Calibration of Timetable Parameters for Rail-Guided Systems

Weiting Zhao, Ullrich Martin†, Yong Cui and Maureen Kösters

Abstract

In order to achieve a comprehensive utilization of railway networks, it is necessary to accurately assess the timetable indicators that affect the train operation. This paper describes the parameter calibration for two timetable indicators: scheduled running time and scheduled dwell time. For the scheduled running time, an existing model is employed and the single timetable parameter (percentage of minimum running time) in that model is optimized. For the scheduled dwell time, two intrinsic characteristics: the significance of stations and the average headway at each station are proposed firstly to form a new model, and the corresponding timetable parameters (the weight of the significance and the weight of the average headway) are calibrated subsequently. The Floyd Algorithm is used to obtain the connectivity among stations, which represents the significance of the stations. A case study is conducted in a light rail transportation system with 17 underground stations. The results of this research show that the optimal value of the scheduled running time parameter can be automatically determined, and the proposed model for the scheduled dwell time works well with a high coefficient of determination and low relative root mean square error through the leave-one-out validation.

Keywords: Calibration, Scheduled Running Time, Scheduled Dwell Time, Floyd Algorithm, Simulation of Railway Operation, Capacity Research

1. Introduction

In order to ensure the accuracy of the scheduled conveyance time\(^1\) in the process of timetable design, recovery time is added to the minimum running time and minimum dwell time to compensate for the influences from stochastic disturbances. Stochastic disturbances include disturbances such as initial delays, running time extensions, dwell time extensions and departure delays. The scheduled running time can be calculated based on the recovery time (running time related) and the pure calculated running time (also known as the minimum running time); the minimum running time is calculated using sufficient information about the infrastructure and the tractive units. The scheduled dwell time can be determined based on the minimum dwell time and the customer-related (passenger-transport and freight-transport) dwell time (including customer-related recovery time) (see Fig. 1).

This research deals with two timetable indicators: the scheduled running time and the scheduled dwell time. As for scheduled running time, one approach is to distribute the recovery time evenly on the whole network. Another alternative is to spread the majority of recovery

\(\text{Fig. 1 Share of scheduled conveyance time (Martin, (9/4/2014))}\)

\(^1\) scheduled conveyance time: It is the sum of scheduled running time and scheduled dwell time.
time between capacity bottlenecks such as large stations/junctions (Ning & Brebbia, 2010). In this paper, recovery time is added over each block section with values calculated based on a certain fixed ratio of the minimum running time. Different block sections can be treated separately because of the different minimum running time simulated with its corresponding infrastructure characteristics and the operating rolling stock. The ratio of minimum running time, taken as timetable parameter associated with scheduling running time in this paper, is widely used both on European railways (3-7% of minimum running time) and on North American railways (6-8% of minimum running time) . As for scheduled dwell time, the number of passengers boarding and alighting at each station is taken as the timetable parameters associated with the scheduled dwell time for decades. A survey on counting these parameters was conducted and the results was further used for building the scheduled dwell time model based on the regression analysis in (Su, 2010). A method for estimating this parameter based on a static origin-destination matrices in dynamic urban traffic models is presented in (Marchal & Palma, 2001). In (Yang et al., 2009), this parameter is assumed as a fuzzy variable and the expected value of this parameter is applied under different scenarios for designing the optimized timetable through heuristic algorithms. Instead of applying the number of passenger boarding and alighting, an innovative model of scheduled dwell time is developed in this paper. The weight of the significance of stations and the weight of the average headway, which serve as the two timetable parameters associated with the scheduled dwell time, are calibrated in the scheduled dwell time model.

Fig. 2 shows the calibration process for the timetable design. Firstly, the initial value of the timetable parameters will be set up. Subsequently, the temporary value of the indicators will be calculated by using the scheduled running time model (see Chapter 3) and the scheduled dwell time model (see Chapter 4). By comparing the temporary values and existing corresponding target values, the timetable parameters can be iteratively adjusted until the error² has been minimized.

Accordingly, this research establishes a more convenient and effective way of calibrating the timetable parameters. The objective of this research is to carry out the calibration system of the timetable parameters in order to accurately estimate the timetable indicators for rail-guided systems during scheduling and simulation of the railway operation process.

An example of the application scenario is explained in Chapter 2. The calibration systems for scheduled running times and scheduled dwell times are discussed in Chapter 3 and Chapter 4 of this paper, respectively. A case study of one railway network is discussed in Chapter 5. This current research is supported by the German Research Foundation (DFG) (project number MA 2326/9-1). The conclusions are stated in Chapter 6.

2. Application Scenario

Given the circumstance that a concrete, detailed timetable of operational simulation is not available (e.g. timetable design for a new railway network or for a given railway network with new train types), the scheduled conveyance time, including the scheduled running time and the scheduled dwell time must first be defined.

Fig. 3 describes an example scenario for the timetable design of a modified railway network. For the two existing stations A and B, the scheduled running time and minimum running time on the track (lines located between two successive stations) between the two stations are known. Therefore, the difference between the scheduled running time and the minimum running time (the recovery time) can be calculated. In this calibration system, the timetable parameter can be determined based on the known recovery time. Once a new station is planned, it will be checked if the track between Station B and the new station is similar to the existing track between Station A and Station B (e.g. similar track condition and geometry). If it is, the adjusted timetable parameter can be directly employed to determine the recovery time (running time related) on the new track section and therefore to estimate the scheduled running time.

²Error: the difference between the temporary values and the corresponding target values.
3. Calibration of Scheduled Running Time

3.1 Convergence criterion and timetable parameters

The scheduled running time model, its relevant timetable parameter (the percentage of the minimum running time), and the convergence criterion will be explained in this chapter. The waiting time, the synchronization time, and the influences caused by individual transport in mixed-use traffic areas are not considered in the calibration process because they are caused mainly by external disturbances.

The scheduled running time is the sum of the minimum running time, the regular recovery time, and the special recovery time (Hansen & Pachl, 2014). Thus, the scheduled running time model can be expressed as follows:

\[
R_T = R_{\text{min}} + R_{rr} + R_{sr}
\]

where

- \(R_T\): the scheduled running time between two successive stations in real operation [s]
- \(R_{\text{min}}\): the minimum running time [s]
- \(R_{rr}\): the regular recovery time [s]
- \(R_{sr}\): the special recovery time [s]

The minimum running time is the shortest possible running time resulting from a running time calculation based on the infrastructure and tractive units used on the infrastructure. In this paper, it is calculated by the simulation tool RailSys® (RMCON Rail Management Consultants, 2010), which is able to model the real operation process based on the structure of the railway network, location of the signal system, and the train operating program. The tractive effort, traction unit resistance, rail vehicle resistance, line resistances, combining efforts, and resistances and rotating masses must be predefined to simulate the acceleration and deceleration process of the tractive units. The regular recovery time is usually added to every train path as a fraction of the minimum running time to compensate for slower train speeds, to avoid departure delays, and to save energy in the acceleration and deceleration process. The special recovery time is used to compensate for the influence of maintenance and construction works and that of sections with temporary bad track conditions and low permissible operation speeds. The special recovery time is different from the regular recovery time because it is not calculated as a percentage of the mini-
current running time but added to the running time as a fixed value based on the specific track section considered. In this paper, the special recovery time will be set as null because no specific construction sites etc. are considered as a part of this research. Therefore, the stochastic disturbances during the operation process will be compensated only by the regular recovery time.

In the calibration system, the target values of the scheduled running times are the scheduled running times in real operation: these have been taken from existing database. The temporary values of the scheduled running time are calculated based on the sum of the minimum running time and the recovery time (running time related). The recovery time (running time related) is calculated as a percentage of the minimum running time (3-7% in Germany). This percentage is taken as the value of the timetable parameter in the calibration process and is adjusted in the following model. The calibration converges at the optimal value of this parameter leading to the minimum error (see Footnote 2).

For a train run with track sections, the temporary value of the scheduled running time for each track section can be calculated with the timetable parameter as shown in Equation (2).

$$R_{Ti}^{δ} = R_{Tmin, i} \cdot (1 + δ)$$  \hspace{1cm} (2)

where

- $i$ the index of track sections between two successive stations with $i \in [1, n]$
- $δ$ the percentage-timetable parameter to calculate the recovery time (running time related) [%]
- $R_{Tmin, i}$ the minimum running time on the $i$-th track section [s]
- $R_{Ti}$ the temporary value of the scheduled running time with parameter $δ$ on $i$-th track section [s]

In order to minimize the error in the calibration process, the difference between the scheduled running time in real operation and the temporary values of scheduled running time based on Equation (2) are calculated. Since the differences are not always the same for every track section, the total difference must be considered. The convergence criterion used in this paper is shown as follows:

$$\min_{\delta} \sum_{i=1}^{n} \frac{R_{Ti} - R_{Tmin, i \cdot (1 + \delta)}}{R_{Tmin, i}}$$ \hspace{1cm} (3)

where

- $R_{Ti}$ the target value (scheduled running time in real operation) on the $i$-th track section [s].

### 3.2 Calibration process

During the calibration process, the value of the timetable parameter $δ$ is adjusted until the convergence criterion is satisfied. Since there is only one timetable parameter that can be adjusted, the calibration process can be directly carried out by calculating the error with different values of the parameter $δ$.

By varying the parameter $δ$ from 0% to 100% with a fixed interval of 0.01%, the optimal $δ$ can be identified that leads to the minimum value of the error. In the circumstance that the parameter is equal to 0%, the train always runs at its maximal speed and no regular recovery time is considered; which is not possible in real operation. On the contrary, if $δ$ reaches 100%, the regular recovery time is long enough to compensate for the stochastic disturbances. However, an unnecessarily long running time will lead to low capacity, which is also not desirable in real operation.

### 4. Calibration of Scheduled Dwell Time

The scheduled dwell time model, its relevant timetable parameters (the weight of the significance of stations and the weight of the average headway), and the convergence criterion will be explained in this chapter.

Scheduled dwell time is a designed time period for passengers alighting and boarding as well as freight handling at a station. It consists of customer-related dwell time, time for the preparation of train departure, and dwell time for operational reasons (see Fig. 1). The customer-related dwell time includes the time for passenger boarding and alighting as well as freight handling and varies according to the number of passengers or quantities of goods. The minimum dwell time is fixed and includes the time for opening and closing the doors, checking and clearance, and response. There are several parameters that influence the number of boarding and alighting passengers or quantities of goods at a certain station, such as the operational time period and the location of the station.

To the best of the authors’ knowledge, there is no existing model to determine the scheduled dwell time at each station in a given network. Therefore, developing a new model is necessary. To alleviate the time consuming and labor intensive work of counting the numbers of boarding and alighting passengers or measuring the time for freight handling at each station, and in order to make this model applicable for various regions, only the intrinsic characteristics, including the connectivity with other important lines (significance of stations) and the average headway at each station, are introduced into the model to determine the length of the scheduled dwell time.

### 4.1. Intrinsic characteristics

A tightly-scheduled dwell time can cause serious depar-
Calibration of Timetable Parameters for Rail-Guided Systems

ture delays in railway operations. Thus, the scheduled dwell time must be estimated accurately. The scheduled dwell time model is firstly built using a nonlinear regression analysis method and based on the minimum dwell time with consideration to the significance of the stations and the average headway at each station. The involved timetable parameters are calibrated subsequently.

4.1.1 Significance of stations

Significance of stations represents the connectivity of the station to others. Normally, the more significant the station, the better its accessibility is to other places. In this study, the significance of a station is defined by the number of times the station appears in the shortest paths of all origin-destination (OD) pairs. As an intrinsic characteristic, the significance of stations is determined by the structure of the railway network. Since it requires modest manual measurements, the method is relatively universally applicable.

The Floyd Algorithm is used to find the shortest paths for all OD pairs between the different stations. Different from the uniform networks usually used with the Floyd Algorithm, in this research, transfers must be considered during the selection of the shortest paths due to the inconvenience. Therefore, the conventional algorithm is not applicable in this paper; some adjustments and modifications have to be done firstly.

A transfer implies not only additional travel time consumption, but also the physical exertion associated with a person physically having to travel to another platform. Inconvenient transfers encourage the passengers to choose another path, which reduce the significance of a well-connected station. Therefore, in the case of only a slight increase in travel time without a transfer, most passengers prefer not to transfer.

In order to quantify the inconvenience of transfer in the simulation process, a penalty parameter is defined as “offset”. The offset is the number of stations the passengers can continuously transit if they use up the time consumption which is assumed for one transfer (Equation (4)). The negative psychological effects of transfers are also considered. The unit of this offset is station.

\[
\text{offset} = \frac{R_{\text{trans}}}{R_{\text{run}}}
\]

where

- \( R_{\text{trans}} \) the impedance of the time consumption for one transfer [s]
- \( R_{\text{run}} \) the impedance associated with the time consumption for continuously driving one additional station[s]

The impedance of one transfer includes three parts: waiting time on the platform, walking time for changing the platform, and the negative psychological effects. The detailed method for calculation is explained in (ITP & VWI, 2006). The time consumption impedance for continuously driving one additional station includes two parts: the average scheduled running time between two adjacent stations and the average scheduled dwell time of the whole railway network.

The flowchart of the significance calculation is shown in Fig. 4. The shortest paths of all OD pairs are firstly determined and listed. Then the number of the times each station appears in all the shortest paths will be counted.

Step 1: Create an adjacency matrix and a matrix of the railway network according to the structure of the railway network. The former provides the accessibility among all stations; the latter records which line a station belongs to.

Step 2: Choose an origin station and a destination sta-
tion, and set the distance between the two stations as the initial value of the shortest distance between them.

Step 3: Choose another station K as a transfer station and determine whether the origin, transfer, and destination stations are on the same line. If so, set the temporary distance as the sum of the distance from the origin station to the transfer station and from the transfer station to the destination station. If not, add a penalty to the aggregated distance.

Step 4: Determine whether the temporary distance is less than the current shortest distance between the origin station and the destination station. If it is less, update the shortest path for (j+1)-th iteration with temporary distance. If not, keep the current shortest distance unchanged.

Step 5: Iterate Step 3–4 until all stations in the railway network have been considered as the transfer station.

Step 6: Iterate Step 2–5 to obtain the shortest distance between every OD pair.

Step 7: Count the number of times each station appears in the shortest paths of all OD pairs and finally calculate the significance of each station in the railway system.

4.1.2 Average headway at each station

According to (Pachl, 2002), headway is the time interval between two successive trains. It is a key factor for estimating the traffic volume based on a fixed carrying capacity of trains. The headway at stations is designed to ensure adequate facilities to handle the peak-hour travel demand of passengers or goods both in short and long term. A shorter headway means a more frequent service, which remains less passengers or goods at stations that need to be transported per train path. Thus, the customer-related dwell time for passengers alighting and boarding as well as the freight handling per train path will be relatively small. Accordingly, the inverse relationship between headway at stations and the scheduled dwell time can be deduced. In general, the average headway at each station is the average time interval between every two successive trains, which can be calculated by the reciprocal of service frequency (Equation (5)).

\[
\overline{H} = \frac{1}{F}
\]

where
\[
\overline{H} \text{ the average headway at the station [s]}
\]
\[
F \text{ the service frequency (trains per hour)}
\]

In order to simplify the automatic calculation process of the average headway at each station, the clockface schedules of railway system is employed in this paper. The average headway at each station can be calculated then based on different fixed time intervals of the related train paths in which the corresponding trains stop at the station.

In this paper, the average headway at each station is considered as another intrinsic characteristic of the station; it is used to build the scheduled dwell time model. For a station belonging to train paths, the average headway can be derived as follows:

\[
\overline{H} = \frac{1}{\sum_{i=1}^{n} \frac{1}{H_i}}
\]

where
\[
i \text{ the index of the train paths in which the corresponding trains stop at the station}
\]
\[
H_i \text{ the fixed time interval of related train paths [s]}
\]

4.2 Calibration process

The target values of the scheduled dwell time are the scheduled dwell time in the real operation, while the temporary values of the scheduled dwell time are calculated based on the minimum dwell time and the intrinsic characteristics with their related timetable parameters. The timetable parameters work as the weight of the two intrinsic characteristics: significance of stations and average headway at each station. Therefore, the temporary values of the scheduled dwell time can be calculated with Equation (7):

\[
D_T^{w_1, w_2} = w_1 \cdot S + \frac{w_2}{\overline{H}} + D_{T_{min}}
\]

where
\[
D_T^{w_1, w_2} \text{ the temporary value of scheduled dwell time at the station [s]}
\]
\[
S \text{ the significance of the station}
\]
\[
w_1 \text{ the timetable parameter associated with the significance of the stations}
\]
\[
w_2 \text{ the timetable parameter associated with the average headway at each station}
\]
\[
D_{T_{min}} \text{ the minimum dwell time at the station [s]}
\]

The timetable parameters will be adjusted in the new developed model until the difference between the scheduled dwell time in the real operation and the scheduled dwell time calculated with Equation (7) is minimal. The convergence criterion of the calibration system with a railway network with stations will be as follows:

\[
\min \left\{ \sum_{i=1}^{m} (D_T^{w_1, w_2} - D_{T_{i}}) \right\}
\]

where

3Clockface schedules: The schedules of a timetable in which trains that belong to the same route are scheduled with fixed time intervals between their train paths (Pachl, 2002).
the index of stations

\[ D_{t_i}^{w_1, w_2} \]  the temporary value of the scheduled dwell time at the \( t_i \)-th station [s]

\[ D_{t_i} \]  the target value of the scheduled dwell time in real operation at the \( t_i \)-th station [s]

The bivariate linear regression analysis will be used based on Equation (7) until the error is minimal. The timetable parameters will be determined by using the linear regression analysis.

4.3 Model validation

In order to examine if the proposed model is applicable for further estimations in macro scopes, it’s essential to perform model validation. K-fold cross-validation is used because it makes good use of the available data both as training and test data (Cawley & Talbot, 2004).

The leave-one-out cross-validation is the extreme case of the k-fold cross-validation, and is used when the sample size of a subset is equal to 1. Due to the simplicity and convenience, it is employed in this research.

5. Case Study of a Light Rail Transportation System

A case study was conducted as part of this research. In the case study, a light rail transportation system with 17 underground stations was selected as the study area in order to avoid the influence from other factors such as individual transport. The scheduled running time and the scheduled dwell time in the real operation were obtained as the target values to calibrate the timetable parameters.

5.1 Calibration results of the scheduled running time

The results can not only be used to determine the appropriate scheduled running time and scheduled dwell time for a new train type or new railway network, but can also be used to optimize the capacity of an existing railway network. The results also tell us that if the timetable parameter is extremely large for one track section, this track section must be a bottleneck in the system or the infrastructure data must be wrong in the simulation tools.

Fig. 5 shows the calibration result of the scheduled running time in the study area. Direction D1 represents the direction from the city center to the suburbs while Direction D2 represents the opposite direction. It indicates that the optimal parameter \( \delta \) is 13.21% for Direction D1 and 20.00% for Direction D2. The optimal parameters \( \delta \) for the two directions are different, which manifests the necessity to differentiate the optimal parameter for different infrastructure operation conditions and different train types. If the calibrated parameter \( \delta \) were to be applied to another railway network, the similarity of the two railway networks must be examined. If they are not similar, the timetable parameter \( \delta \) must be further adjusted.

5.2 Calibration results of scheduled dwell time

The calibration of the scheduled dwell time is performed based on the model illustrated in Chapter 4.2. The offset is set to 10 stations in the Floyd Algorithm to compensate for the transfer inconvenience in the selected study area. There are 78 stations in total in the study area, which means \((78-1)^2 = 5929\) shortest paths for all OD pairs. The significance of stations can be determined by counting the number of times each station appears in the shortest paths of all OD pairs.

The proportional relationship between the significance of stations and the scheduled dwell time in the study area is shown in Fig. 6. The results are in line with what was expected. The significance of stations is high at stations located in the city center and low at stations located in the suburbs. These results indicate the connectivity degree

Fig. 6 Relationship between the significance of stations and the scheduled dwell time
Based on Equation (6), the average headway at each station can be calculated by using the designed fixed time intervals of different train paths in which the corresponding trains stop at the station. The results confirm the assumption of an inverse relationship between the average headway and the scheduled dwell time.

By combining the two intrinsic characteristics, the new scheduled dwell time model can be developed. The timetable parameters $w_1$ and $w_2$ can be adjusted according to the scheduled dwell time in real operation in this study area with the method of least squares, which is shown in Equation (9).

$$D_T = 0.004s \times S + \frac{224.22x^2}{H} + 11.125s$$

(9)

The fixed value is the minimum dwell time and the rest is the recovery time (dwell time related). In this study, the minimum dwell time is unknown, so that its value is also calibrated in the model. A model validation must be conducted before the proposed model with specific parameters can be applied to a large area. The method used for validation is the leave-one-out cross-validation (see Chapter 3.3). There are 15 samples in total, of which 14 are randomly chosen in the first step to create a regression function. Then the output value $\hat{y}$ is predicted by using the regression function. This process is conducted iteratively for 15 times. By comparing all of the predicted values of $\hat{y}$ with the observed values $y$ (original values), the goodness of fit of the model can be estimated.

The results of the validation are shown in Fig. 8, which illustrates the differences between the predicted and observed values. A high coefficient of determination $R^2$ (here $R^2 = 0.705$) shows that the new scheduled dwell time model with specific timetable parameters has a high prediction accuracy, which can be applied in practice.

6. Conclusions

In order to achieve comprehensive utilization of railway networks, it is necessary to accurately assess the timetable indicators that effect the train operation. The research described in this paper was conducted to calibrate two major indicators of the rail-guided system: the scheduled running time and the scheduled dwell time. The scheduled running time is calibrated by adjusting the timetable parameter of an existing model, while the scheduled dwell time is calibrated by building a new model using regression analysis. During the process of model generation, the modified Floyd Algorithm is applied to estimate the significance of the stations. A light rail transportation system with 17 underground stations is selected as a case study as part of this research.

There are two major findings in this paper: 1) The timetable parameter associated with the scheduled running time is able to be optimized. In this paper, the optimal value of this parameter is investigated in the study area. It is worth pointing out that even though the ratio of minimum running time is fixed over all block sections inside the investigation area, the amount of the regular recovery time added on each block section is different due to the different values of the minimum running time simulated with various infrastructure characteristics and the operating rolling stock. Nowadays, a fixed ratio of the minimum running time for all block sections is commonly used in the process of timetabling due to its simplicity. However, applying different ratios for different block sections can be investigated to further minimize the error. In that way, the convergence might be achieved at multiple values of the running time associated parameters. A study on analyzing the relationship between the ratios of the minimum running time and infrastructure characteristics as well as the...
operating rolling stock will become an additional topic in future research with sufficient database. 2) A scheduled dwell time model is built using two intrinsic characteristics of stations. One is the significance of stations in the railway network, and the other is the average headway at each station. For the former, it positively correlates with the scheduled dwell time with a coefficient of determination of 0.851; while the latter inversely correlates with the target value with a coefficient of determination of 0.781. A combination of the two parameters also has strong relationship with the scheduled dwell time, with a high coefficient of determination of 0.702 and a low relative root mean square error of 18.58% through the leave-one-out validation.

Acknowledgement

This work was funded within the framework of a German Research Foundation (DFG) project (with project number MA 2326/9-1).

References