Studies on the impact of the fixed-speed approximation for the real time Railway Traffic Management Problem

Pierre Hosteins, Paola Pellegrini, Grégory Marlière, Joaquín Rodriguez

Studies on the impact of the fixed-speed approximation for the real time Railway Traffic Management Problem*

Pierre Hosteins
Università degli Studi di Torino,
Dipartimento d'Informatica, Torino, Italy
e-mail: pierre.hosteins@unito.it

Paola Pellegrini
COSYS-ESTAS, Univ Gustave Eiffel,
F-59650 Villeneuve d'Ascq, France

Grégory Marlière
COSYS-ESTAS, Univ Gustave Eiffel,
F-59650 Villeneuve d'Ascq, France

Joaquín Rodriguez
COSYS-ESTAS, Univ Gustave Eiffel,
F-59650 Villeneuve d'Ascq, France

Abstract

The real-time Railway Traffic Management Problem is the problem of selecting train routes and schedules to minimize delay propagation in case of perturbation. A frequent approximation used to solve it is named fixed-speed, where the effects of accelerations and decelerations resulting from train rerouting or rescheduling are neglected. In this paper, we assess for the first time the ability of this approximation to drive the search of optimization algorithms on large infrastructures with numerous trains. This is done through a statistical analysis on a number of perturbed scenarios on different railway control areas, for various objective functions commonly used in the literature. For each scenario, we compare the ranking of hundreds of solutions, assessed, on the one hand, with the fixed-speed approximation and, on the other hand, through the variable-speed dynamics obtained through simulation. Our results indicate that the fixed-speed approximation is able to capture the relative solution quality with all considered objective functions, although some differences exist. This justifies the use of the fixed-speed approximation in optimization algorithms, to obtain practically reliable results and exploit its higher efficiency compared to more sophisticated approximations methods. We also assess a modified fixed-speed approximation that better reflects the behavior of train speed dynamics, but no remarkable difference is observable as for the quality of the approximation.

Keywords: Traffic Management, speed profile, fixed-speed approximation, Mathematical Programming, optimization.

1 Introduction

At peak times, in critical parts of the railway network of many European countries, traffic is planned to occupy the infrastructure close to maximum capacity. When this happens, a delay of one train, even of a few seconds,

*This paper has been published: Pierre Hosteins, Paola Pellegrini, Grégory Marlière, Joaquín Rodriguez, Studies on the impact of the fixed-speed approximation for the real time Railway Traffic Management Problem, Journal of Rail Transport Planning & Management, Volume 35, 2025, 100535, https://doi.org/10.1016/j.jrtpm.2025.100535.

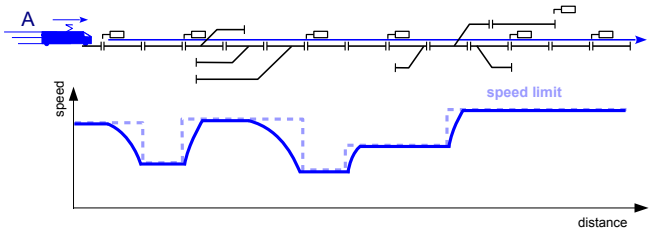
may propagate to several other trains in a snow-ball effect: if two or more trains, running at the planned speed, would require the same railway track concurrently, all but one of them must slow down or even stop to ensure safety. In this case, a *conflict* is said to emerge. Conflicts are particularly critical at *junctions* and stations areas, where multiple lines cross. Here, the precedence between the involved trains must be specified and may have a strong impact on delay propagation. Moreover, it is often possible to route trains in different ways to go through a junction, also impacting delay propagation. The definition of suitable precedences and routes translates into a difficult combinatorial optimization problem.

Today, traffic management is performed mostly manually by dispatchers. Several algorithms have been proposed to solve the described routing and scheduling problem (Cacchiani et al., 2014), which is often named real-time Railway Traffic Management Problem (rtRTMP). Many variants exist to tackle such a problem, be it on the modeling side or on the methodology for recovering normal operation. For example, such models advocate a macroscopic or mesoscopic infrastructure representation, ignoring some details such as the locking and releasing of track sections. Conversely, others advocate a microscopic modeling of the infrastructure, taking into account almost all operational details. Different models focus on different aspects of the problem, such as, e.g., trains or passengers, details of the train speed variation dynamics when brakes and accelerations are necessary due to conflicts, etc. The choice of these aspects results in different objective functions to be optimized and operational constraints. There also exists a wide range of algorithms to solve the rtRTMP. Exact approaches usually make use of commercial branch-and-cut solvers to solve a Mixed Integer Linear Program (MILP), as in e.g. Caimi et al. (2011); Corman et al. (2012); Lamorgese and Mannino (2015); Törnquist and Persson (2007), while a few of them are more specific, such as the cumulative flow variables network model of Meng and Zhou (2014). Heuristic and meta-heuristic approaches have also been devised to tackle the problem. Prominent examples are the works of Khosravi et al. (2012); Dündar and Şahin (2013); Sama et al. (2017). More recently, other technologies and machine learning techniques have found their way to railway traffic management. Notably, Huang et al. (2023) propose a data-driven method for the rtRTMP, Agasucci et al. (2023) introduce Q-learning algorithms, and Xu et al. (2023) rely on a transformer-based model.

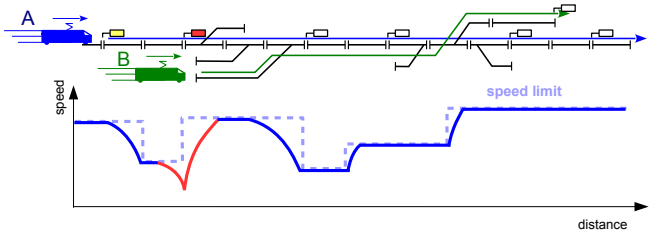
Given the complex behavior of the speed profiles of trains, due to the unexpected brakings and reaccelerations after crossing a restrictive signal, many models need to introduce some form of approximation for such profiles. The impact of the modeling choices on the actual performance of the algorithms due to the quality of these underlying approximations, has not been deeply studied yet. In particular, it is quite unknown whether the solutions selected by optimization algorithms based on these assumptions are of good quality in reality, when such assumptions disappear.

In this paper, we assess the validity of the assumption underlying the so-called *fixed-speed approximation* for the unplanned brakes and accelerations , which we define as follows. Fixed-speed approximation is a non-exact approach to define train speed profiles. Here, the speed profiles of trains traveling undisturbed at their planned speed are precisely computed, considering track and rolling stock characteristics. However, when traffic congestion implies the emergence of conflicts and trains need to slow down or stop, the speed profiles are modified assuming an infinite acceleration and braking rate. This brings to stops and re-accelerations that occur in no space and time. We illustrate in Figure 1a the speed profile of a train when it crosses an infrastructure without conflict. In the same figure, we show the speed profiles when a conflict arises and the train has to slow down, using realistic speed dynamics (Figure 1b) and the fixed-speed approximation (Figure 1c). The red part in the speed profile of train A shows where the train speed diverts from its free-network profile due to the conflict with train B, to which the former has to give precedence. While this part displays a smooth change in speed due to acceleration and deceleration in Figure 1b, we can see that the speed jumps directly to 0 in the fixed-speed approximation and remains null until the train is free to move again, after which it jumps back to its maximum value over the track section. In the following, the models relying on this approximation are called *fixed-speed models*. Solutions of the traffic management problem computed by fixed-speed models are typically thought to be implemented in reality in terms of train routes and precedences, neglecting passing times. In this sense, these solutions are feasible for real life deployment. However, we observe that the intrincic non-exacteness of the approximation implies a lack of precision of the moments at which the train occupies the different tracks. This implies that the models might either miss conflicts which would actually arise in reality, or even tackle conflicts which would not happen when taking into account realistic speed profiles. Remark that, although this is more evident with solutions computed using the fixed-speed approximation, the same limitation applies when any model is used. Indeed, the speed profiles that will be operated by train drivers cannot be exactly replicated when making decisions, as they are very different from one another and depend on many factors. Hence, independently of the quality of the model considered for computing speed-variation dynamics during

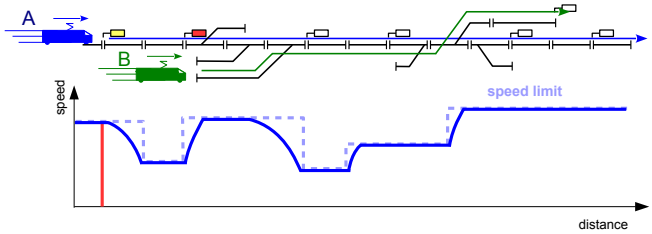
optimization, a perfect correspondence with reality is unachievable. Thus, the only way to guarantee solution applicability is by accounting for routing and precendences and not precise timings.



(a) Free network



(b) Conflict with train B (green) with realistic speed dynamics



(c) Conflict with train B (green) with fixed-speed approximation

Figure 1: Speed profiles of train A (blue) in three cases (from top to bottom): no conflict, conflict with train B (green) with realistic speed dynamics and fixed-speed approximation. The curves represent the value of speed as a function of distance traveled. Dashed lines represent the speed limit on each track section.

Many existing optimization algorithms for the rtRTMP are based on the fixed-speed approximation, as e.g. Corman et al. (2010); Pellegrini et al. (2014). The assumption behind this modeling choice is the preservation of the actual relative quality of solutions: if one takes two solutions α and β, then if the routing and scheduling decisions in α are better than those in β according to the fixed-speed approximation, they will be better also in reality. Although the intuition suggests that this is likely to be true most often, no deep analysis has been performed so far to support this intuition. To the best of our knowledge, the only attempt to quantitatively study the impact of the fixed-speed approximation is the work of Sobieraj et al. (2011), which proposes experiments to characterize relations between some specific traffic conditions and the preservation of the relative solution quality. A form of comparison between the maximum and total delay obtained by respectively a fixed-speed model and a model exploiting realistic speed variation computation was attempted in D'Ariano et al. (2007b). However, the fact of considering two different models does not allow the actual assessment of the impact of the fixed-speed approximation: the fact that a model outperforms the other may be due to the quality of the models themselves rather than on the validity of the approximation. Some papers exist which propose an iterative loop solving a fixed-speed model and then computing the realistic speed profiles of the current solutions (D'Ariano et al., 2007b; Mazzarello et al., 2007; Lüthi, 2009). Few other approaches consider realistic speed profiles and must then deal with exponential numbers of alternatives (Caimi et al., 2011). Some of these works are focused on the reduction of power consumption on the railway network, which is directly linked to the trains' speed and accelerations-decelerations. Examples of such energy consumption oriented works can be found in D'Ariano and Albrecht (2006); Albrecht (2009). Recently, several works appeared which try to take into account the possibility of controlling the speed profile of the trains in the optimization procedure. The work of Zhou et al. (2017), for example, addresses the simultaneous optimization of train timetables and speed profiles to account for tight power supply and temporal capacity constraints. The problem is solved through the use of a multi-flow network model with multi-commodity variables, built upon a discretized space-time-grid network, where sub-problems are solved by Dynamic Programming (DP). Another work considers train disturbances and possible rescheduling. It is the one of Xu et al. (2017), where a MILP is provided to solve the problem on a Chinese infrastructure with a quasi-moving block system. It discretizes the possible values for the speed of the trains on the relevant portions of the railway. This is particularly relevant for Chinese high-speed trains due to the quasi-moving block convention where the number of block sections to reserve in front of a given train depends on its actual speed. The model allows the authors to consider a specific type of disruption, in which temporary

speed limit is put in place for a limited duration. The two works of Luan et al. (2018a,b) are concerned with a very similar problem where the non-linear dependence of running times and energy consumption on the trains' speed is modeled through a Mixed Integer Non-Linear Program (MINLP), which is linearized through the use of piecewise affine functions. An alternative model is also obtained by selecting a subset of appropriate possible speed profiles through a specific preprocessing method. Specific two-step approaches are proposed to solve the different models and compare them. The models are able to find interesting solutions in under three minutes to help reschedule trains after a disruption. Simultaneous optimization of train scheduling, routing and speed profile is also the object of the work of Long et al. (2023). Here, the authors introduce a MINLP for the problem, and a MILP reformulation which optimizes speed profiles approximated as piecewise linear functions. This reformulation is shown to provide solutions for instances representing networks that include many small stations and traffic running in a single direction. On the other hand, the work of Reynolds and Maher (2022) proposes a data-driven approach to account for realistic speed profiles in the rtRTMP. Specifically, the possible speed profiles are inferred from historical observations. The proposed data-driven approach results in a parsimonious model that is less onerous than existing physics-based approaches, where speed profiles need to be computed depending on the chosen routes and precedences. A basic assumption of the approach is that all pertinent speed profiles occur in historical data, and no bias is introduced by the choices of dispatchers who manually managed traffic that generated those data. Finally, other works propose simpler, approximate modeling approaches to account for the additional stops or decelerations of the trains when solving conflicts. Rodriguez (2007) proposes adding extra time to the running time of the next track-circuit, i.e., track segments where trains are automatically detected. The time is determined linearly based on the duration for which the train is hindered in the current track-circuit by the preceding train, as described by the blocking time model. For single-track lines, Krasemann (2015) suggests adding a fixed amount of running time to the track sections near stations. These forfeits approximate the additional braking and re-acceleration times of the trains and are derived from empirical data and runtime profiles in the timetable. Similarly, Bach et al. (2019) defines additional running times for single-track sections, accounting for each additional stop at connected stations. The extra times depend on the characteristics of both the trains and the track sections. All these works show possible directions to integrate realistic speed profiles in the management of disturbances on railway networks. However, most state-of-the-art approaches for the rtRTMP, especially when also re-routing is taken into account, are based on fixed-speed models (Versluis et al., 2024).

In the following, we will assess statistically, on general grounds, the validity of the fixed-speed approximation on different control areas , i.e. portions of railway networks where train traffic is controlled by one dispatcher, and for different objective functions. Indeed, all solutions found are feasible from the point of view of the model constraints and, as mentioned, can be practically implemented by considering their routes and precedences. However, the fixed-speed approximation implies that the value of the train delays and therefore the value of the objective function differs from the one that will be observed when implementing the solution in reality. In a thorough experimental analysis, we consider large sets of traffic management solutions and we evaluate the delays of the trains in each of them, considering, on the one side, the fixed-speed approximation. On the other side, we make the parallel evaluation considering realistic speed profiles. To do so, we run a microscopic traffic simulation imposing the routing and precedence relations extracted by each solution, and we observe the resulting delays. We then compare both results to evaluate whether the fixed-speed approximation allows us to actually find the optimal solution of the true problem. The contributions of the paper are therefore the following:

- We generate a large number of routing and scheduling solutions for the rtRTMP on different railway infrastructures and evaluate them using both a MILP with fixed-speed approximation and a microscopic railway simulator.
- We provide a statistical analysis over such solutions to assess whether the fixed-speed approximation allows us to single out the best solutions for the rtRTMP with optimization models.
- We find out that, in the above sense, the fixed-speed approximation is valid on a variety of infrastructures with different kinds of trains behavior and speed (with or without stops at stations, terminal station,...).
- We highlight the large discrepancies that can be introduced in the trains total travel time because of the blocking time model due to a forced delayed entrance on the infrastructure.

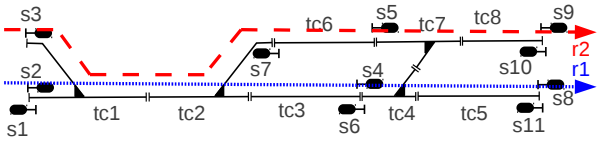


Figure 2: Example of microscopic representation of a railway infrastructure.

We stress that the focus of our work is not to provide specific models or algorithms to improve the fixed-speed approximation but, rather, to assess the validity of the approximation itself, which is scarcely done in the existing scientific literature. Indeed, the few existing works on the subject never considered a large infrastructure like those considered in this work with a large number of interacting trains. Rather, they focused on a small junction, see e.g., Sobieraj et al. (2011). Therefore, we propose the first systematic statistical study on the quality of the fixed-speed approximation.

The rest of the paper is organized as follows. In Section 2, we report a fixed-speed formulation of the rtRTMP and introduce a refined model to try to take into account the effect of deceleration on running times. In Section 3, we detail our methodology to assess the quality of the fixed-speed approximation by comparing its results to simulation ones. In Section 4, we present the different control areas that we use for our numerical experiments and then discuss the results of our analysis. Finally, we conclude in Section 5.

2 Integer Linear formulation for the rtRTMP

2.1 Formulation for the classic rtRTMP with fixed-speed approximation

To formally define a fixed-speed model for the rtRTMP, we refer to the RECIFE-MILP formulation presented in the previous works of Pellegrini et al. (2014, 2015). We recall the corresponding MILP formulation in Appendix A. The RECIFE-MILP algorithm based on the solution of this formulation is typically considered as representative of the state of the art for the rtRTMP. The formulation models the infrastructure at the microscopic level and implements the route-lock sectional-release interlocking system (Pachl, 2008). The tracks are divided into *track-circuits*. *Block sections* represent sequences of track-circuits whose access is controlled by a signal. Moreover, before a train can *occupy* a block section, all its track-circuits must be reserved for the train itself. Figure 2 depicts an example of this microscopic model of the infrastructure for a simple control area. Track-circuits are named tc and signals are named s, both indexed with a progressive number. Signals indicate the availability of block sections in a precise direction: for example, signal s2 is for block sections s2-s4 including tc1, tc2 and tc3, in this order, and s2-s5 including tc1, tc2 and tc6, in this order. Suppose that two trains (t1 and t2) cross the control area: t1 going from s2 to s8 (using route r1 including block sections s2-s4 and s4-s8) and t2 going from s3 to s9 (using route r2 including block sections s3-s5 and s5-s9). We will see in the following how their passing through track-circuits is represented in the model.

Depending on the objective function used, following the literature, the model assesses either overall or consecutive delay. The former, simply named *delay* from now on, is the delay suffered by a train when exiting the control area: it is the difference between the exit arrival time and the one scheduled in the timetable. Let us remark that if a train reaches its destination within the control area itself, it will still be considered as if it was exiting. The *consecutive delay* is the delay cumulated in the control area due to traffic. For example, if a train enters the control area with five minutes of delay and it exits five minutes later than planned not encountering traffic, then its delay will be of five minutes and its consecutive delay will be null. Instead, if within the control area the train has to give precedence to others and it cumulates two further minutes, it will have a delay of seven minutes and a consecutive delay of only two.

The fixed-speed approximation is used in delay definition, in Constraints (5) of the model in appendix. Indeed, in this model, delays can have any value: there is no relation with the duration of deceleration and acceleration. Moreover, no link is set between the delay of a train on different track-circuits, while decelerations and subsequent accelerations due to conflicts involve sequences of track-circuits. Another aspect on which the fixed-speed approximation has an impact is the behavior in case of restrictive signal aspect. In reality, a train starts braking when it meets a restrictive signal, and accelerates when the restriction is removed and the next

signal is opened. In this model, we apply the blocking time model (Pachl, 2008), e.g. in Constraints (12): the utilization of a track-circuit by a train starts shortly before the train itself enters the reference track-circuit. The latter is the first track-circuit of the $n-2^{\text{nd}}$ preceding block section, if it is opened by a signal with n aspects. As a consequence, the train cannot enter this $n-2^{\text{nd}}$ preceding block section as long as the considered track-circuit is utilized by another train. This simulates the fact that trains can only enter block sections opened by green signals whereas in reality if $n > 2$, a train can enter a block whose entry signal has a restrictive aspect, for example the yellow aspect for 3-aspect signalling. By doing so, the model imposes a train distancing that is longer than the one that can be kept in reality. To the best of our knowledge, all microscopic approaches based on a fixed-speed model implement the blocking time model, to somehow balance the absence of constraints on delays. In the example of Figure 2, supposing t1 precedes t2 and all signals have three aspects, the former will be able to reserve tc1, tc2 and tc3 only after the latter leaves tc2. It is at this time that t1 can enter the preceding block section. In the subsequent experimental analysis, we will exemplify how the blocking-time theory can have an important impact on, e.g. the total travel time of a train on a specific infrastructure and is therefore an important aspect of the model to assess and study, in concordance with the fixed-speed approximation.

In our analysis, we consider the following four objective functions:

- the total delays,
- the maximum consecutive delay,
- the number of delayed trains,
- the total travel times of all trains in the control area.

We choose these functions following previous studies on alternative optimization objectives in railway traffic management (Pascariu et al., 2024; Sama et al., 2015), as they are the ones that are most often used in the literature. Let us remark that these measures do not directly consider the number of conflicts encountered during the time horizon considered in the instances but rather the impact of these conflicts on the delays induced on trains.

2.2 Formulation for the modified fixed-speed approximation

To evaluate the simplest existing modifications that partially include the effects of deceleration and re-acceleration due to conflicts (and their possible added value), we consider a different approximation based on a slight modification of the fixed-speed one described above, which we call the *min-fixed-speed approximation* and is very similar in spirit to the modification of Krasemann (2015), adapted to the model previously described. In the fixed-speed approximation, in some situations a conflict might be resolved by stopping a train for a very short amount of time. Instead, in reality the train dynamics will have to follow a deceleration and re-acceleration pattern which would generate a larger delay. For example, if we focus on the example in Figure 1, when assessed with the fixed-speed approximation, train A has to shortly stop to let train B go first in the common track section, and it can start travelling again at the planned speed right after (figure on the third row, second column). Any stop duration is possible for the fixed-speed model, provided that it is compatible with the blocking time model and the fact that train B goes first. When realistic speed dynamics are considered, though, train A has to brake to be able to stop at the end of the block section it enters with the yellow signal. It can accelerate again when encountering the next signal in a yellow or green state after train B has released one or two downstream blocks respectively. This holds independently of the fact that train B has occupied the block section for one additional second or for one additional minute after train A has passed the yellow signal. As it can be observed in the figure on the second row and second column, the speed is reduced, either for deceleration or re-acceleration, for a longer time than it happens in fixed-speed. To somehow model this, inspired by the model of Rodriguez (2007) and Krasemann (2015), we introduce a minimum *forfait* f for the stopping time of a train: when a train has to stop due to a conflict, we impose that it stays on its current track-circuit for at least f additional seconds. The value of f is a parameter that might depend on the control area and the type of train considered.

To account for this minimum stop duration, we introduce new binary variables to the model described in Appendix A: $\sigma_{t,r,tc} = 1$ if train $t \in T$ needs to stop on track-circuit $tc \in TC_t$ on route $r \in R_t$. We add the

```
1: for d = 1, ..., D do                          ▷ Loop through the number of daily perturbed scenarios desired
2:     daily perturbed scenario ← randomly perturbed timetable
3:     for p = p_1, ..., p_P do                  ▷ Loop through the starting times of the scenarios desired
4:         perturbation scenario ← trains entering the control area between times p and p+duration
5:         for s = 1, ..., S do                  ▷ Loop through the number of solutions desired
6:             generate solution as in Figure 4a
7:         end for
8:         generate optimal solution
9:         discard equivalent solutions
10:        for all setting of the interlocking system do
11:            for all solutions do
12:                compute simulation, fixed-speed and min-fix-speed assessment (Figure 4b)
13:            end for
14:        end for
15:        Analyze results and aggregated results as in Figure 5
16:    end for
17: end for
```

Figure 3: Pseudo-code of the methodology employed

following constraints to the fixed-speed model described in Appendix A:

$$l_{t,r,tc} \leq M\sigma_{t,r,tc}, \qquad t \in T, r \in R_t, tc \in TC_t, \tag{1}$$

$$l_{t,r,tc} \geq f\sigma_{t,r,tc}, \qquad t \in T, r \in R_t, tc \in TC_t. \tag{2}$$

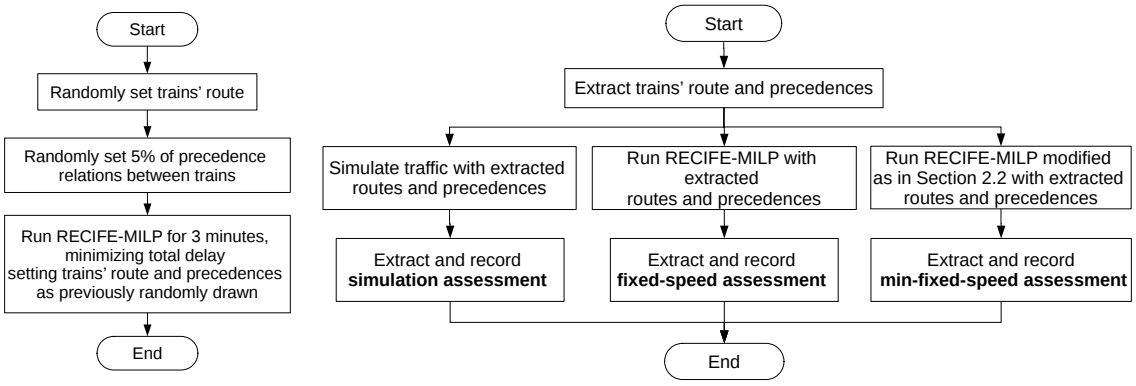
Hence, when variable l must be greater than 0 on a certain track-circuit due to some conflict, the corresponding binary variable σ is equal to 1 due to Constraints (1), thus forcing l to be at least equal to the forfait value f through the use of Constraints (2).

3 Methodology

We now present the methodology chosen for our numerical analysis. The aim of this analysis is understanding if optimizing traffic management considering a model based on the (min-)fixed-speed approximation is equivalent to do so considering realistic train speed dynamics in case of conflict. As a perfect equivalence is very likely impossible, we aim to measure the correlation between the "goodness" of a solution computed using the fixed-speed approximation and a simulation tool. In other words, given two solutions α and β, we wish to understand if, or how likely, α better than β under the approximation corresponds to α better than β in simulation. Indeed, the two solutions are intended as having the same routes and precedences when assessed with the approximation or the realistic speed dynamics, but different passing times of trains along track-circuits if conflicts emerge.

To obtain a general conclusion, as far as possible, we consider different control areas and different objective functions, as reported in Section 2. For each of them, we run an equivalent experimental analysis, as described hereinafter and schematized in the pseudocode in Figure 3.

We start by generating a large set of solutions to get a statistically relevant sample. First, we consider a real one-day timetable, and we randomly select a number of trains. We assign them a delay at the entrance of the considered control area (Figure 3 line 2). We repeat this procedure D times, thus obtaining different *daily perturbed scenarios*. Second, from each daily perturbed scenario, we extract a predefined number P of *perturbation scenarios* (or simply *scenarios*) by considering all trains entering the infrastructure in time intervals with pre-determined start and duration (Figure 3 line 4). Third, for each perturbation scenario we generate a set of different *solutions*, in terms of train routes and precedences on common tracks (Figure 3 line 6). To generate a solution, we apply the procedure reported in Figure 4 (a). Specifically, we randomly fix the route of all trains, among all available routes connecting their origin to their destination in the control area and that may be used by the train. Then, we randomly set 5% of precedences (y variables in in Appendix A). The remaining precedence decisions are made by solving RECIFE-MILP minimizing total delay. To do this, we use CPLEX 12.8 and we set a 3-minute time limit. By repeating this procedure, we generate $S = 500$ solutions. In addition,



(a) Solution generation

(b) Solution Assessment

Figure 4: Solution generation and assessment procedures.

we generate the optimal solution minimizing total delay (Figure 3 line 8), by running RECIFE-MILP with no time limit and no pre-set variables. Adding the optimal solution acts as a point of comparison for the rest of the solution set. From these 501 solutions, we discard those featuring equal routes and precedences for all trains (Figure 3 line 9). Through this procedure, thanks to the random selection of routes and precedences, we obtain solutions that are diversified enough to be representative of the search space available for an optimization algorithm. By applying RECIFE-MILP to set the remaining precedences, we obtain solutions that we expect to be of good quality, at least to some extent. Undoubtedly, this procedure introduces a certain bias, which reflects the fact that we wish to account for solutions that may be generated by a sophisticated optimization algorithm, which will ideally explore good quality regions of the search space. Moreover, let us remark that the same procedure may be repeated by modifying the portion of precedences set and the RECIFE-MILP time limit. By doing so, the number of instances may be increased and specific types of solutions may be generated. This may be particulary relevant, for example, to study the quality of the (min-)fixed-speed approximation in very particular situations.

We then assess the obtained solutions for two settings of the interlocking system (Figure 3 line 10). Indeed, as the modeling of the interlocking system can have a significant impact on train stopping and travel times, we consider the classic model described in Appendix A and a model in which each signal has only two aspects. Here, trains will not reserve the use of several track-circuits in advance. In particular, for each of the solutions, we apply the procedure shown in Figure 4 (b) (Figure 3 line 12) . We begin by extracting the train routes and precedences. Then, we compute the *fixed-speed assessment*: it is the objective function value obtained by running RECIFE-MILP and setting all routes and precedences between trains as extracted. Similarly, we compute the *min-fixed-speed assessment* by running RECIFE-MILP modified as reported in Section 2.2. Finally, to obtain the objective function value corresponding to realistic train speed dynamics, which we call *simulation assessment*, we simulate traffic imposing the extracted routes and precedences with OpenTrack (Nash and Huerlimann, 2004). Here, using the default setting of the simulator, trains run as close as possible to their planned speed unless they meet restrictive signals. In this case, the behavior to be held in reality with the fixed-block interlocking system is reproduced: if a train needs to give precedence to another one in a block section, the signal opening it will be red, the precedent one will be yellow, and so on if the system has more than three-aspect signals (Pachl, 2008). A train crossing a yellow signal will have to start braking to be able to stop by the next signal.

Then, for each perturbation scenario, we perform the analysis depicted in Figure 5 (Figure 3 line 15). For both the (min-)fixed-speed and the simulated assessments, we compute both the order and the rankings of solutions. When a solution has better objective function value than another, the former is assigned lower order or ranking than the latter. When the objective function values are equal, different orders are randomly assigned to the solutions; however, they are assigned the same rank. Hence, given four solutions A, B, C and D, with A being the best, B and C being equivalent, and D the worst, they will have order 1, 2, 3 and 4 (or 1, 3, 2 and 4 depending on the random drawing) and ranking 1, 3, 3 and 4. Hence, solution order is a biunivoque

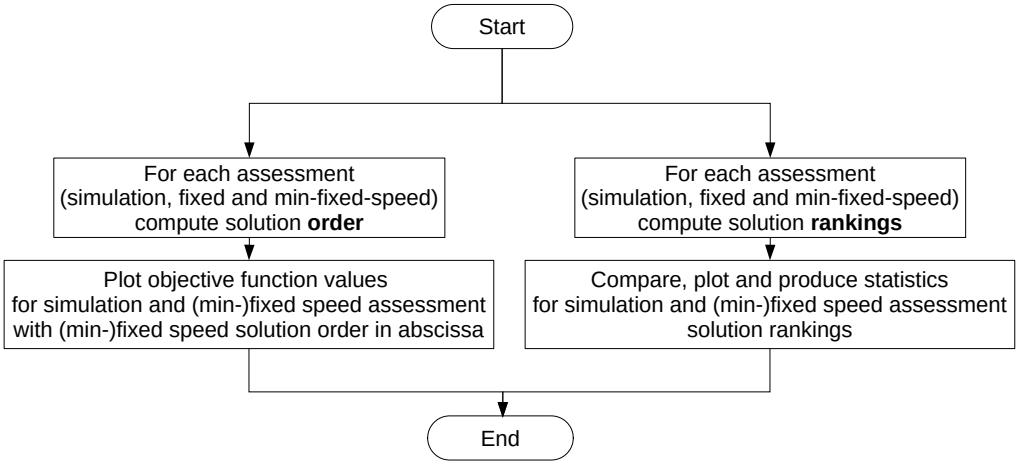


Figure 5: Flow chart representing the methodology used in the paper.

relation, while ranking is not. On the one hand, we use the order to visually compare objective function values of (min-)fixed-speed and simulation assessments. To do so, we plot the solution values considering the abscissa of each solution equal to its order according to the (min-)fixed-speed assessment. The order allows the plot to show a "function-like" trend, in which each point has a different abscissa. On the other hand, we use the ranking to run statistics to understand the correlation between assessments. Although the loss of the "function-like" behavior slightly complicates visual appreciations, rankings do not suffer from any bias due to random draws. Indeed, if (min-)fixed-speed and simulation assessments equally rank all solutions, when we represent the solutions in a plane where the x and y-axes are respectively the ranking of each solution according to the two assessments, the points will be located on the diagonal, or around the diagonal for solutions of equal value. If this is the case, the (min-)fixed-speed approximation will have no impact on the actual quality of the scheduling and routing strategies selected in optimization. To understand how far reality is from this ideal situation, we quantify the departure from the diagonal by performing a linear regression and by extracting the linear coefficient and the correlation coefficient of the different rankings. This is the Pearson correlation coefficient (Akoglu, 2011). The coefficient is between -1 and 1: a null value indicates that there is no correlation between the two variables analyzed. A positive value indicates a positive correlation: as the value of one variable increases, so does the value of the other variable. Conversely, a negative value shows a negative correlation. The interpretation of the value of this coefficient is not unanimous for the whole range $[-1, 1]$. However, it is widely accepted that values higher (in absolute value) than 0.8 indicate very strong correlation. For values between 0.6 and 0.8, the literature interprets the indication as moderate to very strong correlation, while for the range $[0.4, 0.6]$ as fair to strong (Nettleton , 2014).

We also perform an aggregation of solutions which have very close objective values, in order to avoid the artificial discrimination of solutions that are equivalent in practical terms. For each objective function, we choose a threshold value θ which fixes the precision we aim to when aggregating solutions. In practice, we divide the value of the objective function by θ, round the obtained value to the nearest integer and multiply the result by θ. This operation creates more solutions with the same objective function value, which won't be discriminated by the ranking procedure. We fix the value of θ to 100 seconds for the total delay and total travel time objective functions, 10 seconds for the maximum consecutive delay and 1 for the number of late trains. This means that we do not perform any aggregation for the number of late trains as we feel that the difference is already clear enough between two different solutions. Instead, for example, two solutions having a difference in total delay smaller than 100 seconds will be considered equivalent in terms of this objective function, to represent the fact that such difference is irrelevant in practice.

	Pierrefitte Gonesse junction	St. Lazare station	Mantes-la-Jolie–Rouen line
Infrastructure			
Length (km)	15	4,5	80
Routes	37	84	187
Blocks	79	197	157
Track-circuits	89	212	236
Timetable			
Trains/Day	336	1212	237
Routes/Train	7 (5-13)	5 (1-9)	13 (1-24)
Track-circuits/Route/Train	28 (16-34)	18 (13-25)	57 (5-97)

Table 1: Control area characteristics. For the last two rows, we indicate the average and, in parentheses, minimum and maximum values observed.

4 Statistical analysis

In this section, we report the details on the statistical analysis we propose in the paper. In particular, we perform the statistical analysis described in Section 3 for the four different objective functions described in Section 2. We set the penalization weight of the entrance delay $W = 1000$ when minimizing total travel time. Moreover, we test two forfeit values in the min-fixed-speed approximation: 30 and 60 seconds. As the results we obtain with the different values are qualitatively equivalent, we only report those we obtain when setting 30 seconds.

In the following, we first depict the control areas and perturbation scenarios that we consider, and we then focus on the obtained results.

4.1 Control areas and perturbation scenarios

We propose an experimental analysis based on three French control areas: the Pierrefitte Gonesse junction, the Parisian St. Lazare station and a portion of the line between Mantes-la-Jolie and Rouen, connecting Paris and Le Havre. These three control areas have very different characteristics , as summarized in Table 1. The first one is a complex junction about 15 km long where freight, conventional and high speed lines cross. A weekday timetable includes 336 trains. The second one is a terminal station area of slightly more than 7 km, with 27 platforms. A weekday timetable includes 1212 trains, most of which linked by rolling stock re-utilization constraints. The last one is roughly 80 km long, has 13 stations and its weekday timetable has 237 trains. These control areas are modeled at a microscopic level considering the lengths, speed limits, slopes, track-circuit positions, block section and signalling system specificities (signal aspects and visibility, block formation and release time, ...). Rolling stock is also represented in detail at microscopic level within the simulation model (tractive effort curves, mass, running resistance, ...). Simplified maps of these three control areas are shown in Figures 6 to 8.

For each control area, we consider $D = 4$ different daily perturbed scenarios. These scenarios are obtained by randomly assigning an entrance delay between 5 and 15 minutes to 20% of trains (Pellegrini et al., 2014). From each of them, we extract $P = 3$ perturbation scenarios, starting at 6, 7 and 8 a.m. and lasting one hour. Therefore we have a total of twelve different perturbation scenarios for each control area. The number of trains present in these scenarios is between 14 and 27 in the Pierrefitte Gonesse junction, 60 and 105 for the St. Lazare station, and 13 and 19 for the Mantes-la-Jolie–Rouen line.

The distributions of the objective function values of the generated solutions vary across objective function and control area. For example, for Pierrefitte Gonesse, the total delay of more than 7% of solutions is distant from less than 20 minutes from the optimum, with a maximum consecutive delay difference below 3 minutes. For St.Lazare, the difference of total delay with respect to the optimum is around 40 minutes for 0.6% of solutions, corresponding to an almost 10 minute difference in maximum consecutive delay. For Mantes-la-Jolie–Rouen, a difference of up to 20 minutes of total delay concerns 0.9% of solutions, with 6 minutes of maximum consecutive delay. For the three case studies, the total delay of 50% of the solutions is less than 113, 187 and 73 minutes from the optimum, respectively, and it gets to at most 201, 267 and 120 minutes when we look at the best

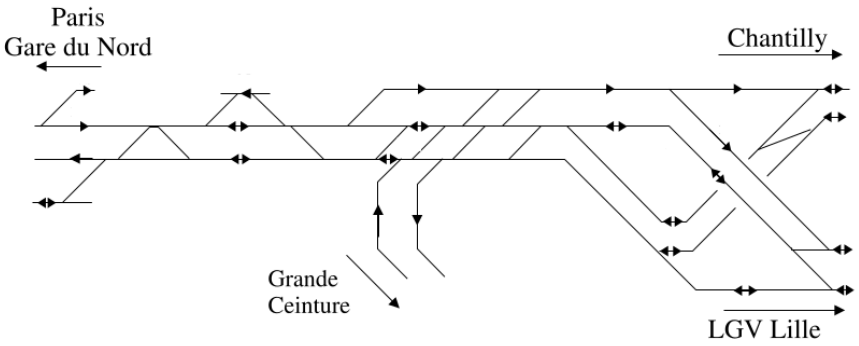


Figure 6: Simplified map of the Pierrefitte Gonesse junction.

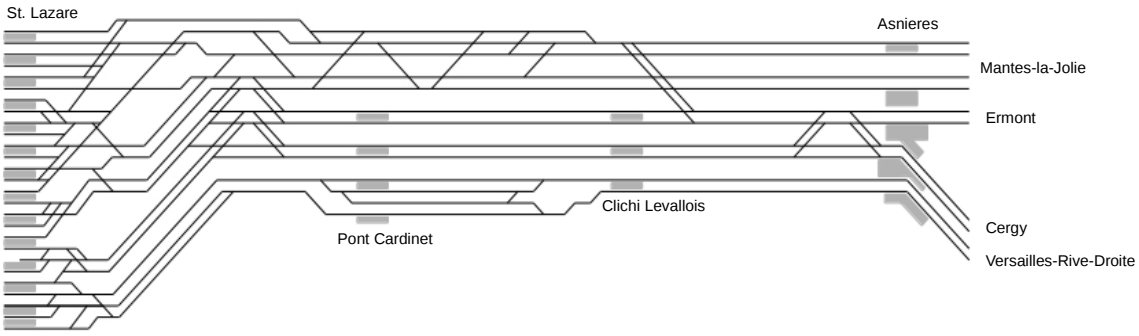


Figure 7: Simplified map of the St. Lazare station.

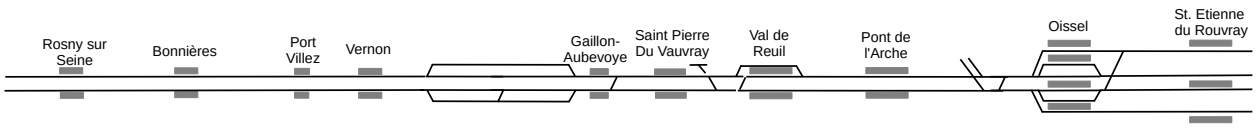


Figure 8: Simplified map of the Mantes-la-Jolie–Rouen line.

80% solution. The corresponding values for the worst generated solutions are 32, 333 and 66 hours. For all objective functions and all control areas, most generated solutions are in parts of the search space that are not immediately next to the optimum, which may be difficult to spot for most algorithms, but that are also very far from very low quality regions. A few solutions, instead, are immediately next to the optimal ones.

4.2 Numerical Results

4.2.1 Pierrefitte Gonesse junction

We first present the results for the Pierrefitte Gonesse junction. We start by providing a graphical illustration of our results on a representative perturbation scenario. Specifically, in Figure 9 (left panels) we report the objective function values for both fixed-speed and simulation assessments for each solution. Four plots are shown, one for each objective function considered in this study. Solutions are ordered according to the fixed-speed ordering, so that the solution with abscissa equal to 1 is the best one according to the particular objective function. In the first two graphs, concerning the total and maximum consecutive delay, the objective function values of about 90% of the solutions are very similar in the two assessments. A bigger difference appears for the worse solutions, which have, for both assessments, objective function values much larger than the rest of the set. The extremely bad quality of these solutions is due to the random choice of 5% of scheduling decisions, which here happen to be extremely penalizing and impossible to compensate by other smart decisions (this happens e.g. when a train which needs to cross a junction early has to wait for much later train). In general, the objective function values computed after simulation appear to be slightly smaller than those of the fixed-speed assessment. This suggests that the impact of the overestimation of train distancing due to the use of the blocking time model, as discussed in Section 2, is stronger than the effect of underestimating braking and acceleration times. In the third graph, which concerns the number of delayed trains objective function, the range of variation of the objective function is much smaller than in the other cases, and this puts more visual emphasis on the differences. As with the two first objective functions, the measure only involves delays, and the fixed-speed assessment ends up overestimating the value of solutions. When measuring total travel time (fourth graph), it is again possible to observe a remarkable similarity of the objective function for good solutions and a larger difference for bad solutions. However, here, the differences between the fixed-speed and simulation assessments are generally bigger. The relation between the two assessments is generally inverted, with respect to total and maximum consecutive delay: the total travel time is typically longer when computed in the simulation assessment than in the fixed-speed one.

When looking at individual scenarios, a large difference can sometimes be observed in the computation of train movements, due to the application of the blocking time model. This imprecision emerges quite often in the Pierrefitte-Gonesse junction, causing the main differences observed. It is due to the peculiar characteristics of the control area, where several junctions are located very close to the boundary, often only after the first block section. As a result, in some solutions, some trains give priority to another train at such junctions, right after their arrival in the control area. In simulation, this corresponds to the trains entering their first block section with a restrictive signal aspect. Instead, in the fixed-speed assessment, trains cannot pass through restrictive signals due to the application of the blocking time model. Hence, they will enter the control area later than scheduled, thereby reducing their travel time on that same control area, sometimes significantly. One such example is illustrated in two space-time Gantt diagram s in Figure 10 for train 848502 (displayed in red).

In each diagram, lines correspond to track-circuits and time evolves from left to right. The focus is on the route followed by the train under analysis: the sequence of track-circuits depicted is the one corresponding to the train's route. The blocking time stairways of all trains using these track-circuits are shown: the light-colored part of the rectangle for a track-circuit stands for reservation, clearing and release time, while the dark-colored one stands for running time. Figure 10a concerns the simulation assessment: the red train enters the infrastructure at time 6:55:05 and waits several minutes before being able to enter the second track-circuit of its route. It then exits at 7:10:19, with a travel time of 914 seconds. Figure 10b concerns the fixed-speed assessment, where the train enters the first track-circuit of its route at 7:04:03, when it can reserve also the following four track-circuits respecting the blocking time model. As it exits at 7:11:21, the travel time in this case is 438 seconds. Differences such as the one shown for this train may concern several trains in a scenario, and this may have a strong impact on the coherence of the observed quality of a solution in the two assessments. As the other objective functions are much less sensitive to the moment when trains actually enter the control area, they are much less affected than the total travel time.

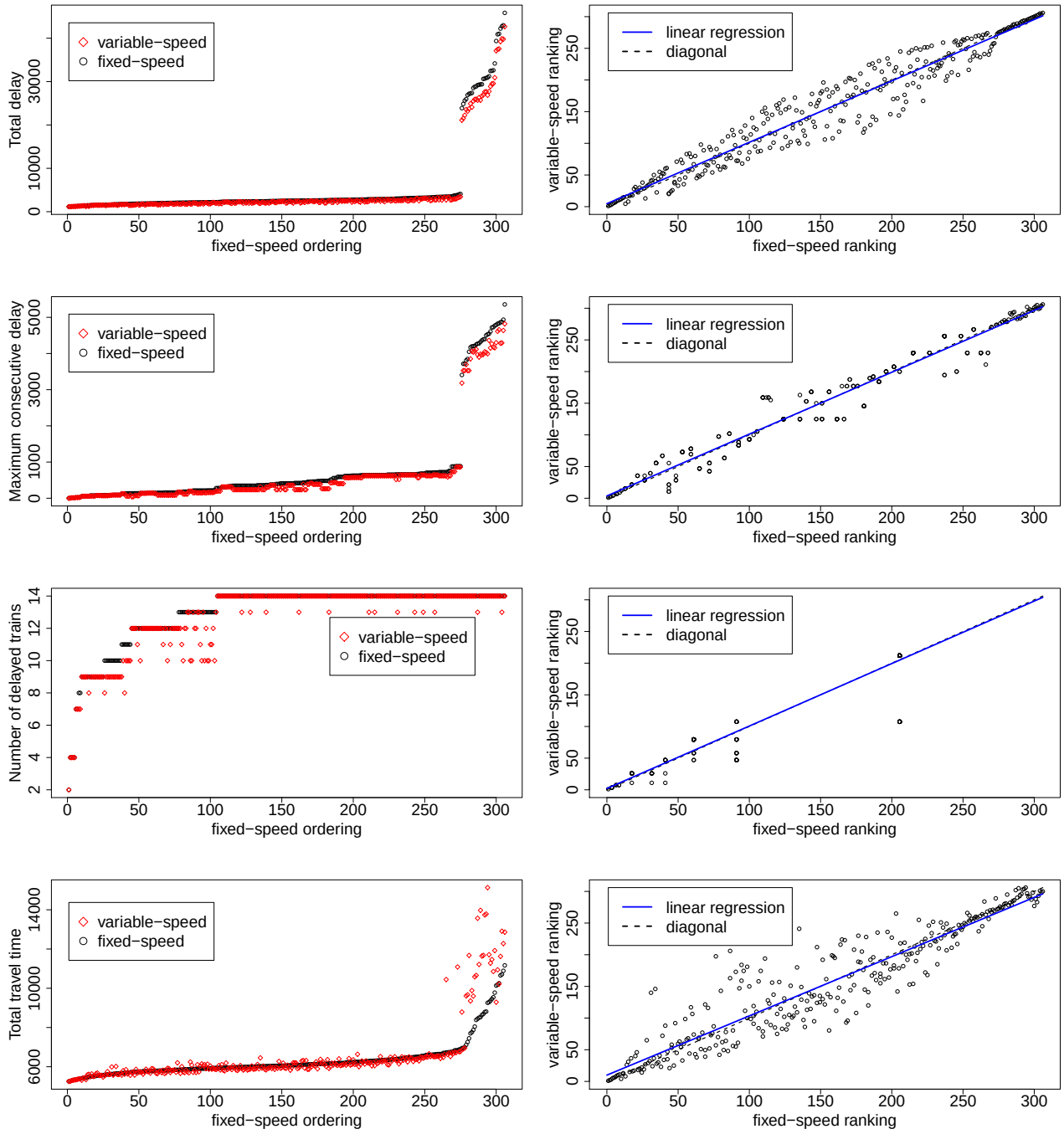
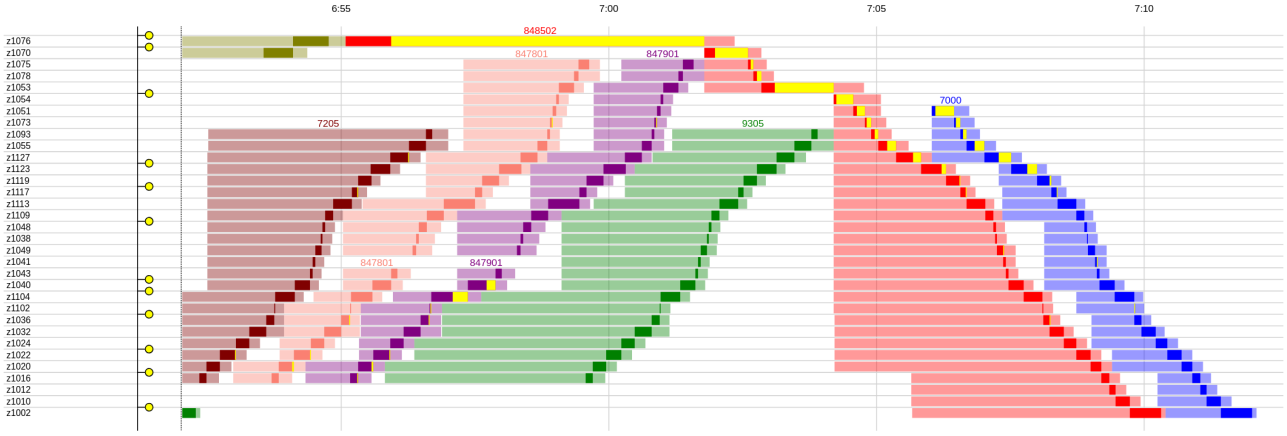
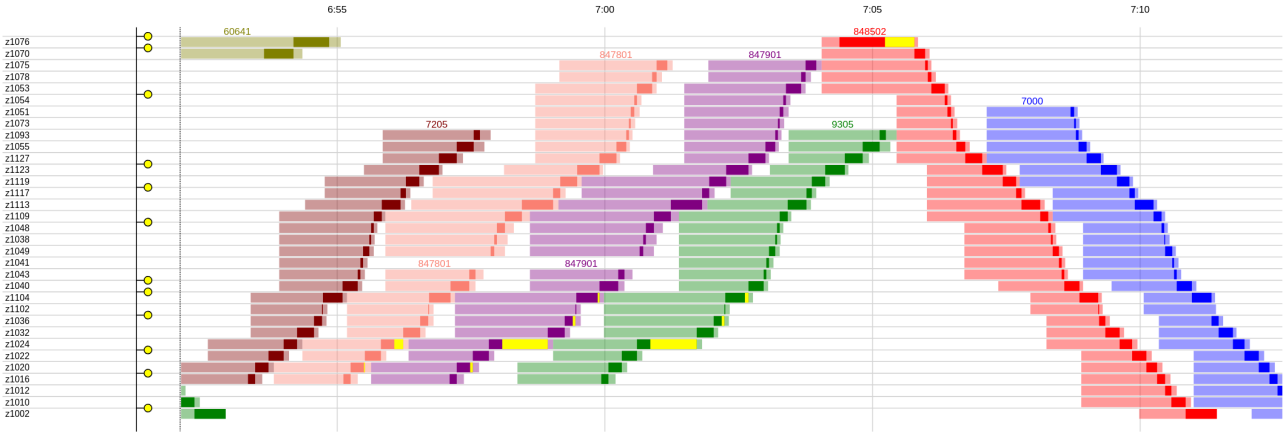


Figure 9: Results for a perturbation scenario on the Pierrefitte Gonesse junction. Value of the objective function (left panels) of fixed-speed (black circle) and simulation (red diamond) assessments as a function of the ordering in the fixed-speed assessment. Ranking of solutions (right panels) according to the simulation assessment as a function of the ranking in the fixed-speed assessment. The different objective functions from top to bottom are: total delay, maximum consecutive delay, number of delayed trains and total travel time. But for the number of delayed trains, objective function values are expressed in seconds.

In Figure 9 (right panels), we then plot the ranking according to the simulation assessment of the same solutions. Abscissas represent fixed-speed assessment rankings. As advocated in Section 3, we perform a linear



(a) Simulation assessment



(b) Fixed-speed assessment

Figure 10: Gantt space-time diagrams for train 848502 in a solution for the Pierrefitte Gonesse junction, from the simulation and the fixed-speed assessments. The yellow part represents the delay incurred by the train in case of a conflict.

	fixed-speed			min-fixed-speed		
	whole sample	50% best	25% best	whole sample	50% best	25% best
totDel	0.98	0.93	0.91	0.98	0.93	0.91
maxConsDel	0.98	0.94	0.88	0.98	0.94	0.88
numTr	0.91	0.89	**0.76**	0.91	0.89	**0.74**
totTravTime	**0.73**	**0.48**	**0.41**	**0.72**	**0.47**	**0.38**

Table 2: Average correlation coefficients with respect to the simulation assessment on the Pierrefitte Gonesse junction, over the whole sample, the best 50% and the best 25% solutions, for the four objective functions. The weaker correlations are displayed in bold font. Results are displayed both for the fixed-speed (left) and min-fixed-speed (right) assessments.

regression on the obtained cloud of points. We display in Figure 9 the straight line obtained, together with the diagonal. As can be seen, even though the linear regression and the diagonal are very close, there is a substantial dispersion of the solutions around it. This dispersion is different for the four objective functions. However, one common characteristic of the results is that, in general, very good and very bad solutions are quite correctly ranked by the fixed-speed assessment, while intermediate quality solutions appear to be more difficult to assess properly.

For readability, we only provide these plots for a representative perturbation scenario for the Pierrefitte Gonesse control area. All the similar figures for the single perturbation scenarios for all control areas are available in the supplementary material.

In order to quantify the dispersion, we compute the average correlation coefficient over the 12 perturbation scenarios and display it in Table 2. We report one line for each objective function and one set of columns for each of the fixed-speed and min-fixed-speed assessments, the reference being the ranking of solutions in the simulation assessment. Indeed, it is reasonable to suppose that an optimization algorithm will tend to provide good quality solutions. Hence, in a sense, one may consider the correlation of the set of best solutions as more important than the one of just any solution. To show what happens for these sets, in addition to the values concerning the whole sample (first column for each assessment) we also indicate the correlation coefficient computed on a fraction of the best solutions, i.e. on the 50% and 25% best solutions (in the second and third columns).

We can infer from this table that, on this particular control area, the ability of the fixed-speed assessment to capture the actual quality of the solution is the lowest when the total travel time is the objective function considered: the average correlation coefficient is at most 0.73, which means that the correlation between the ranking of solutions according to the fixed-speed and simulation assessments can be interpreted as, at most, moderate to strong. This is in line with the behavior discussed above. The other three objective functions provide very strong correlation coefficients on the whole sample, but the correlation decreases when we restrict the analysis to a fraction of the best solutions, getting to a moderate to strong level for the number of delayed trains (but remaining strong for total delay and maximum consecutive delay). When controlling the results in detail, we see that this is due to the relatively small number of solutions for some scenarios, preventing the regression from providing a reliable correlation when the rankings are not perfectly aligned. When restricting to the best solutions, the number of delayed trains seems to be more sensitive than total delay and maximum consecutive delay to this lack of stabilization. The results in the min-fixed-speed assessment are very similar to those of the fixed-speed one. This seems to imply that differences between the fixed-speed and simulation assessments do not come from neglecting decelerations and accelerations (akin to a short stop in the fixed-speed one), but from more complicated dynamics involving sometimes wrong conflict detections. In general, the correlation coefficients show that both the fixed-speed and the min-fixed-speed approximations allow the proper distinction of very bad and very good solutions, with all objective functions but the total travel time. However, the discrimination ability decreases when only good solutions are concerned. This says that an optimization algorithm based on a (min-)fixed-speed model may possibly not return the very best solution, evaluating it slightly worse than other ones. However, these other solutions will still be among the best ones. Indeed, by looking at Figure 9, we can conjecture that, if considering only the small set of very best solutions, the correlation coefficients might be higher. However, we do not present the corresponding values because the sample size is too small to be significant, and because it would be relevant only for algorithms spotting these very best solutions, while we aim at more general observations.

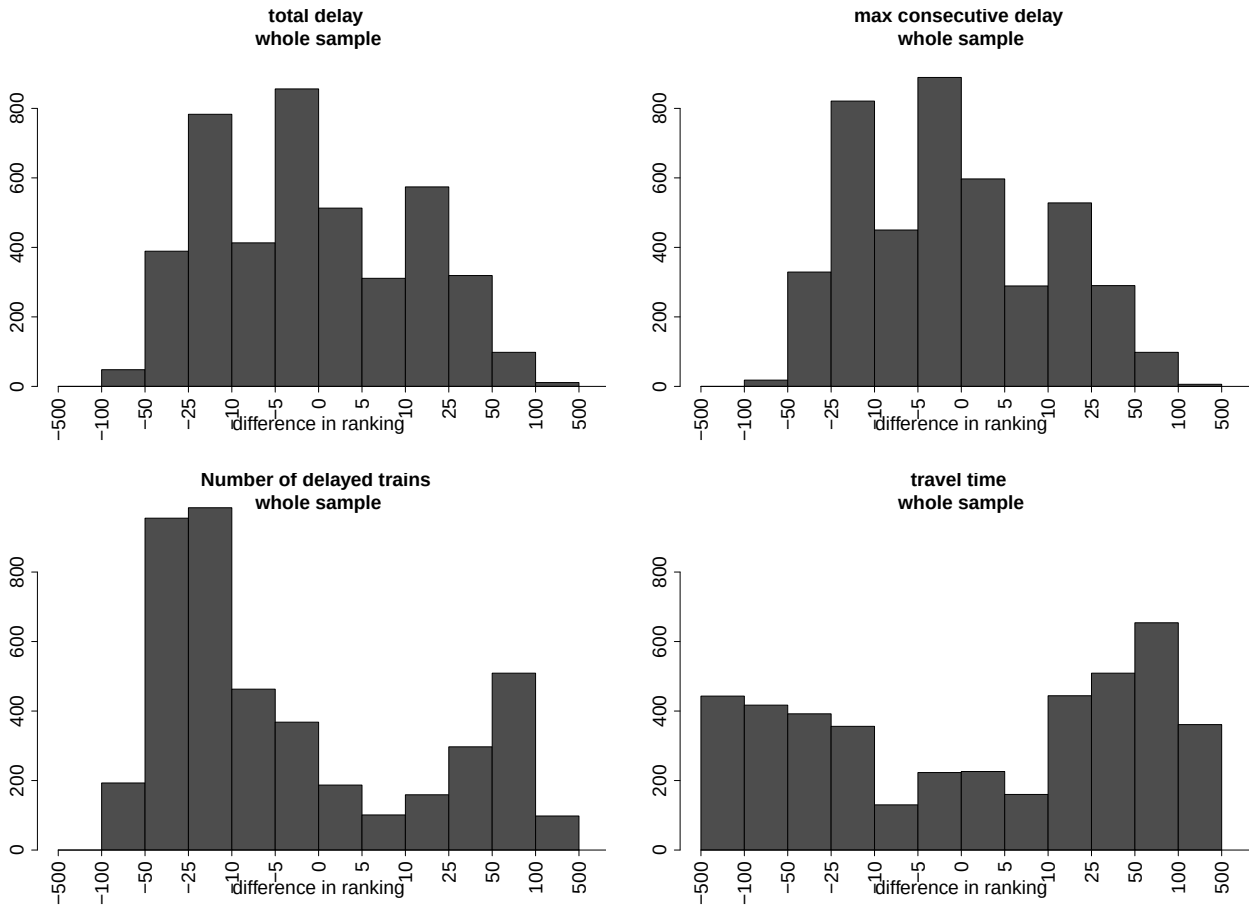


Figure 11: Histograms of the difference between the fixed-speed and simulation rankings for the four objective functions, on the Pierrefitte Gonesse junction (total over the twelve perturbation scenarios). Each bar represents the total number of solutions with a difference in ranking located in the range displayed on the x axis.

To provide further insights on the distribution of the results, we show in Figure 11 histograms indicating the difference for each solution between the ranking in the fixed-speed and the simulation assessments. In these histograms, each bar shows the number of solutions for which the difference in the two rankings is between the values indicated on the left and right side of its basis. We report one histogram per objective function. They are based on the solutions for all twelve perturbation scenarios. Ideally, one would expect a bell-shaped, narrow distribution centered on value zero, with short and thin tails. This would mean that the difference in the ranking of a solution in the assessments is often close to zero, which would testify of the good quality of the approximation. What we can see in the figure is that the distributions for the total delay and maximum consecutive delay are roughly bell-shaped and centered around zero, though their tails are fatter than one would hope for. The distribution for the total number of delayed trains has a thinner width. This may explain the good correlation coefficients for this particular objective function. Nonetheless, the distribution is not centered around zero. Finally, the shape of the distribution for the total travel time sheds some light on the weak correlation coefficients displayed in Table 2, since it has a shape almost opposite to what one could hope for. Specifically, the number of solutions with very different ranking in fixed-speed and simulation assessments is much higher than the one of those for which the ranking is similar. Over the about 4000 solutions assessed in the twelve perturbation scenarios, more than 500 appear to be very good in fixed-speed while they are among the worst in simulation (difference in ranking between -100 and -500), and the opposite relation holds about as many times (difference in ranking between 100 and 500). The ranking is the same or almost the same only for about 700 solutions (difference between -10 and 10), so in average less than 60 solutions per perturbation scenario.

	Diff. in tot. trav. time (sim. - fixed-speed)	Dispersion	Nb of brakings (sim.)	Nb of stops (sim.)	Nb of stops (fixed-speed)
inst. 1	49.66	54.78	0.65	0.26	0.67
inst. 2	70.17	59.35	1.38	0.41	1.25
inst. 3	95.21	88.76	1.37	0.42	1.26
inst. 4	16.63	26.31	0.43	0.12	0.45
inst. 5	78.49	88.84	1.29	0.41	1.26
inst. 6	109.93	95.87	1.54	0.87	1.40
inst. 7	30.82	37.28	0.73	0.24	0.86
inst. 8	74.87	74.89	1.32	0.45	1.31
inst. 9	100.16	73.68	1.11	0.40	1.04
inst. 10	38.36	46.88	0.52	0.35	0.50
inst. 11	77.77	60.83	1.31	0.41	1.21
inst. 12	113.22	75.66	1.38	0.47	1.23
average	71.27	-	1.09	0.35	1.04

Table 3: Average results over the trains and solutions, for each scenario and for all scenarios, for the Pierrefitte Gonesse junction, of the following indicators: difference in travel time (in seconds) between the simulation and the fixed-speed assessments and for each scenario the dispersion of the same quantity with respect to the different trains; the number of brakings and the number of stops for the simulation assessment; the number of stops/brakings for the fixed-speed assessment.

In accordance with the methodology described in Section 3, we perform the same statistical analysis after rounding the objective function values with respect to a precision threshold θ. The effect of this rounding is to aggregate some solutions together, so we do not differentiate them in the ranking procedure when the difference is negligible in practical terms. The results show that the correlation coefficients do not change significantly for the four objectives, with respect to what is observed in Table 2. Therefore, we do not display them here. This means that the ranking procedure does not introduce artificial noise in the statistical analysis.

Finally, we provide some additional indicators, averaged over the solutions and trains of the twelve scenarios, in order to better grasp the differences encountered in the fixed-speed and simulation assessments. In Table 3, we display (for each perturbation scenario and on average): i) the difference in total travel time of the trains (a positive number denotes a longer travel time in simulation than in the fixed-speed assessment); ii) its dispersion; iii) the number of brakings and iv) the number of stops, in the simulation assessment; and v) the number of stops in the fixed-speed assessment, which we recall is the same as the number of brakings since, according to the fixed-speed approximation, any train which brakes stops immediately. For each of the twelve perturbation scenarios, we also provide the dispersion, computed over all trains, of the average difference in travel times. This quantity reveals the cases where there is a larger dispersion of values between the different trains, even though the average travel time difference computed over all trains might seem close to zero. As a reference, the average total travel time in simulation for the Pierrefitte Gonesse junction is 543.72s while it is equal to 472.35s for the fixed-speed assessment. Such a large difference is due to the imprecision due to the blocking time model, discussed earlier in this section and illustrated in Figure 10. The details of Table 3 show, without surprise, that the number of stops in the fixed-speed assessment is larger than in the simulation one. However, we can see clearly that the number of brakings in both situations is very close (1.09 against 1.04 on average), which is a positive hint that the fixed-speed approximation manages to capture the interaction between the trains more than it appeared with some of the correlation coefficients displayed earlier.

4.2.2 St. Lazare station

We now replicate the analysis on the St. Lazare station, where trains circulate at lower speed and typically encounter more conflicts with trains running in the opposite direction, compared to the Pierrefitte Gonesse junction. The results, as displayed in Table 4, are in line with the ones observed for the first control area. The values of the average correlation coefficients are similar to the ones of Table 2, slightly higher or slightly lower but indicating the same general behavior. The only remarkable difference concerns the total travel time objective

	fixed-speed			min-fixed-speed		
	whole sample	50% best	25% best	whole sample	50% best	25% best
totDel	1.00	0.99	0.97	1.00	0.99	0.98
maxConsDel	0.99	0.97	0.94	0.99	0.97	0.94
numTr	0.92	0.80	**0.71**	0.92	0.80	**0.70**
totTravTime	0.98	0.92	0.88	0.98	0.92	0.88

Table 4: Average correlation coefficient with respect to the simulation assessment on the St. Lazare station, over the whole sample, the best 50% and the best 25% solutions, for the four objective functions. The weaker correlations are displayed in bold font. Results are displayed both for the fixed-speed (left) and min-fixed-speed (right) assessments.

function, which here is in line with the other ones. Indeed, no junctions are present at the border of the control area, which led to the frequent imprecision of the fixed-speed assessment in the Pierrefitte Gonesse junction. The total delay now becomes the objective function with the highest correlation coefficient for the whole sample, with a maximum value of 1.00. It remains such also for the 50% and 25% best solutions. The lowest average correlation is observable for the number of delayed trains when only the best solutions are considered, and is equal to 0.71, which still corresponds to a strong correlation. Even though, for this same objective function, the values of the correlation coefficients show a strong correlation between fixed-speed and simulation assessments, the values decrease faster than for the Pierrefitte Gonesse junction when we restrict the number of solutions considered. As for the Pierrefitte Gonesse junction, the min-fixed-speed assessment does not improve over the fixed-speed one.

A possible interpretation of the better correlation for the objective functions measured in time units, compared to the Pierrefitte Gonesse junction, is linked to the speed of trains. In the Pierrefitte Gonesse junction, trains typically run at line speed , that is, the maximum speed allowed on this portion of the infrastructure, and no stops are scheduled. Instead, in the St. Lazare station, trains are halting or departing and are hence running quite slowly in general. Moreover, only passenger trains circulate here, with quite effective braking and acceleration curves. On the contrary, freight trains are present on the Pierrefitte Gonesse junctions, and the time necessary for a freight train to stop and re-accelerate to its free-network speed is not at all negligible. This implies that, in the St. Lazare station, braking and acceleration due to traffic generally do not imply a strong difference with respect to the free-network speed profiles. In the Pierrefitte Gonesse junction, instead, the difference between free-network and traffic-perturbed speed profiles is intuitively much stronger. Hence, the error made when neglecting this difference as in the fixed-speed assessment is higher in the Pierrefitte Gonesse junction, in terms of both the impact of conflicts on train runs, and the number and characteristics of conflicts detected.

Here again, we provide the histograms of the distribution of the difference in fixed-speed and simulation rankings in Figure 12. The histograms show that the distributions for the total delay and maximum consecutive delay are still roughly bell-shaped, with a higher concentration of solutions around zero for the latter. For this control area, both the number of delayed trains and the total travel time look like double bell-shaped symmetric distributions, which explains the slight degradation of the correlation coefficient for the former objective function. As for the previous case of the Pierrefitte Gonesse junction, the correlation coefficients for aggregated solutions do not change with respect to those displayed in Table 4 and are hence not displayed here. However, they are available in the supplementary material.

We complete the analysis of the second control area through the analysis of Table 5, similar to Table 3 for the Pierrefitte Gonesse junction, with average differences in total travel time and information of the brakings and stops of the trains in simulation and fixed-speed assessments. The average total travel time in simulation for the St. Lazare station is 429.89s while it is equal to 441.73s for the fixed-speed assessment, which demonstrates the absence of the imprecision discussed for the Pierrefitte Gonesse junction, as illustrated in Figure 10. The relative difference in the number of brakings in the two assessments is once again quite small (few percentage points). Our analysis of specific solutions of both the simulation and the fixed-speed assessments tends to show that the larger travel times of trains in the fixed-speed one are due to the blocking time model, and to the fact that they cannot enter some block sections with a restrictive signal.

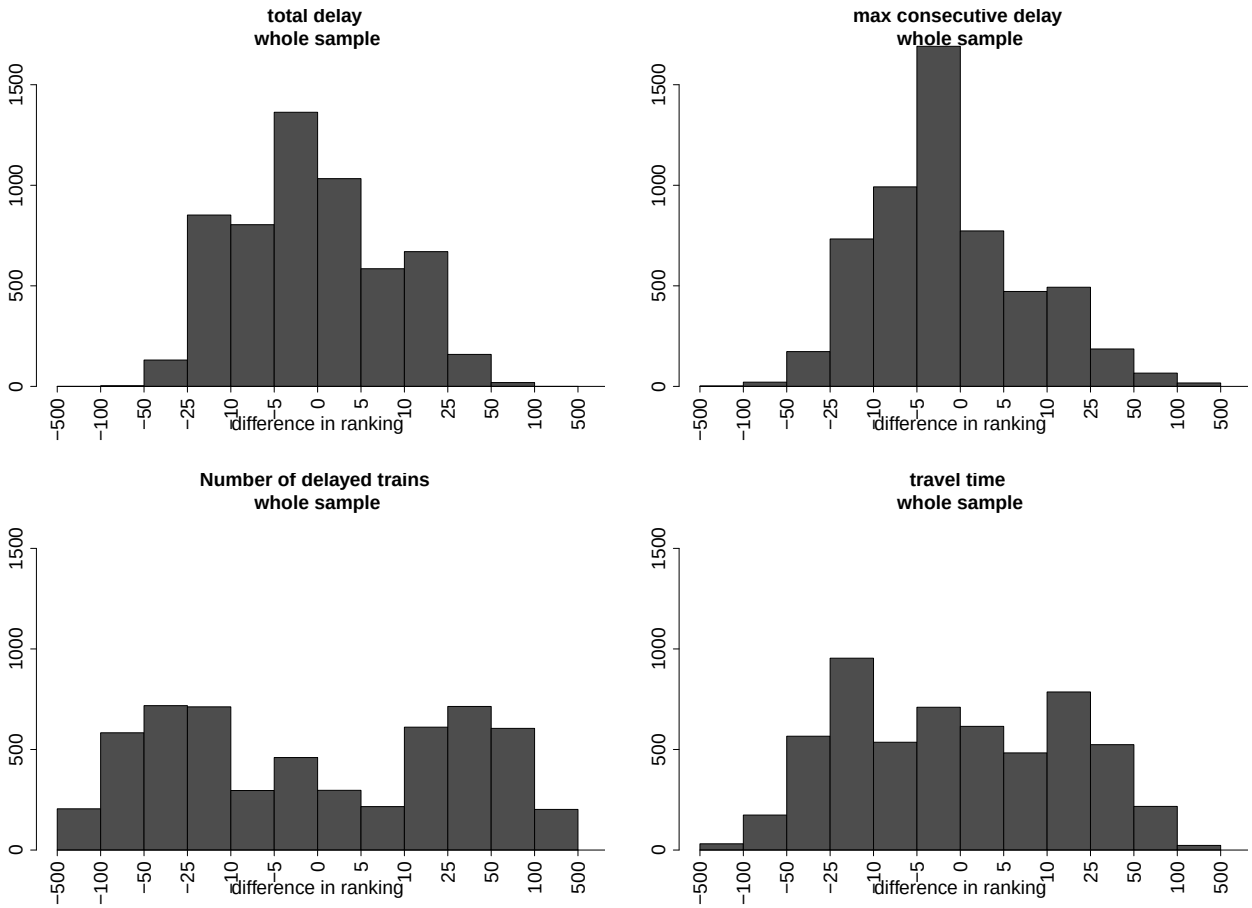


Figure 12: Histograms of the difference between the fixed-speed and simulation rankings for the four objective functions, on the St. Lazare station (total over the twelve perturbation scenarios). Each bar represents the total number of solutions with a difference in ranking located in the range displayed on the x axis.

4.2.3 Mantes-la-Jolie–Rouen line

We finally display the results for the last control area, which is the longest one and includes several stations. The average ranking correlation coefficients are displayed in Table 6 and the histogram distributions are shown in Figure 13. For the case of the Mantes-la-Jolie–Rouen line, we can see that the fixed-speed approximation is capable of capturing the actual relative quality of solutions for all four objective functions. The correlation coefficients are always higher than 0.80, and for the total delay they reach the impressive values of 0.98, 0.93 and 0.92 for the whole sample, the 50% and 25% best solutions, respectively. This capability can also be assessed from the histograms of the distributions of the differences in ranking. Contrary to the previous control areas, the four distributions are closer to bell shaped ones, with thinner and shorter tails. Once again, the min-fixed approximation does not help improve the correlation coefficients and neither does the aggregation of solutions (which we do not display here, but are available in the supplementary material).

A peculiarity of this control area is the presence of buffer times associated to train stops in the timetable. Typically, some seconds of time allowance are included in the schedule on this line, and this allows absorbing inaccuracies in running time computations in the fixed and min-fixed-speed assessments. The presence of stops along the train trips, moreover, often forces the train speed profiles computed with the approximations to catch-up with the realistic speed profiles and with the timetable: no earlier departures from stations are possible. Hence, if the fixed or the min-fixed-speed assessments underestimate the utilization of buffers before a scheduled stop, this will have no impact on the departure time from the stop itself unless these buffers are insufficient. Moreover, this control area being a line portion, rather than including a large station or junction, trains mostly

	Diff. in tot. trav. time (sim. - fixed)	Dispersion	Nb of brakings (sim.)	Nb of stops (sim.)	Nb of stops (fixed-speed)
inst. 1	-8.14	20.52	0.89	0.29	0.70
inst. 2	-14.08	24.65	1.13	0.40	1.05
inst. 3	-13.86	23.86	1.24	0.41	1.13
inst. 4	-12.58	17.38	0.72	0.28	0.81
inst. 5	-12.76	26.66	1.15	0.44	1.13
inst. 6	-13.40	22.66	1.01	0.33	0.94
inst. 7	-9.66	17.62	0.72	0.27	0.70
inst. 8	-12.04	21.57	1.21	0.40	1.17
inst. 9	-12.53	24.13	1.31	0.36	1.11
inst. 10	-11.66	16.74	0.68	0.30	0.74
inst. 11	-10.13	23.68	1.09	0.38	1.06
inst. 12	-11.22	22.08	1.42	0.41	1.21
average	-11.84	-	1.05	0.36	0.98

Table 5: Average results over the trains and solutions, for each scenario and for all scenarios, for the St. Lazare station, of the following indicators: difference in travel time (in seconds) between the simulation and the fixed-speed assessments and for each scenario the dispersion of the same quantity with respect to the different trains; the number of brakings and the number of stops for the simulation assessment; the number of stops/brakings for the fixed-speed assessment.

	fixed-speed			min-fixed-speed		
	whole sample	50% best	25% best	whole sample	50% best	25% best
totDel	0.98	0.93	0.92	0.98	0.93	0.92
maxConsDel	0.93	0.89	0.80	0.93	0.89	0.80
numTr	0.93	0.85	0.80	0.93	0.85	0.80
totTravTime	0.97	0.91	0.87	0.97	0.91	0.87

Table 6: Average correlation coefficients with respect to the simulation assessment on the Mantes-la-Jolie–Rouen line, over the whole sample, the best 50% and the best 25% solutions, for the four objective functions. The results are displayed both for the fixed-speed (left) and min-fixed-speed (right) assessments.

follow each other. Few crossings between trains running in opposite directions occur, and it is not very frequent to have train routes merge and then depart from each other. Hence, the impact of a conflict is somehow easier to evaluate, even with a likely error on the exact number of seconds that undeniably characterizes the fixed and min-fixed-speed assessments.

Finally, we provide in Table 7 information similar to Tables 3 and 5, i.e. the average differences in total travel time and the number of brakings and stops of the trains in the simulation and fixed-speed assessments. The average total travel time in simulation for the Mantes-la-Jolie–Rouen line is 2023.13s, while it is equal to 1947.52s for the fixed-speed assessment, which is around four times larger than for the two previous control areas. The relative difference in the number of brakings is similar to the one encountered for previous control areas. Since the dispersions in the total travel time difference are much larger than the average values, it means we have trains with both (substantial) negative or positive average difference of total travel time, a piece of information not apparent in the global average difference of total travel time. An investigation of specific solutions shows that this control area sometimes suffers from the blocking time model imprecision encountered for the Pierrefitte Gonesse junction (see discussion on Figure 10), i.e. some trains are forced to wait before entering the control area to give priority to other trains, while they can enter with a restrictive signal in simulation. However, this problem has less of an impact on the correlation coefficients as a whole, possibly because here trains usually make several stops, increasing substantially the travel times and diminishing the influence of a delayed entrance in the infrastructure.

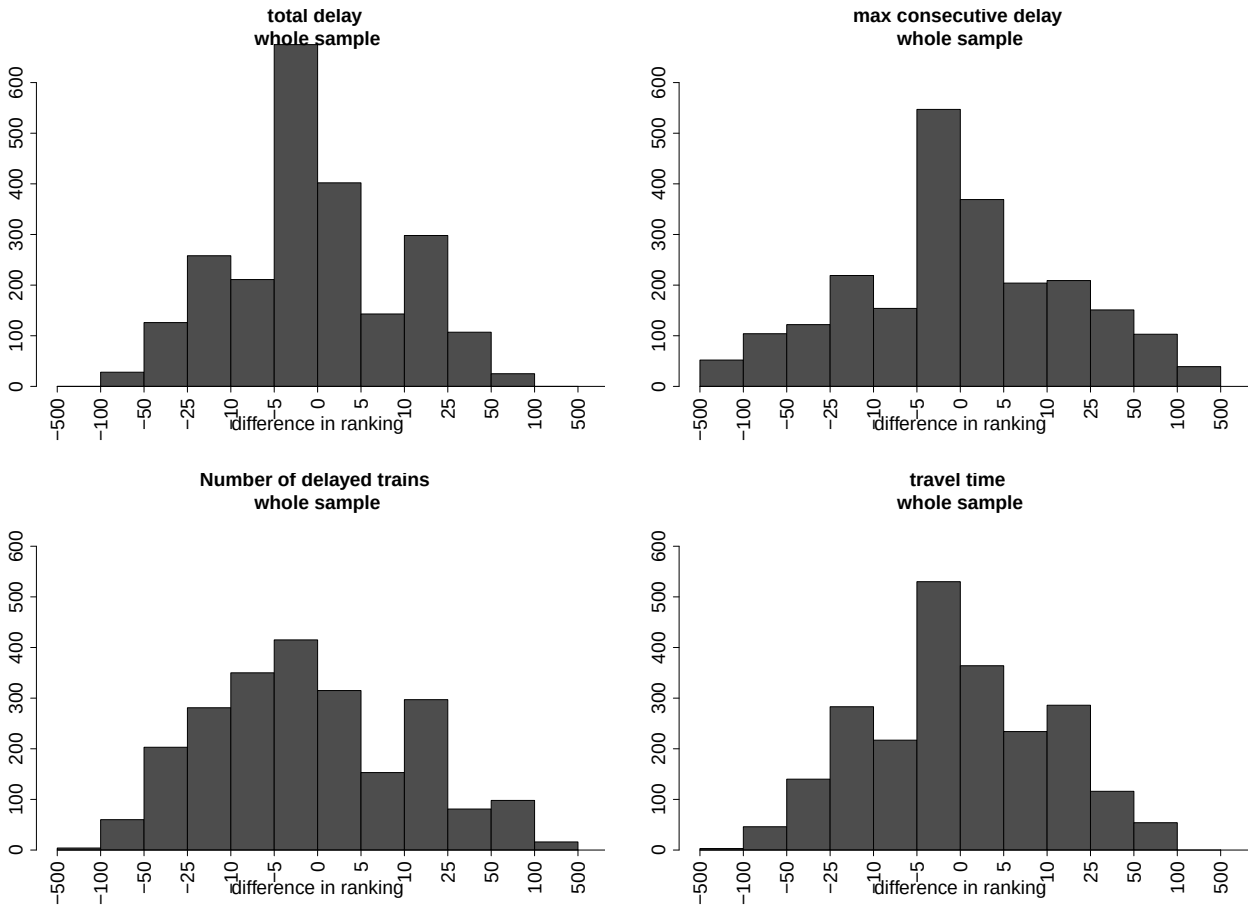


Figure 13: Histograms of the difference between the fixed-speed and simulation rankings for the four objective functions, on the Mantes-la-Jolie–Rouen line (total over the twelve perturbation scenarios). Each bar represents the total number of solutions with a difference in ranking located in the range displayed on the x axis.

4.2.4 Results with limited blocking time model

In some solutions discussed above in this section, we observed large differences on train entrance in control areas due to the blocking time model, e.g. as illustrated in Figure 10. In order to elucidate the impact of this aspect on the correlation factors, we replicate the whole analysis considering a limited version of the blocking time model: we consider a simplified signalling system with only two aspects. Trains will now only stop when seeing a restrictive signal that displays a red aspect, i.e. before entering the first track-circuit of a block section where another train is to pass first.

When considering such a version, we can see a sharp evolution of travel times for the Pierrefitte Gonesse case. The average travel time now increases to 516.59s, which represents only a difference of 27.14s with respect to the average travel time in simulation. When we recompute the correlation coefficients as in Table 2 and display the results in Table 8, we can see a solid improvement for the total travel time (as well as a marginal improvement for other objective functions). These results shed some light on the interplay of the fixed-speed approximation and the blocking time model, and the possible undesirable effects arising from their interaction.

The results for the other two control areas are displayed in Tables 9 and 10. While the correlation coefficients are similar for the St. Lazare station (and in fact tend to improve), for the Mantes-la-Jolie–Rouen line the table shows a slight deterioration for the subsets of the 50% or 25% best solutions. By checking the correlation coefficients scenario per scenario, however, we can confirm that this decrease comes from scenarios where the final number of solutions considered is around (or even under) 100, so the 25% best solutions correspond to a somewhat small sample, which tends to easily worsen the correlation coefficient whenever the ranking is not

	Diff. in tot. trav. time (sim. - fixed)	Dispersion	Nb of brakings (sim.)	Nb of stops (sim.)	Nb of stops (fixed-speed)
inst. 1	103.65	236.68	0.93	0.36	0.81
inst. 2	88.51	130.10	1.30	0.43	1.21
inst. 3	84.24	112.21	1.75	0.53	1.57
inst. 4	64.76	143.45	0.91	0.34	0.93
inst. 5	36.23	75.48	1.37	0.42	1.21
inst. 6	29.82	90.31	1.32	0.36	1.03
inst. 7	68.38	178.91	1.01	0.41	1.08
inst. 8	82.65	125.68	1.45	0.42	1.33
inst. 9	104.80	114.40	1.93	0.53	1.98
inst. 10	43.53	51.87	0.80	0.34	0.88
inst. 11	133.45	130.07	1.65	0.56	1.60
inst. 12	67.27	109.82	1.34	0.38	1.05
average	75.61	-	1.31	0.42	1.22

Table 7: Average results over the trains and solutions, for each scenario and for all scenarios, for the Mantes-la-Jolie–Rouen line, of the following indicators: difference in travel time (in seconds) between the simulation and the fixed-speed assessments and for each scenario the dispersion of the same quantity with respect to the different trains; the number of brakings and the number of stops for the simulation assessment; the number of stops/brakings for the fixed-speed assessment.

	fixed-speed			min-fixed-speed		
	whole sample	50% best	25% best	whole sample	50% best	25% best
totDel	0.99	0.97	0.94	0.99	0.97	0.94
maxConsDel	0.99	0.97	0.92	0.99	0.97	0.92
numTr	0.97	0.93	0.85	0.97	0.93	0.85
totTravTime	0.98	0.94	0.90	0.99	0.95	0.91

Table 8: Average correlation coefficients with respect to the simulation assessment on the Pierrefitte Gonesse junction, over the whole sample, the best 50% and the best 25% solutions, for the four objective functions and a two aspects signalling system. Results are displayed both for the fixed-speed (left) and min-fixed-speed (right) assessments.

	fixed-speed			min-fixed-speed		
	whole sample	50% best	25% best	whole sample	50% best	25% best
totDel	1.00	0.99	0.99	1.00	0.99	0.99
maxConsDel	0.99	0.98	0.97	0.99	0.98	0.97
numTr	0.94	0.82	**0.75**	0.94	0.82	**0.75**
totTravTime	0.99	0.96	0.93	0.99	0.96	0.93

Table 9: Average correlation coefficients with respect to the simulation assessment on the St. Lazare station, over the whole sample, the best 50% and the best 25% solutions, for the four objective functions and a two aspect signalling system. Results are displayed both for the fixed-speed (left) and min-fixed-speed (right) assessments.

perfectly maintained between simulation and fixed-speed. An indication of the good correspondence between the simulation and the fixed-speed assessments with the limited version of the blocking time model is the small difference in the average travel time, now reduced to 2.42s for St. Lazare and 50.68s for Mantes-la-Jolie–Rouen with a reduced dispersion between the different trains. Overall, we confirm that the role of the blocking time model is in general minor with respect to the quality of the fixed-speed approximation, except in the specific cases related to the train entrance times.

	fixed-speed			min-fixed-speed		
	whole sample	50% best	25% best	whole sample	50% best	25% best
totDel	0.98	0.93	0.85	0.98	0.93	0.85
maxConsDel	0.97	0.88	**0.69**	0.97	0.88	**0.69**
numTr	0.87	**0.74**	**0.63**	0.87	**0.74**	**0.63**
totTravTime	0.98	0.90	**0.78**	0.98	0.90	**0.79**

Table 10: Average correlation coefficients with respect to the simulation assessment on the Mantes-la-Jolie–Rouen line, over the whole sample, the best 50% and the best 25% solutions, for the four objective functions and a two aspect signalling system. Results are displayed both for the fixed-speed (left) and min-fixed-speed (right) assessments.

5 Conclusion

In this paper, we have proposed a study of the quality of the so-called fixed-speed approximation for train dynamics in optimization models of the real time Railway Traffic Management Problem. In the often used fixed-speed approximation, trains running times are precisely known if the network is free, as they are typically computed *a priori*. Instead, instantaneous stops and accelerations are considered in case of conflicts. Hence, when traffic is perturbed, this approximation may be quite inaccurate in computing train running times, compared to the realistic results obtained using realistic speed profiles.

We used three different railway control areas with various train behaviors, generated a dozen perturbation scenarios for each, and hundreds of solutions for said scenarios. A statistical analysis was performed over the solutions for each scenario and four different objective functions used in the literature, to assess whether the ranking of solutions was the same with both the fixed-speed and realistic speed profiles, the latter being computed through microscopic simulation. We also considered a slightly more refined approximation, in which the fixed-speed one is brought a step closer to the realistic speed dynamics. However, this approximation did not prove to behave differently from the fixed-speed one.

The average results on all perturbation scenarios indicate that the fixed-speed approximation is capable of discriminating between good and bad solutions for most objective functions, especially with solutions of very different qualities. However, when one looks specifically at the best solutions, the correlation coefficients between fixed-speed and simulation rankings tend to decrease, which is partly due to the reduced cardinality of the sets of very best solutions. Overall, the capability of the fixed-speed approximation of properly identifying good solutions appears high for all objective functions considered, being the highest with the total delay. Instead, on specific control areas, the correlation coefficients with respect to the simulation assessments are the lowest for the total travel time. By computing the travel times with a limited version of the blocking time model, which considers a two-aspect signaling system, we demonstrated that this effect is mainly due to the application of said model, rather than to the fixed-speed approximation.

The conclusion of our analysis is that the fixed-speed approximation overall correctly captures the quality of the different routing and scheduling decisions, though in some occurrences the blocking time model may have non-negligible effects that one should be careful about.

In future research, we will aim at understanding the link between the characteristics of a control area and the quality of the fixed-speed approximation with respect to different objective functions. This may allow the proposal of tailored improvements of the approximation itself. Such a tailoring may be pertinent also with respect to clusters of instances, for example in terms of the quantity or intensity of the conflicts to be dealt with. A sensitivity analysis with respect to different characteristics and set-ups of the signaling system may allow an even finer tailoring of the approximation to the application at hand. Furthermore, it will be interesting to consider stochastic variations, e.g., in the running time of the trains: in this case, it will be necessary to analyze the quality of the fixed-speed approximation by implementing a closed-loop between optimization and simulation to reoptimize the train routing and scheduling, which is in general technically challenging. With such a closed-loop implementation, it will also be possible to evaluate the impact of real-time data latency effects on the quality of the approximation. The conducted fixed-speed approximation analysis did not include benchmarking against methods employing train dynamics formulations or data-driven approaches like neural networks. Future benchmarking among these methods will offer significant insights into computational efficiency versus solution quality trade-offs. Since interpretability lacks a quantifiable metric, the principal

observable outcome will likely be a lower ranking of data-driven approaches in terms of interpretability relative to the other two.

References

Agasucci, V., Grani, G., and Lamorgese, L., 2023. "Solving the train dispatching problem via deep reinforcement learning". *Journal of Rail Transport Planning & Management*, vol. 26, 100394.

Akoglu, H. 2018. "User's guide to correlation coefficients". *Turkish Journal of Emergency Medicine*, vol. 18:3, pp. 91–93.

Albrecht, T., 2009. "The influence of anticipating train driving on the dispatching process in railway conflict situations". *Networks Spatial Economics*, vol. 9:1, pp. 85–101.

Bach, L., Mannino, C. and Sartor, G., 2019 "MILP Approaches to Practical Real-Time Train Scheduling". In: *OpenProceedings.org*, DOI: 10.5441/002/INOC.2019.15.

Cacchiani, V., Huisman, D., Kidd, M., Kroon, L., Toth, P., Veelenturf, L., and Wagenaar, J., 2014. "An overview of recovery models and algorithms for real-time railway rescheduling". *Transportation Research Part B: Methodological*, vol. 63, pp. 15–37.

Caimi, G., Chudak, F., Fuchsberger, M., Laumanns, M., and Zenklusen, R., 2011. "A new resource-constrained multicommodity flow model for conflict-free train routing and scheduling". *Transportation Science*, vol. 45:2, pp. 212–227.

Corman, F., D'Ariano, A., Pacciarelli, D., and Pranzo, M., 2010. "A tabu search algorithm for rerouting trains during rail operations". *Transportation Research Part B*, vol. 44:1, pp. 175–192.

Corman, F., D'Ariano, A., Pacciarelli, D., and Pranzo, M., 2012. "Optimal inter-area coordination of train rescheduling decisions". *Transportation Research Part E*, vol. 48, pp. 71–88.

D'Ariano, A., and Albrecht, T., 2006. "Running time re-optimization during real-time timetable perturbations". In: *WIT Transactions on the Built Environment*, vol. 88. WIT Press, pp. 532–540.

D'Ariano, A., Pacciarelli, D., and Pranzo, M., 2007. "A branch and bound algorithm for scheduling trains in a railway network". *European Journal of Operational Research*, vol. 183, pp. 643–657.

D'Ariano, A., Pranzo, M., and Hansen, I.A., 2007. "Conflict Resolution and Train Speed Coordination for Solving Real-Time Timetable Perturbations". *IEEE Transactions on Intelligent Transportation Systems*, vol. 8:2, pp. 208–222.

Dündar, S., and Şahin, I., 2013. "Train re-scheduling with genetic algorithms and artificial neural networks". *Transportation Research Part C: Emerging Technologies*, vol. 27:0, pp. 1–15.

Huang, P., Li, Z., Zhu, Y., Wen, C., and Corman, F., 2023. "Train traffic control in merging stations: A data-driven approach". *Transportation Research Part C: Emerging Technologies*, vol. 152, 104155.

Khosravi, B., Bennell, J.A., and Potts, C.N., 2012. "Train scheduling and rescheduling in the UK with a modified shifting bottleneck procedure". In: *Delling, D., Liberti L. (Eds), Proceedings of the 12th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization and Systems*, vol. 25 of OpenAccess Series in Informatics (OASIs), Dagstuhl, Germany, Schloss Dagstuhl-Leibniz-Zentrum für Informatik, pp. 120–131.

Krasemann, J.T., 2015. "Computational decision-support for railway traffic management and associated configuration challenges: An experimental study". *Journal of Rail Transport Planning & Management*, vol. 5, pp. 95–109.

Lamorgese, L. and Mannino, C., 2015. "An exact decomposition approach for the real-time train dispatching problem". *Operations Research*, vol. 63:1, pp. 48–64.

Long, S., Meng, L., Wang, Y., Miao, J., Luan, X., and Corman, F. (2023).
Integrated speed modeling and traffic management to precisely model the effect and dynamics of temporary speed restrictions to high-speed railway traffic". *Transportation Research Part C: Emerging Technologies*, vol. 152:104148.

Luan, X., Wang, Y., De Schutter, B., Meng, L., Lodewijks, G. and Corman, F., 2018. "Integration of real-time traffic management and train control for rail networks - Part 1: Optimization problems and solution approaches". *Transportation Research Part B*, vol. 115, pp. 41–71.

Luan, X., Wang, Y., De Schutter, B., Meng, L., Lodewijks, G. and Corman, F., 2018. "Integration of real-time traffic management and train control for rail networks - Part 2: Extensions towards energy-efficient train operations". *Transportation Research Part B*, vol. 115, pp. 72–794.

Lüthi, M., 2009. "Improving the efficiency of heavily used railway networks through integrated real-time rescheduling". *Diss., Egenössische Technische Hochschule Zürich.*

Mazzarello, M., and Ottaviani, E., 2007. "A traffic management system for real-time traffic optimization in railways". *Transportation Research Part B*, vol. 41:2, pp. 246–274.

Meng, L., and Zhou, X., 2014. "Simultaneous train rerouting and rescheduling on an n-track network: A model reformulation with network-based cumulative flow variables". *Transportation Research Part B: Methodological*, vol. 67, pp. 208–234.

Nash, A. and Huerlimann, D. (2004). " Railroad simulation using OpenTrack". In Brebbia, C., Allan, J., Sciutto, G., and Scone, S., editors, *Computers in Railways IX*, pages 45–54. WIT Press, Southampton, United Kingdom.

Nettleton, D., 2014. "Selection of Variables and Factor Derivation", In Nettleton, D., editor, *Commercial Data Mining*, chapter 6, pages 79–104, Morgan Kaufmann, Boston, USA.

Pachl, J., 2008. "Timetable design principles". In Hansen, I. and Pachl, J., editors, *Railway Timetable & Traffic*, chapter 2, pages 9–42. Eurailpress — DVV Rail Media, Hambourg, Germany.

Pascariu, B., Samà, M., Pellegrini, P., D'Ariano, A., Rodriguez, J., and Pacciarelli, D., 2024. "Formulation of train routing selection problem for different real-time traffic management objectives". *Journal of Rail Transport Planning & Management*, vol. 31, 100460.

Pellegrini, P., Marlière, G., and Rodriguez, J., 2014. "Optimal train routing and scheduling for managing traffic perturbations in complex junctions". *Transportation Research Part B: Methodological*, vol. 59, pp. 58–80.

Pellegrini, P., Marlière, G., Pesenti, R., and Rodriguez, J., 2015. "RECIFE-MILP: An Effective MILP-Based Heuristic for the Real-Time Railway Traffic Management Problem". *IEEE Transactions on Intelligent Transportation Systems*, vol. 16:5, pp. 2609–2619.

Reynolds, E., and Maher, S.J., 2022. "A data-driven, variable-speed model for the train timetable rescheduling problem", *Computers & Operations Research*, vol. 142, 105719,

Rodriguez, J., 2007. "A constraint programming model for real-time train scheduling at junctions". *Transportation Research Part B: Methodological*, vol. 41, pp. 231–245.

Samà, M., Meloni, C., D'Ariano, A., Corman, F., 2015. "A multi-criteria decision support methodology for real-time train scheduling". *Journal of Rail Transport Planning & Management*, vol. 5, no. 3, pp. 146–162.

Samà, M., D'Ariano, A., Corman, F., and Pacciarelli, D., 2017. "A variable neighbourhood search for fast train scheduling and routing during disturbed railway traffic situations". *Computers & Operations Research*, vol. 78, pp. 480–499.

Sobieraj, S., Marlière, G., and Rodriguez, J., 2011. "Simulation of solutions of a fixed-speed model for the real-time railway traffic optimization problem". In: *Proceedings of The 4th International Seminar on Railway Operations Modelling and Analysis (RailRome2011)*, Rome, Italy.

Törnquist, J., and Persson, J., 2007. "N-tracked railway traffic re-scheduling during disturbances". *Transportation Research Part B*, vol. 41, pp. 342–362.

Versluis, N.D., Quaglietta, E., Goverde, R.M.P., Pellegrini, P., and Rodriguez, J., 2024. "Real-time railway traffic management under moving-block signalling: A literature review and research agenda", *Transportation Research Part C: Emerging Technologies*, vol. 158:104438

Xu, P., Corman, F., Peng, Q. and Luan, X., 2017. "A train rescheduling model integrating speed management during disruptions of high-speed traffic under a quasi-moving block system". *Transportation Research Part B*, vol. 104, pp. 636–666.

Xu, W., Zhao, C., Cheng, J., Wang, Y., Tang, Y., Zhang, T., Yuan, Z., Lv, Y., and Wang, F.-Y., 2023. "Transformer-based macroscopic regulation for high-speed railway timetable rescheduling". *IEEE/CAA Journal of Automatica Sinica*, vol. 10, no. 9 pp. 1822-1833.

Zhou, L., Tong, L., Chen, J., Tang, J. and Zhou, X., 2017. "Joint optimization of high-speed train timetables and speed profiles: A unified modeling approach using space-time-speed grid networks". *Transportation Research Part B*, vol. 97, pp. 157–181.

A MILP formulation for the rtRTMP

In this appendix we detail the MILP formulation at the basis of the RECIFE-MILP algorithm, used to asses the impact of the Fixed-Speed approximation in the rest of the paper. RECIFE-MILP uses the following sets:

- T: the set of trains;
- Θ: the set of train types;
- R_t: the set of routes available to train $t \in T$, with $R = \cup_{t \in T} R_t$ the total set of routes;
- TC_t: the set of track-circuits which can be used by train $t \in T$;
- TC^r: the set of track-circuits belonging to route $r \in R$;
- $OTC_{ty,r,tc}$: the set of track-circuits such that, if a train $t \in T$ of type $ty \in \Theta$ traverses them along $r \in R_t$ and has its head at their end, it holds that t's tail has not yet left tc. $OTC_{ty,r,tc} = \{tc\}$ if t is shorter than tc itself;
- S_t, $TCS_{t,s}$: the set of stations where $t \in T$ has a scheduled stop and set of track-circuits that can be used by t for stopping at $s \in S_t$;

and parameters:

- tc_0 and tc_∞: dummy track-circuits representing entry and the exit locations of the control area considered;
- $sched_t$: scheduled arrival time of traint $t \in T$ at destination;
- ty_t: type corresponding to train t (train characteristics);
- $init_t$, $exit_t$: earliest time at which train $t \in T$ can be operated and earliest time at which it can reach its destination given $init_t$, the route assigned in the timetable and the intermediate stops;
- $i(t', t)$: indicator function equal to 1 if t' and t use the same rolling stock and t results from the turnaround, join or split of t', 0 otherwise;
- $ms_{t,t'}$: minimum separation between the arrival and the departure of trains t and t' using the same rolling stock;
- $rt_{r,ty,tc}$, $ct_{r,ty,tc}$: free-network running and clearing time of $tc \in TC^r$ along $r \in R$ for a train of type $ty \in \Theta$;
- $ref_{r,tc}$: reference track-circuit for the reservation of $tc \in TC^r$ along $r \in R$, depending on block section structure and interlocking system;
- $e(tc, r)$: indicator function equal to 1 if track-circuit $tc \in TC^r$ belongs to either the first or the last block section of $r \in R$, 0 otherwise;
- $bs_{r,tc}$: block section including track-circuit $tc \in TC^r$ along route $r \in R$;
- for_{bs}, rel_{bs}: formation and release time for block section bs;
- $dw_{t,s}, a_{t,s}, d_{t,s}$: minimum dwell time, scheduled arrival and scheduled departure times for train $t \in T$ at station $s \in S_t$;
- $p_{r,tc}$, $s_{r,tc}$: track-circuits preceding and following $tc \in TC^r$ along $r \in R$;
- M: a large constant.

Remark that all trains of the same type are considered to be planned to travel in the same way across a track-circuit along a route. If train-specific running and clearing times need to be considered, $rt_{r,ty,tc}$ and $ct_{r,ty,tc}$ can be indexed on the train itself rather than on its type.

We also make use of the following variables, which include the binary routing and scheduling decisions as well as continuous variables used to evaluate the travel time and potential delays of the trains on each track-circuit:

- $sU_{t,tc}$, $eU_{t,tc}$: continuous positive variables representing the time at which $tc \in TC_t$ starts and ends being utilized by $t \in T$;

- $x_{t,r}$: binary variable equal to 1 if train $t \in T$ uses route $r \in R_t$, 0 otherwise;
- $y_{t,t',tc}$: binary variable equal to 1 if train $t \in T$ utilizes track-circuit tc before train t', such that index t is smaller than index t' ($t \prec t'$), with $tc \in TC_t \cap TC_{t'}$, and 0 otherwise;
- $o_{t,r,tc}$: time at which $t \in T$ starts the occupation of $tc \in TC^r$ along $r \in R_t$;
- $l_{t,r,tc}$: longer stay of $t \in T$'s head on $tc \in TC^r$ along $r \in R_t$, due to dwell time and scheduling decisions (delay).

To model the different objective functions we are interested in, we define the following variables:

- D_t: delay suffered by train t when exiting the control area;
- $\bar{D}_t$: consecutive delay suffered by train t when exiting the control area;
- $Dentr_t$: delay imposed to t at its entrance in the control area;
- Δ: maximum consecutive delay among all trains;
- δ_t: binary variable equal to 1 if train $t \in T$ suffers some delay compared to its original timetable.

The model has to respect the following sets of constraints:

$$o_{t,r,tc} \geq init_t\, x_{t,r} \qquad \forall t \in T, r \in R_t, tc \in TC^r, \tag{3}$$

$$o_{t,r,tc} \leq M x_{t,r} \qquad \forall t \in T, r \in R_t, tc \in TC^r, \tag{4}$$

$$o_{t,r,tc} = o_{t,r,p_{r,tc}} + l_{t,r,p_{r,tc}} + rt_{r,ty_t,p_{r,tc}} x_{t,r} \qquad \forall t \in T, r \in R_t, tc \in TC^r, \tag{5}$$

$$o_{t,r,s_{r,tc}} \geq \sum_{s \in S_t : tc \in TCS_{t,s} \cap TC^r} d_{t,s} x_{t,r} \qquad \forall t \in T, r \in R_t, tc \in \bigcup_{s \in S_t} TCS_{t,s}, \tag{6}$$

$$l_{t,r,s_{r,tc}} \geq \sum_{s \in S_t : tc \in TCS_{t,s} \cap TC^r} dw_{t,s} x_{t,r} \qquad \forall t \in T, r \in R_t, tc \in \bigcup_{s \in S_t} TCS_{t,s}, \tag{7}$$

$$\sum_{r \in R_t} x_{t,r} = 1 \qquad \forall t \in T, \tag{8}$$

$$\sum_{r \in R_t, tc \in TC^r : p_{r,tc} = tc_0} o_{t,r,tc} \geq \sum_{r \in R_{t'}, tc \in TC^r : s_{r,tc} = tc_\infty} o_{t',r,tc} + (ms_{t,t'} + rt_{r,ty_{t'},tc}) x_{t',r} \qquad \forall t, t' \in T : i(t',t) = 1, \tag{9}$$

$$\sum_{r \in R : s_{r,tc_0} = tc} x_{t,r} = \sum_{r \in R_{t'} : p_{r,tc_\infty} = tc} x_{t',r} \qquad \forall t, t' \in T : i(t',t) = 1, tc \in TC_t : p_{r,tc} = tc_0, \tag{10}$$

$$\sum_{tc \in TC_t : \exists r \in R_t, p_{r,tc} = tc_0} sU_{t,tc} \leq \sum_{tc \in TC_{t'} : \exists r \in R_{t'}, s_{r,tc} = tc_\infty} eU_{t',tc} \qquad \forall t, t' \in T : i(t',t) = 1, \tag{11}$$

$$sU_{t,tc} = \sum_{r \in R_t : tc \in TC^r} \left(o_{t,r,ref_{r,tc}} - for_{bs_{r,tc}}\, x_{t,r} \right) \qquad \forall t \in T,\ tc \in TC_t : \\ (\nexists t' \in T : i(t',t) = 1) \vee (\forall r \in R_t : ref_{r,tc} \neq s_{r,tc_0}), \tag{12}$$

$$eU_{t,tc} = \sum_{\substack{r \in R_t : \\ tc \in TC^r}} o_{t,r,tc} + \sum_{\substack{tc' \in TC^r : \\ tc \in OTC_{ty_t,r,tc}}} l_{t,r,tc'} + (rt_{r,ty_t,tc} + ct_{r,ty_t,tc} + rel_{bs_{r,tc}}) x_{t,r} \\ \forall t \in T, tc \in \cup_{r \in R_t} TC^r, \tag{13}$$

$$eU_{t,tc} - M(1 - y_{t,t',tc}) \leq sU_{t',tc} \qquad \forall t, t' \in T, \text{index } t < \text{index } t',\ tc \in TC_t \cap TC_{t'} : \\ i(t,t') \sum_{r \in R_t} e(tc,r) = 0 \ \wedge \ i(t',t) \sum_{r \in R_{t'}} e(tc,r) = 0, \tag{14}$$

$$eU_{t',tc} - M y_{t,t',tc} \leq sU_{t,tc} \qquad \forall t, t' \in T, \text{index } t < \text{index } t',\ tc \in TC_t \cap TC_{t'} : \\ i(t,t') \sum_{r \in R_t} e(tc,r) = 0 \ \wedge \ i(t',t) \sum_{r \in R_{t'}} e(tc,r) = 0. \tag{15}$$

Constraints (3) and (4) force train t to be operated no earlier than $init_t$ on its chosen route and set all track-circuit occupations to 0 on the alternative routes. In Constraints (5), a train starts occupying a given

track-circuit after spending its free-network running time in the preceding one plus the longer stay cumulated there (if the route is used). Constraints (6) and (7) ensure that train t, which stops at station s along route r, does not leave track-circuit $tc \in TCS_{t,s}$ before the scheduled departure time from s, and in any case spends at least its minimum dwell time on tc. In (8), a single route is chosen for train t. Constraints (9), (10) and (11) are used to guarantee consistency for trains using the same rolling stock, i.e. the respect of the minimum separation time between arrival and departure of such trains, the use of the same arrival and departure track-circuit, and the overlapping utilization times of this track-circuit. In Constraints (12), a train's utilization of a track-circuit starts as soon as the train starts occupying track-circuit $ref_{r,tc}$ along one of the routes including it, minus the formation time. Constraints (12) are imposed as inequalities ($\leq$) when they concern a track-circuit of the first block sections of the route ($ref_{r,tc} = s_{r,tc_0}$) and the train t results from the turnaround, join or split of one or more other trains ($\exists t' \in T : i(t', t) = 1$). This is a consequence of the need of keeping platforms utilized. Indeed, if t results from t', Constraints (11) ensure that the track-circuit where the turnaround takes place starts being reserved by t as soon as t' arrives. However, t needs to wait at least for a time ms before departing. The occupation of the track-circuit by t is however starting from its actual departure, for guaranteeing the coherence of the occupation variables and the running time (Constraints (5)). Hence, t's reservation starts much earlier than its occupation. Furthermore, in Constraints (13), the utilization of a track-circuit lasts till the train exits it along any route, plus the release time. If the train is long enough to keep occupying the track-circuit when its head is at the end of the following ones (the ones included in set $OTC_{ty_t,r,tc}$), also the longer stay of the train on these further track-circuits has to be accounted for. Finally, Constraints (14) and (15) ensure that the track-circuit utilizations by two trains do not overlap. This must hold unless they use the same rolling-stock and the track-circuit is at the extreme part of their routes, where the reutilization must take place. We refer the reader to Pellegrini et al. (2015) for additional details about the above formulation.

The objective functions usually considered in the literature are typically:

- the weighted total delays (where w_t is a weight linked to the importance of the train):

$$\min \quad \sum_{t \in T} w_t D_t, \tag{16}$$

$$D_t \geq \sum_{r \in R_t} o_{t,r,tc_\infty} - sched_t \qquad \forall t \in T; \tag{17}$$

- the maximum consecutive delay:

$$\min \quad \Delta, \tag{18}$$

$$\Delta \geq \bar{D}_t \qquad \forall t \in T, \tag{19}$$

$$\bar{D}_t \geq \sum_{r \in R_t} o_{t,r,tc_\infty} - exit_t \qquad \forall t \in T; \tag{20}$$

- the number of delayed trains:

$$\min \quad \sum_{t \in T} \delta_t, \tag{21}$$

$$M\delta_t \geq \sum_{r \in R_t} o_{t,r,tc_\infty} - exit_t \qquad \forall t \in T; \tag{22}$$

- the total travel times of all trains in the control area:

$$\min \quad \sum_{t \in T} \sum_{r \in R_t} (o_{t,r,tc_\infty} - o_{t,r,tc_0}) + W Dentr_t, \tag{23}$$

$$\sum_{r \in R_t} o_{t,r,s_{r,tc_0}} - init_t \leq Dentr_t \qquad \forall t \in T. \tag{24}$$