Available online at www.sciencedirect.com

SciVerse ScienceDirect

Transportation Research Procedia 10 (2015) 625–633

Transportation Research Procedia

www.elsevier.com/locate/procedia

18th Euro Working Group on Transportation, EWGT 2015, 14-16 July 2015, Delft, The Netherlands

Practical Macroscopic Evaluation and Comparison of Railway Timetables

Peter Sels [a,b,*], Dirk Cattrysse [a], Pieter Vansteenwegen [a]

[a] *KU Leuven, Leuven Mobility Research Centre, CIB, Celestijnenlaan 300, 3001 Leuven, Belgium*
[b] *Logically Yours BVBA, Plankenbergstraat 112 bus L7, 2100 Antwerp, Belgium*

Abstract

From the perspective of passengers, a railway timetable can be called better than another if its expected passenger time is lower in practice. So, we constructed an analytical function that evaluates a timetable on this criterion: total expected passenger time in practice. Other methods to evaluate timetables invariably describe different performance indicators: realisability, conflict-freeness, stability, efficiency, robustness, resilience, but mostly do not indicate how to score and weigh these different performance indicators. This means that when comparing two timetables, deciding which one is preferable remains hard. Our objective of expected passenger time in practice resolves these issues.

Also, compared to a simulation approach, our analytical stochastic approach has a major practical advantage. It decouples all actions (ride, dwell, transfer, knock-on) in the timetable and in doing so, can evaluate the expected time in every action separately and simply add all expected times of composing actions afterwards. So the exponential amount of combinations of primary delays over all actions that standard simulation packages explicitly iterate over is dealt with implicitly and much more efficiently. This makes that the evaluation made by our method requires less time.

Our method is applied to two timetables of all passenger trains in Belgium. Both timetables were manually planned and then put into operation in practice. With our method, we can conclude that one timetable has considerably lower total expected passenger time in practice than the other one. We also show that this is caused mainly by better passenger transfer planning, but also partly by a changed line planning. Comparison of the reported results for both timetables also suggests that advantages of each could maybe be combined.

Peer-review under responsibility of Delft University of Technology
doi: 10.1016/j.trpro.2015.09.016.

Keywords:
Comparing Cyclic Timetables, Expected Passenger Time, Objective Function Evaluation, Periodic Event Scheduling Problem (PESP)

* Corresponding author. Tel.: +32-486-956797.
E-mail address: sels.peter@gmail.com

Peer-review under responsibility of Delft University of Technology
doi: 10.1016/j.trpro.2015.09.016.

1. Introduction

Whenever a railway company updates or completely overhauls a timetable, it is highly important to evaluate the new timetable and compare it to the previous one, before the new timetable is implemented in practice. Today, evaluation methods are typically restricted to simulation methods that do not report expected passenger time but mainly focus on train time, ignoring passenger numbers in ride, dwell and transfer actions and in knock-on delays (Weston et al., 2006; Parbo et al., 2015). More recently, this shortcoming has been remedied by some (Weston et al., 2006; Kanai et al., 2011). Also, often no measuring of positive effects of temporal spreading of alternative trains is present. So these methods do not answer some relevant questions that should be asked when developing and evaluating a new timetable.

Questions to be asked about the correctness of any timetable are: Are all minimum ride, dwell, transfer and headway times respected? Questions about whether the new timetable is an improvement compared to the previous one are: Is total expected passenger time in practice reduced? Is the average probability for passengers of missing transfers reduced? Are secondary delays for passengers diminished?

In this paper, we provide a methodology and a tool to answer all these questions. The output is presented graphically, so that the effects - both in size and in sign - of differences between timetables become more visually obvious. As such, strong and weak points in the new timetable, for example better transfer planning or a worse headway planning, will be noticed quickly. Our method is applied to two different timetables, both of all passenger trains in Belgium and shows that the newer of the two timetables is indeed improved. We believe our tool is innovative in the sense that it shows how much better or worse a timetable is for passengers and on which aspects. It also pinpoints where expected passenger time is spent: during ride, dwell, transfer or knock-on actions. In an early stage, our evaluation method could indicate where to further improve the timetable and in the end it could deliver crucial arguments to convince anyone about the benefits of implementing the new timetable in practice.

Section 2 compares traditional timetable evaluation methods with our approach. Section 3 demonstrates the results of our evaluation approach applied to two manually constructed timetables, T1 and T2 as received from Infrabel, the Belgian infrastructure management company. T2 is a reworked version of T1. Due to confidentiality issues, we cannot give more details about these timetables. Sections 4 and 5 conclude this work and indicate some potential further work.

2. Comparing Timetable Evaluation Methods

Both our quality criterion for and our practical implementation of evaluating timetables are quite different from the traditional criteria and practical methods. So, in this section, we first contrast and compare them. We then show that our approach has quite some theoretical as well as practical advantages and next, explain that we cover most performance indicators present in the traditional approaches.

2.1. Traditional Approach: Estimating Performance Indicators by Simulation, mainly Microscopic

Goverde and Hansen (2013) propose a list of timetable performance indicators and derived quality levels. Widely used terms for these indicators are realisability, conflict-freeness, stability, efficiency, robustness and resilience. We are convinced that these six performance indicators should actually be considered as means to obtain the final goal of serving the passengers in the best possible way. According to us that corresponds to minimising the expected passenger time in practice. The performance indicators realisability and conflict-freeness can be checked by verifying that the minima for ride and dwell respectively for headways are respected. Together, realisability and conflict-freeness could be seen as operational feasibility. So these two criteria are easy to check. Contrary to these, the four other performance indicators depend on primary delays and on how these are propagated in the train network. Because of the stochastic nature of primary delays, timetable evaluation on these criteria is usually done via simulation, where a single simulation randomly takes one sample for each primary delay distribution and then applies these combined delays on the system and propagates them to secondary delays on affected trains. This way, simulations can calculate total delays. As an example, a low total delay then indicates a good robustness against these primary delays. For large networks, to cover all cases or rather to obtain a representative collection of all cases, this means that many simulations have to be performed.

Barber et al. (2007) give a list of available software simulation packages (e.g.: RailSys of RMCon, SIMONE of NS). More recently, quite some more software packages have become available (e.g.: LUKS of ViaCon, and OnTime of TrafIT Solutions & ViaCon). Some packages simulate on the macroscopic or mesoscopic level, but quite some also simulate on the microscopic level. This means that they will detect conflicts between trains in the timetable, also inside stations, on routings or platform tracks and via the application of time shifts, they then calculate secondary and total delays. These conflicts arise due to primary delays in practice or due to the timetable not being realisable.

2.2. Our Approach: Expected Passenger Time Evaluation and its Advantages

Our evaluation approach is entirely macroscopic. Slightly simplified, it consists of evaluating the expected passenger time in practice for every action in the timetable (ride, dwell, transfer, knock-on delay) independently. This is possible by application of the expected passenger time per action which is an analytic expression of only the time supplement variable to be assigned to that action. For a given timetable, each action's supplement is known, so also its associated expected passenger time is known. The total expected passenger time of the whole timetable is then simply the sum of all expected passenger time of each of its actions. A complete derivation of our evaluation function can be found in Sels et al. (2013b,a, 2015b), where it plays the role of objective function in an optimisation context.

The more global focus on expected passenger time in practice, also shared by Dewilde et al. (2011) instead of the direct focus on stability, efficiency, robustness and resilience has at least three advantages. (i) Since minimising expected passenger time improves stability, efficiency, robustness and resilience, it is a combined performance indicator and the problems with evaluating a timetable over multiple objectives and setting priorities for objectives are resolved. (ii) Communicating that a new timetable has some percentage less expected passenger time in practice than the previous one is a very meaningful and clear way to communicate about timetables to management and passengers. (iii) We can *separately* evaluate the expected passenger time of every planned action: ride, dwell, transfer, knock-on delay, considering only its own primary delays. Indeed, the total expected passenger time in the system is then simply the sum of all expected passenger times of its constituent components. This means that, for evaluation or optimisation, we can decouple all actions and so do not suffer from having to make combinations of samplings of primary delays over all actions. There is one underlying assumption here, and that is that primary delays of different actions are independent. But in fact, that is also what most simulation packages assume, since they allow primary delay distributions to be entered per action (ride, dwell) and in simulations then combine primary delays over different actions. Only some research, like Kroon (2008) has taken great care of not assuming any primary delay dependencies by strictly keeping together all delays as measured in practice of one train occurrence as one 'trace'. Individual delays of different traces of one train are then not combined with each other in any simulation. This means that all unknown delay dependencies existing in reality, if there would be any, are preserved in the simulation or optimisation. The idea of avoiding the problem of finding a representative set of simulation traces by taking a more analytical approach is similar to what is done in the robustness analysis tool PETER, as described in Soto y Koelemeijer et al. (2000).

Note that our evaluation function - expected passenger time in practice - is already present as the objective function of our Periodic Event Scheduling Problem (PESP) Mixed Integer Linear Programming (MILP) model as published in Sels et al. (2013b,a, 2015b). In these papers, the goal was always to automatically find a new timetable that minimises the objective function. However, in this paper, we do not optimise a timetable, but compare two already existing ones. This means that we only need to evaluate the evaluation function for both timetables and see which timetable results in the lowest value to know which one is better for passengers.

2.3. Our Timetable Evaluation Method Contains Most Traditional Performance Indicators

To show that, with our evaluation function, we cover most of the traditional performance indicators for a timetable, and one more, we discuss these performance indicators one by one and relate them to our evaluation function.

2.3.1. Realisability

First of all we evaluate both timetables on *realisability*. This means that we check whether the time planned for each ride and dwell action is at least as large as is minimally required. In other words, ride and dwell actions should be assigned positive time supplements on top of the minima required for them. Our tool performs these checks.

2.3.2. Conflict-Freeness and Stability

Conflict-freeness, also sometimes simply called feasibility, means that in the timetable, without assuming delays, there are no conflicts between trains. On the macroscopic level this means that planned headways, between train departure times on the same resource and similarly between train arrival times, are larger than an assumed safe minimum. A common value for this macroscopic minimum is 3 minutes. Stability is defined as having sufficient time supplements, on top of the headway minimum time, between subsequent trains on the same track section to absorb delays on the first train. To find out where stability bottlenecks occur, this property can be analysed over every *cycle* in the *event activity graph* corresponding to the timetable (Soto y Koelemeijer et al., 2000). The event activity graph has a node for every train arrival or departure time in a station, representing an event and has an edge for every action (ride, dwell, transfer, headway) between two events representing an activity. In our objective function of expected passenger time, we do not explicitly check every *cycle* for this quality criterion, but for every inter-train edge (knock-on or transfer edge), we penalise a supplement that is chosen too low via the expected passenger time of this edge. This guarantees that every cycle will have a reasonable amount of supplement for it to become stable. This means that we do not only answer the question if every cycle has enough supplements present on its edges but also obtain a global score for how bad the stability is in terms of the summed expected passenger knock-on time and transfer time. If for some headway edges, the timetable would not be feasible, our objective function will give a very high penalty in terms of expected knock-on time. So in our evaluation method, both conflict-freeness and stability are handled by the terms of expected passenger knock-on delay. Note that, since we treat transfer edges similarly to knock-on edges, also with a minimum of three minutes but with a different evaluation function, we also obtain a feasibility for transfers which we could call *transfer guarantee* and also obtain a *transfer stability*.

2.3.3. Efficiency versus Robustness

The total expected passenger time in practice is also the property we use to evaluate both efficiency and robustness together. Very low supplements will make the timetable efficient for passengers, but not robust against primary delays so passenger travel times will vary. Very high supplements wil make the timetable very robust but not efficient. In this sense, we think it is not very practical to try to measure efficiency and robustness separately. Our method makes the trade-off in a passenger weighted way, because both efficiency and robustness are more important for trains with many than with few passengers. As for robustness against potentially occurring primary delays, we assume the primary delays on ride, dwell and transfer actions to occur according to a negative exponential distribution, as do Meng (1991); Goverde (1998); Vansteenwegen and Van Oudheusden (2006). We further assume the average 'a' of these distributions to be a certain percentage of the minimum time of that action. The value of 'a' is typically chosen in the range of 1% to 5% Goverde (1998). In this paper, we assume $a = 2\%$.

2.3.4. Temporal Spreading of Alternative Trains

In Sels et al. (2015a), it is shown how the inter-departure waiting time between alternative trains from a passenger's first station to this passenger's arrival station depends on the temporal spreading of these trains' departure times. So, this can be different in different timetables as well. Sels et al. (2015a) also indicate, that only some categories of passengers benefit from good temporal spreading. We set the percentage of passengers that do benefit to 'r'. Traditional simulators do not report on the level of temporal spreading and to which degree it benefits passengers, because traditionally, spreading is either enforced as a hard constraint for all passengers or not at all. However, quite some research papers did concentrate on evaluating the amount of inter-departure waiting time induced by a timetable (e.g.:Zhao et al. (2013)). Our method evaluates temporal spreading with the function described in Sels et al. (2015a).

2.3.5. Resilience

A timetable that is resilient against primary delays and disruptions recovers quickly, in a number of timetable periods, to its planned state without delays thanks to (i) good supplement choices that, to some degree, prevent knock-on delays and (ii) real time interventions of dispatchers (Goverde and Hansen, 2013). This means that, to properly evaluate the expected passenger time associated with resilience, one needs to be able to model both the recovery behaviour and also the precise dispatching rules. This is work that we did not address yet.

3. Results

We now apply the method as described in section 2.2 to the timetables T1 and T2. Both timetables have a planning horizon of 1 year. They each contain around 200 hourly passenger trains. In total, for about a 1000 physical points, being stations and stop places, train arrival and departure times are specified.

3.1. Realisability

Both in timetable T1 and in timetable T2, we find some cases where the process times assigned to ride actions are smaller than the required minimal run times. Timetable T1 contains 823 such cases and timetable T2 reduced this amount to 570 cases. Table 1 shows the size and amounts of these violations for both timetables. For timetable T1 there are 320 ride edges where the assigned process time is 6 seconds lower than the minimum allowed. The total violation time for all violations together for timetable T1 is 1909 times 6 seconds, or 190.9 minutes. For timetable T2 this is 1084 times 6 seconds or 108.4 minutes. Of course, the degree to which this is problematic is proportional to the number of passengers on the ride actions where these violations occur.

Table 1. Realisability. Reduction of the number and size of minimum runtime violations from timetable T1 to timetable T2.

timetable	distribution: # actions with a violation per size of violation in seconds											weighted sum (s)	tot.#	avg. (s)
	6s	12s	18s	24s	30s	36s	42s	48s	56s	60s	66s			
T1	320	219	126	93	24	27	3	6	1	3	1	11454	823	13.9
T2	277	155	84	37	11	2	2	2				6504	570	11.4

So realisability in terms of both number and size of violations improved from timetable T1 to T2. In practice, a non-realisable timetable will not cause train collisions but will cause primary delays, even in the absence of any other external factors interfering with the train and infrastructure system. Remark that the final goal should of course be to have no such minimum time violations at all in the timetable. Note that realisability is automatically 100% achieved for any timetable that is automatically produced by the optimisation method as described in Sels et al. (2015b).

3.2. Conflict-Freeness, Stability, Efficiency, Robustness & Temporal Spreading via Total Expected Passenger Time in Practice

In fact, the minimum ride and dwell time violations mentioned in paragraph 3.1 should be solved first before one can accurately evaluate passenger time. We achieve this by, before evaluation, correcting the assigned process times in the timetable to the minimum required durations, whenever they are lower. This is also what will happen in practice.

For the case $a = 2\%$, the total expected passenger time for both timetables is shown in figure 1. The left bar indicates timetable T1 and the right bar indicates timetable T2. The vertical dimension represents expected passenger time, also for its constituent components: ride (blue), dwell (yellow), transfer (orange), knock-on (purple). For dwell and transfer time, all ride time of the ride action preceding it, is convoluted with it, which is what the blue shading refers to. On the left of each bar, the percentages (T1.m and T2.m) indicate the ratio of the total expected passenger time due to the minima (m) to its total bar height. On the right, the percentages (T1.s and T2.s) indicate the ratio of the total expected passenger time due to supplements (s) to its total bar height. For each color, the minima are shown in a darker tone of the color and the supplements in a lighter tone of the same color.

For the case $r = 0\%$, we see that the total time reduction between timetables T1 and T2 amounts to 2.47%. This is a considerable improvement. This 2.47% reduction can be attributed for 70.3% to better timetabling and for 29.7% to a lower sum of minimum ride and dwell times. This latter part of the reduction is caused by the also changed line planning associated with T1 and T2. This changed line planning can cause that more passengers could choose a route with fewer ride and dwell actions or now even have a direct connection without transfer from their origin to

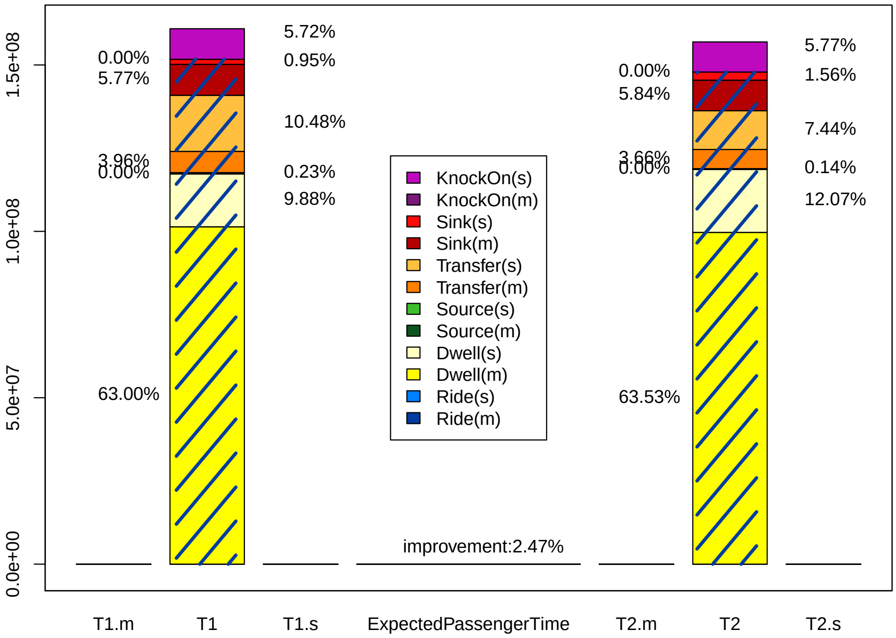

Fig. 1. Reduction of 2.47% of total expected passenger time from T1 to T2. All time units are in 6 second multiples.

destination. Other causes of the decreased minimal ride and dwell passenger time could be the reduction of run time minima by use of faster train material or the assignment of different train material on different lines (e.g. faster trains on lines with more passengers).

Figure 1 also shows that the transfer component in timetable T1 was responsible for 10.48% of the total expected passenger time, while in timetable T2 it only amounts to 7.44%. This corresponds to $1.687 \cdot 10^6$ min and $1.167 \cdot 10^6$ min respectively in absolute numbers, which is a considerable reduction of 31%. This could be due to more people having a direct connection instead of a transfer and also with better supplement planning for transfers. Figure 1 also shows that the total knock-on time did not change much between T1 and T2. It also shows that, in expected passenger time, more time is spent by passengers 'sitting out' these extra supplements. The goal of these extra supplements should be that the timetable becomes more robust and this should be visible in lower expected secondary delays. However, the purple component, representing expected knock-on delay does not decrease from T1 to T2. The expected passenger time spent in the ride and dwell supplements is 9.88% for timetable T1 and increases to 12.07% in the timetable T2. This larger expected time is caused by, on average over passengers, larger *planned* ride and dwell supplements and is intended to make the timetable more robust. If this higher robustness was obtained, it should be visible as a decrease of expected knock-on delay time that at least offsets the increased expected passenger dwell and ride time. However, as mentioned, the expected knock-on time does not decrease but remains roughly the same. Timetable T2 had as one of the major goals the improvement of punctuality over timetable T1. To obtain this, larger ride and dwell supplements were inserted. In practice, during implementation of timetable T2, it was indeed noted that punctuality, as measured by Infrabel has risen with a few percentage points compared to during implementation of timetable T1. This could be expected from the larger inserted ride and dwell supplements, but not from the expected knock-on delay which stays similar from T1 to T2. Note that Infrabel's punctuality measure currently only represents measurements at end stations of train lines and in Brussels, when these trains are at least 6 minutes late. It also does not weigh trains delays with affected passenger numbers. This could explain different findings between model and reality.

Using the method described in Sels et al. (2015a), we also evaluated both timetables on inter-departure waiting time between alternative trains and found that there is no considerable difference between these times for both timetables. This means that, when we add this time component to the total expected passenger time, the total reduction percentage will be somewhat smaller than 2.47%. However, we prefer to postpone detailed reporting of total expected inter-departure and arrival time until it is clear what the value of '*r*' is in practice.

3.3. Expected Passenger Transfer Time Improvement

Because the expected transfer time for all passenger together decreased considerably form timetable T1 to timetable T2, we want to investigate how this was achieved in some more detail. We do this graphically in figure 2 by plotting different histograms of transfer times, on the top row, per transfer action, on the middle row, per transfer passenger in the planned domain and on the bottom row, per transfer passenger in the expected time domain. The left half of figure 2 again represents results for timetable T1 and the right half for timetable T2. The top row shows that there is little variation in planned time (minimum + supplement) assigned to transfer edges. The distribution is indeed quite uniform. Note that the total number of planned train time is increased from timetable T1 to T2 by $(668115 - 573426)/573426 = 16.5\%$. One should not conclude yet that this is a deterioration, since, for all time components, the *expected passenger* time domain instead of planned train time domain is the domain we should evaluate a timetable in and these results are shown in the bottom row.

The middle row shows that, when looking at the planned time assigned per transfer passenger, a pattern starts to appear. More specifically, transfers taken by many passengers tend to be assigned 'good' supplement values. 'Good' here means that the chosen supplements, together with the transfer minimum time, separate the arrival time of the feeder train and the departure time of the connecting train, are neither too small and neither too large. Too small would lead to a frequently missed transfer and our model penalises this with a waiting time of 1 hour for the next connecting train. Too large would mean that too much time has to be spent waiting at a platform for the connecting train. The middle row of figure 2 also shows that this good pattern is even better in timetable T2 than in T1. Indeed, the average planned passenger duration went down from 8.02 minutes to 7.76 minutes. This decrease of total planned passenger time (middle row) associated with an increase in total train time (top row) shows how important it is to weigh with passenger numbers. It also demonstrates that human timetablers also take this into consideration. The bottom row of figure 2 shows that in the *expected* time domain, the pattern is even more pronounced and the expected average transfer duration per passenger goes down from 19.2 minutes to 15.8 minutes. This reduction is a big improvement. The values 19.2 and 15.8 may seem large, but note that (i) this includes *all* possible (passenger weighted) transfers and not just the 'important' transfers from a restricted list and that (ii) our penalty of 1 hour for a missed transfer is a *worst case* one. In line with these findings of improved transfers is that we also calculated that, for $a = 2\%$, the average probability for a passenger to miss a transfer is reduced from 14.41% in timetable T1 to 5.51% in timetable T2. These calculations are based on a similar quick evaluation over all transfer edges, now not of the expected passenger time function per edge, but of the analytical function representing the probability of missing a transfer. The average probability of missing a transfer is then simply the passenger weighted average over all transfer edges.

4. Conclusions

As for our particular timetable evaluation methodology, we have shown that our evaluation function, expected passenger time in practice, has quite some advantages. These are essentially derived from the fact that weighing of different performance indicators is done in a natural way because all constituent components are expressed as passenger time. These advantages are (i) the full timetable evaluation problem can be decomposed in evaluating separate actions, (ii) the evaluation results of actions can be directly added, and ultimately, (iii) simulation is replaced with simple addition.

As for the application on the Belgian timetables, assuming that for each ride, dwell and transfer action, the average primary delay is 2% of the action minimum time, we can conclude that timetable T2 reduces total expected passenger time by 2.47% compared to timetable T1. This is a considerable improvement. This paper also shows that this 2.47% improvement is caused for 70.3% by timetabling improvements, chiefly by improvement of passenger transfer planning. Improved line planning is responsible for 29.7% of the 2.47% reduction. Under the same primary delay

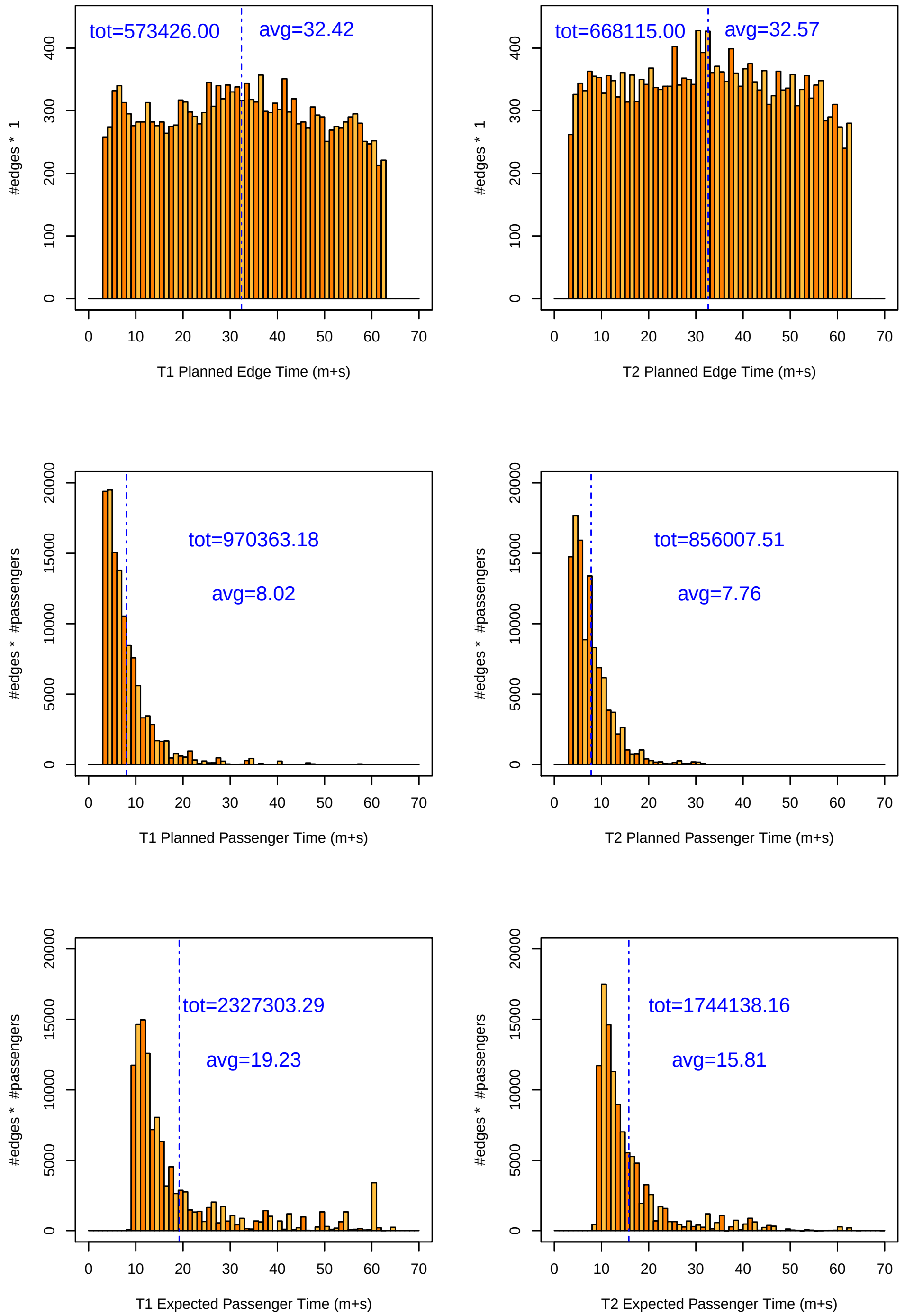

Fig. 2. Transfer time histograms for timetable T1 on the left and timetable T2 on the right. All time units are in minutes.

assumption, the average probability for a passenger of missing a transfer is reduced from 14.41% in timetable T1 to 5.51% in timetable T2. This is a huge improvement. The new timetable T2 has more expected passenger time spent in ride and dwell supplements than timetable T1. So the question arises if, in timetable T2, these supplements can be reduced to the level of timetable T1 while still keeping the advantage of the reduced time spent in transfers of timetable T2.

5. Further Work

As for our method of evaluation, evaluating the same timetables but for different values of the parameter 'a' would be useful. Also, resilience of a timetable is not directly part of our method yet. This would require more research.

Our evaluation is conclusive about timetable T2 being preferable for passengers over timetable T1. When further manual improvement of timetable T2 is required, the question arises how to give the user feedback on how to achieve this. Answering this question is not easy. Indeed, indicating that, for example a transfer is badly planned, can be easily done. How to improve the transfer locally can also be indicated easily. However, making a local change will almost always affect other trains and it is generally hard to know if the potentially negative effect on other trains will not outweigh the desirable positive effect of the improved transfer planning. How to give hints for manual improvement could be considered in further research. However, note that via automatic timetabling optimisation, Sels et al. (2015b) were able to reduce the expected passenger time by 3.81% in 2 hours of computation time, by only changing the timetable and not the line planning. In the resulting timetable, expected knock-on time was decreased by about 50% and expected time of ride and dwell supplements was decreased, but expected transfer time increased.

References

Barber, F., Abril, M., Salido, M., Ingolotti, L., Tormos, P., Lova, A., 2007. Survey of automated systems for railway management. Technical Report. Universitat Politècnica de València.

Dewilde, T., Sels, P., Cattrysse, D., Vansteenwegen, P., 2011. Defining Robustness of a Railway Timetable. Proceedings of 4th International Seminar on Railway Operations Modelling and Analysis (IAROR): RailRome2011, February 16-18, Rome, Italy. .

Goverde, R., 1998. Optimal Scheduling of Connections in Railway Systems. Paper presented at the 8th WTCR, Antwerp, Belgium .

Goverde, R., Hansen, I., 2013. Performance indicators for railway timetables. Proceedings of IEEE International Conference on Intelligent Rail Transportation: ICIRT2013, August 30-September 1, 2013, Beijing, China. , 301–306.

Kanai, S., Shiina, K., , Harada, S., Tomii, N., 2011. An optimal delay management algorithm from passengers' viewpoints considering the whole railway network. Journal of Rail Transport Planning & Management 1, 25–37.

Soto y Koelemeijer, G., Iounoussov, A.R., Goverde, R., Van Egmond, R.J., 2000. PETER, a performance evaluator for railway timetables, in: Computers in Railways VII, WIT Press, Southampton, UK. pp. 405–414.

Kroon, L., 2008. Stochastic improvement of cyclic railway timetables. Transportation Research Part B: Methodological 42, 553–570.

Meng, Y., 1991. Bemessung von Pufferzeiten in Anschlüssen von Reisezügen. Veröffentlichungen des Verkehrswissenschaftlichen Institutes der Rheinischen-Westfälischen Technischen Hochschule Aachen 46, 29–36.

Parbo, J., Anker, O., Prato, C.G., 2015. Passenger Perspectives in Railway Planning: A Literature Review. submitted to Transport Reviews .

Sels, P., Dewilde, T., Cattrysse, D., Vansteenwegen, P., 2013a. Knock-On Delay Modelling and Optimisation: The Passenger Perspective, in: Proceedings of the 3rd International Conference on Models and Technologies for Intelligent Transport Systems (MT-ITS 2013), December 2-4, Dresden, Germany., pp. 1–10.

Sels, P., Dewilde, T., Cattrysse, D., Vansteenwegen, P., 2013b. Expected Passenger Travel Time as Objective Function for Train Schedule Optimization. Proceedings of 5th International Seminar on Railway Operations Modelling and Analysis (IAROR): RailCopenhagen2013, May 13-15, Copenhagen, Denmark. .

Sels, P., Dewilde, T., Cattrysse, D., Vansteenwegen, P., 2015a. Optimal Temporal Spreading of Alternative Trains in order to Minimise Passenger Travel Time in Practice. Proceedings of 6th International Seminar on Railway Operations Modelling and Analysis (IAROR): RailTokyo2015, March 23-26, Tokyo, Japan. .

Sels, P., Dewilde, T., Cattrysse, D., Vansteenwegen, P., 2015b. Reducing the Passenger Travel Time in Practice by the Automated Construction of a Robust Railway Timetable. submitted to Transportation Research Part B .

Vansteenwegen, P., Van Oudheusden, D., 2006. Developing Railway Timetables that Guarantee a Better Service, journal = European Journal of Operational Research 173(1), 337–350.

Weston, P., Goodman, C., Takagi, R., Armstrong, J., Preston, J., 2006. Minimising train delays in a mixed traffic railway network with consideration of passenger delay. Proceedings of the 7th World Congress on Railway Research: WCRR2006, June 4-8, 2006, Montreal, Canada .

Zhao, J., Frumin, M., Wilson, N., Zhao, Z., 2013. Unified Estimator for Excess Journey Time under Heterogeneous Passenger Incidence Behaviour using Smartcard Data. Transportation Research, Part C: Emerging Technologies. 34, 70–88.