University of Technology, The Netherlands. http://resolver.tudelft.nl/uuid:4245d6ab-cf07-4daa-a1d4-607f5a53190c

Schöbel A (2012) Line planning in public transportation: models and methods. OR Spectrum 34(3):491–510. doi:10.1007/s00291-011-0251-6

Trap ML (2014) The Dutch winter timetable: assessment of alternative line systems for the Dutch Railway Network during winter weather. Master Thesis, Delft University of Technology, Delft. http://resolver.tudelft.nl/uuid:10ed75ca-76dd-4293-aeb1-f2005a1d30e1

Van Oort N, Van Nes R (2009a) Controlling operations of public transport to improve reliability: theory and practice. Washington DC

Van Oort N, Van Nes R (2009b) Line length versus operational reliability: network design dilemma in urban public transportation. Transp Res Record 2112(3):104–110. doi:10.3141/2112-13

Warmerdam J (2004) Specificaties TRANS toedeler. Internal Document. QQQ Delft, Delft