

Towards a comprehensive model for track allocation and roll-time scheduling at marshalling yards *

Sara Gestrelus ^{a,1}, Martin Aronsson ^a, Martin Joborn ^a, Markus Bohlin ^a

^a RISE SICS AB

Box 1263, SE-164 29 Kista, Sweden

¹ E-mail: sara.gestrelus@ri.se

Abstract

This paper considers multi-stage train formation with mixed usage tracks at a marshalling yard without departure yard. A novel integer programming model for scheduling shunting tasks as well as allocating arrival yard tracks and classification bowl tracks is presented. By taking a comprehensive view of the marshalling yard operations, more effective schedules can be found, and a variety of characteristics can be optimised, including shunting work effort, number or cost of tracks, and shunting task start times. Two different objective functions are evaluated: minimising work effort in terms of wagon pull-backs and minimising track costs. A procedure for finding a hot-start solution with few wagon pull-backs is also presented. The proposed model is tested on real data from Sävenäs marshalling yard in Sweden. The results show that the method is able to return an optimal schedule for a planning period of 4 days if the hot-start solution is optimal or the remaining problem is tractable for the heuristics in CPLEX.

Keywords

Shunting, Marshalling, Railways, Optimisation, Integer programming

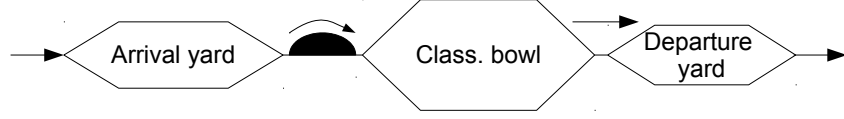
1 Introduction

Freight transportation on railways is realised by unit trains or wagonload trains. Unit trains service origin-destination pairs whose freight volumes are large enough to fill complete trains (point-to-point distribution), while the remaining customers are served by wagonload trains in a hub-and-spoke network.

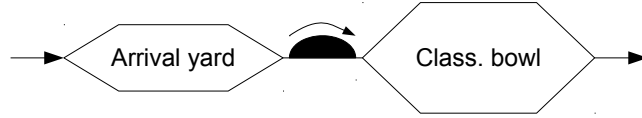
The hubs of the wagonload system are marshalling yards where wagons from inbound trains are sorted into new outbound trains. Effective planning and operation of marshalling yards is therefore a decisive factor for reliable and punctual freight transportation, and the effectiveness of the marshalling yard also affects the overall rail system fluidity (Dirnberger and Barkan (2007)). Unfortunately, scheduling the operations of a marshalling yard is hard as there are different shunting tasks with complex dependencies.

In addition to generating a hard operational planning problem, marshalling yards require expensive infrastructure. The yards consist of an arrival yard and a classification bowl with a sophisticated braking system, and sometimes also a departure yard (see Figure 1). The

*This work was financed by the Shift2Rail Joint Undertaking under the Horizon 2020 program with Grant Agreement number 730813 and by the Swedish Transport Administration (Trafikverket) under grant TRV 2015/102676.



(a) A schematic layout of a hump yard with both arrival and departure yard. The trains depart from the departure yard.



(b) A schematic layout of a hump yard with only an arrival yard. The trains depart from the classification bowl.

Figure 1: Typical layout of hump yards.

arrival yard and classification bowl are often connected by a hump, and a switching system directs the wagons to their assigned classification bowl tracks. The lay-out of the sub-yards affects the work required to sort a given traffic, and it's a daunting task to decide on track investments when building or maintaining a yard. The high cost of tracks and switches must be weighed against the yard's ability to provide the capacity needed to effectively sort the wagons.

In this paper, a novel mixed-integer optimisation model is used for planning the operations of a marshalling yard without a departure sub-yard. The model is based on multi-stage train formation with mixed usage tracks (mixing tracks) and it allocates trains to tracks in the arrival yard and in the classification bowl, and also schedules the shunting tasks. Previous models for train formation with mixing tracks have assumed a given fixed shunting task schedule and only considered classification bowl track allocation. By taking a more comprehensive view of the marshalling yard operations, more effective schedules can be found, and a variety of characteristics can be optimised, including shunting work effort, number or cost of tracks, and shunting task start times.

1.1 Marshalling yard operations

Trains arriving to a marshalling yard are parked on the arrival yard where their wagons are decoupled and inspected. The wagons are then rolled into the classification bowl where they are sorted into new outbound trains. Outbound trains have fixed departure times, and all wagons and the engine must be coupled and all departure preparations must be finished before the departure time of an outbound train. Outbound trains depart from the classification bowl, or if the marshalling yard has a departure yard, they may be rolled to the departure yard and depart from there. Figure 2 shows the timings of the different shunting tasks at a marshalling yard without departure yard (corresponding to Figure 1b), as this is the case considered in this paper.

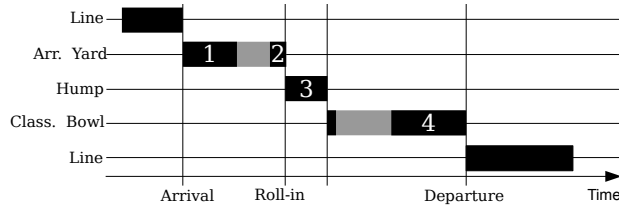


Figure 2: Timing requirements for shunting and work tasks for wagons at different parts of the yard. The graph follows a wagon that arrives to the arrival yard with an inbound train, and is pushed over the hump to a classification bowl track where its outbound train is built. The black boxes represent time durations that have to be respected while grey boxes are non-required occupation time. 1 = Engine decoupling and wagon inspection and preparation. 2 = Shunt engine coupling and push preparation. 3 = Time for rolling over hump. 4 = Coupling, brake preparation and test, time for rolling to line.

The arrival yard and the classification bowl are connected by a hump which allows the wagons to roll from the arrival yard to the classification bowl by means of gravity. A switching system at the end of the hump distributes the wagons on the classification bowl tracks. Further wagon sorting can be accomplished by pulling wagons on a classification bowl track back to the arrival yard and pushing them over the hump once more. There are numerous tactics for sorting the wagons using roll-ins and pull-backs (see e.g. Boysen et al. (2012) or Gatto et al. (2009) for an overview), and this paper considers the practice used in Sweden: multi-stage train formation with mixing tracks (see next section for an introduction, or Bohlin et al. (2015) for a more thorough description).

Multi-stage train formation with mixing tracks

There are two shunting actions for sorting wagons in a marshalling yard without departure yard. The first action is the *roll-in*, which is when all wagons from an inbound train are pushed over the hump and distributed on the classification bowl tracks. The second action is the *pull-back*, which is when all wagons on a classification bowl track are pulled back to the arrival yard and immediately pushed over the hump to be re-distributed on the classification bowl tracks. If possible, pull-backs should be avoided as they require work and hump-time, and wear the yard. In this paper, the number of *wagon pull-backs* is used as a measure for shunting work. The number of wagon pull-backs rather than the number of pull-backs is chosen as e.g. the coupling work effort depends on the number of wagons that take part in each pull-back.

The operational practice considered in this paper is based on designating classification bowl tracks as either *train formation tracks* or *mixing tracks*. The majority of the classification bowl tracks are train formation tracks. Train formation tracks are used for compounding outbound trains, and wagons that have been rolled to a train formation track will remain there until their departure. That is, train formation tracks are never part of a pull-back. A further restriction is that at any point in time, a train formation track may contain wagons for one and only one outbound train. A few classification bowl tracks are designated *mixing tracks*. Mixing tracks are used for temporarily storing wagons until their outbound trains

are being built. All wagons on mixing tracks have to be pulled back and rolled to their train formation track before their departure. In reality the designation of tracks may change as needed, but this paper assumes that tracks are in advance firmly designated as either train formation tracks or mixing tracks.

1.2 Related Work

As marshalling yard operations are critical for an effective wagonload system, and different yards have different operational and physical constraints, a plethora of planning problems exists in the literature. Extensive (but non-comprehensive) classification schemes for marshalling yard planning problems are found in Hansmann and Zimmerman (2008) and Stefano et al. (2007). Further, Gatto et al. (2009) and Boysen et al. (2012), provide overviews of marshalling procedures and sorting methods.

An important characteristic of marshalling yard planning methods is the number of pull-backs allowed. Marshalling methods are either single-stage, where the wagons can be pulled back only once, or multi-stage, where the wagons can be pulled back multiple times (Gatto et al. (2009)). Another dividing factor is the objective function. A common objective is to minimise the dwell-time of wagons on yards (e.g. Jaehn, Rieder and Wiehl (2015); Haahr and Lusby (2016)). The problem formulation of this paper assumes that wagons are pre-scheduled on departing trains and delays are not allowed. Therefore the wagon dwell-time is given. Other common objectives are to minimise the number of tracks required for shunting or the number of pull-backs. The model of this paper could be used to minimise both the number of pull-backs and the number of required tracks, but our main objective is to minimise work effort measured as the number of wagon pull-backs.

Sorting schemes are early planning methods for multi-stage train formation. The four most commonly mentioned sorting schemes, *Sorting by trains*, *Sorting by block*, *Triangular Sorting* and *Geometric Sorting*, are described in e.g. Gatto et al. (2009), Boysen et al. (2012) and Jacob et al. (2011). The sorting schemes take a set of inbound wagons and sort them into outbound trains without considering arrival or departure times. The required track capacity and work effort differs between the schemes, and an overview of the trade-offs can be found in Gatto et al. (2009).

There are many aspects of marshalling yard planning including e.g. track allocation, wagon-to-train allocation, wagon ordering within trains, roll-in times, pull-back times, roll-out times, and personnel and engine scheduling. Often one, or a few, of the aspects rather than all are considered. For example, Jaehn, Rieder and Wiehl (2015) focuses on hump-scheduling in a setting with given wagon-to-train assignments, no wagon ordering in trains, and unrestricted arrival and marshalling yard capacity. Jaehn, Rieder and Wiehl (2015) presents two heuristics and two exact branch-and-bound algorithms for finding a roll-in order that minimises the weighted tardiness of outbound trains. They also model the problem as an integer program and solve it using CPLEX 12.3. Their exact methods are faster than CPLEX and the heuristic tabu search yields near optimal solutions.

Bohlin et al. (2015) presents three integer programming models, D-IP, GC-IP and A-IP, for solving the multi-stage train formation problem with mixing tracks given a wagon-to-train allocation, fixed roll-in, pull-back and roll-out times, and no wagon ordering in inbound or outbound trains. The models find the classification bowl track allocation that minimises the work effort required. The three different integer programming models are tested using real data from Hallsberg marshalling yard for a five-month period from December

2010 to May 2011. Instances of 2 to 5 days are solved, and a rolling horizon framework is implemented for generating schedules for even longer time periods. The best model, AI-IP, returns optimal solutions for all instances within 340s. Gestrelus et al. (2013) introduces an extension to GC-IP for respecting the wagon ordering in outbound trains.

Jacob et al. (2011) establishes a binary representation of a marshalling schedule and use it to analyse various marshalling planning methods. The encoding incorporates the ordering of wagons in the arriving and departing trains, and assumes that the roll-in order has been pre-determined. In Maue and Nunkesser (2009) the binary encoding is used in an integer program to generate marshalling schedules that minimise the number of extra roll-ins. Various operational practices used at Lausanne Triage Shunting Yard are modelled, including two separate yard sections and time-dependent use of tracks. Further, the model can handle uniform classification bowl track capacity and a train dependent upper bound on departure times. The model is tested on real data for one day and on a synthetic data set. An optimal solution is found within 3 minutes for the real data set. In Marton, Maue and Nunkesser (2009) the model is further adapted for the operational practices used at Lausanne Triage Shunting Yard. This more detailed model took 5.75h to solve and prove optimality for a complete day's worth of traffic.

A paper that comprises many of the aspects of the marshalling process is Haahr and Lusby (2016). Similarly to this paper, Haahr and Lusby (2016) considers the marshalling yard comprehensively. However, the problem considered in Haahr and Lusby (2016) is the hump yard block-to-track assignment (HYBA) problem, which is different from the planning problem considered in this paper. In HYBA blocks of wagons are assigned to classification bowl tracks. The blocks stay on the classification bowl tracks until they are pulled out to the departure yard where outbound trains with a specific wagon ordering are built. That is, in HYBA no pull-backs are allowed, a departure yard is required and the wagon order in outbound trains must be respected. Further, there is no pre-determined wagon-to-train allocation but wagons have to be assigned to departing trains going to their destination, and the objective is to minimise the average wagon dwell-time on the yard. Haahr and Lusby (2016) decompose the problem into three interdependent problems: the hump sequencing problem, the block-to-track-assignment problem and the pull-out allocation problem. Heuristics are proposed for all three problems, and optimisation approaches for the first and last. The hump sequence is decided first using rolling horizon planning and a simplified allocation of wagons to departing trains. The hump sequence is then fixed and the algorithm moves forward to greedily assigning blocks to tracks as the inbound trains are rolled in, and, when the classification bowl is full, deciding on a pull-out schedule. The pull-out schedule is either another greedy approach or a MIP that maximises the length of wagons that are moved from the classification bowl to the departure yard. Haahr and Lusby (2016) test different combinations of their heuristics and MIPs on three problem instances with different classification bowl sizes (42, 59 and 58 tracks). A problem instance consists of 42 days, and there are 702 inbound trains and 18 daily outbound trains. To gauge the quality of the returned schedule Haahr and Lusby (2016) propose a method for calculating a lower bound. The results show that using only heuristics gives execution times of less than 2s and yields solutions that have optimality gaps of 30-185%. By using the MIP approach for the pull-out scheduling the solution quality is improved (optimality gaps of 5-14%) but the execution times are prolonged (273-228s), and using both MIP approaches yields the best results (optimality gaps from 4-13%) but takes longest time (507-570s). Haahr and Lusby (2016) also use their model in a more strategic setting, showing that longer trains

and tracks, as well as delaying the train departure times, would improve the average wagon dwell time.

1.3 Contributions

The contributions of this paper are summarised below.

1. A definition of the extended multi-stage train formation problem with mixing tracks for marshalling yards with an arrival yard but no departure yard, E-MSTF(a).
2. An new integer programming model, E-IP, for the E-MSTF(a).
3. A routine for finding a hot-start solution to E-IP.
4. An evaluation of E-IP based on 12 months' worth of data from the Sävsnäs marshalling yard in Sweden. Two different objective functions are tested, one for minimising the shunting work effort and one for minimising the cost of tracks. The execution time and solution quality are presented.
5. A comparison of schedule quality when optimising the roll-in order in combination with the track allocation and when using a first-come first-served roll-in order.

1.4 Limitations

The model presented in this paper is for multi-stage train formation with mixing tracks on a marshalling yard without departure yard. It has been developed for the typical practice used in Sweden, and assumes that the wagon-to-train assignment is given in advance by the train bookings and must not be changed in the marshalling process.

In reality, outbound trains often have an internal wagon ordering that should be respected, but this is not included in the current model. Further, shunting task times are assumed to be constant rather than dependent on the number of wagons handled. The model can handle preparation and roll-times that depends on the number of wagons for arriving and departing trains. However, making pull-back times depend on the number of mixed wagons would require further modelling.

Another simplification when it comes to the mixing track modelling is that all wagons on all mixing tracks are assumed to be pulled-back at the same time. In reality only one or a few of the mixed wagons may have to be pulled back. The number of wagon pull-backs reported in this paper is therefore an upper bound on the number of wagon pull-backs that would have to be executed in real life operations. Further, a pull-back moving wagons to a departing train is assumed to have to start after a train formation track has been allocated to the train. In reality the pull-back may start earlier as long as the wagons are not rolled to the train formation track before it is available. Also, as stated before, the designation of tracks as train formation tracks or mixing tracks is assumed to be predetermined and fixed.

When it comes to track lengths, the arrival yard tracks are assumed to be long enough to accommodate any arriving train, and the classification bowl tracks are split into track groups. A track group consists of tracks of similar length, and the shortest track length is used to decide if a train can be allocated to the track group or not. Using track groups instead of individual tracks makes the allocation problem easier, but may result in less effective marshalling schedules as some track capacity is relinquished.

Finally, engine and personnel resources are assumed to be non-restrictive, and all trains are assumed to depart from the classification bowl without requiring hump or arrival yard capacity.

1.5 Outline of paper

The remainder of the paper is structured as follows. In Section 2, a definition of the problem, E-MSTF(a), is provided, and Section 3 introduces E-IP, a mixed-integer programming model for solving E-MSTF(a). Notably, two different objective functions, one that minimises the number of wagon pull-backs and one that minimises the cost of tracks, are presented in Sections 3.3 and 3.4 respectively. Section 4 outlines the experimental runs, including a procedure for finding a hot-start solution and a procedure for generating schedules with first-come first-served roll-in order. The results are presented and discussed in Section 4.5, and the paper ends with conclusions and suggestions for future work in Section 5.

2 Problem definition

In this section the extended multi-stage train formation problem with mixing tracks for a marshalling yard without a departure yard, E-MSTF(a), is defined. The aim is to schedule all shunting tasks and also allocate trains to arrival yard tracks and classification bowl tracks. The original multi-stage train formation problem was presented in Bohlin et al. (2015).

There are five important *event* types during marshalling. First of all, there are *arrival* events and *roll-in* events of arriving trains. An arrival is a fixed event as it has pre-planned time that the marshalling schedule should respect. On the other hand, a roll-in is a shunting task that should be scheduled, and it doesn't have a pre-planned fixed time but rather a time period during which it can be executed. Let a_i and r_i denote the arrival and roll-in of an inbound train i . For the departing trains there are two important event types. The first type is *build* events, which is when a train formation track is allocated to a departing train, and the compounding of the train can start. Builds are variable events. The second event type is *departure* events. A departure is a fixed event with a pre-planned time that should be respected. Let b_j and d_j denote the build and departure of an outbound train j . The fifth type of event, which is not connected to any specific train, is *pull-back* events. Pull-backs are variable events, denoted $p \in \mathcal{P}$.

Events have timing constraints that must be fulfilled (see Figure 2). The time required between an event e_1 and e_2 is given by $\Delta_{e_1 e_2}$. For a marshalling schedule to be feasible it must provide a time feasible ordering of events that allows for all departing trains to be compounded before their departure time. For example, if wagons for a departing train j are rolled in before the build of j , then there must be a pull-back after the build of j but before its departure.

But from a feasible event schedule a complete marshalling plan also requires track allocations. Assume a marshalling yard with an arrival yard and a classification bowl. The arrival yard consists of T^a number of tracks that are of sufficient length to accommodate any arriving train. The classification bowl consists of $g \in \mathcal{G}$ train formation track groups with unique lengths l_g . Each track group $g \in \mathcal{G}$ consists of N_g tracks, and all tracks have a length of at least l_g . A train formation track is included in one and only one track group, notably the one with the largest length l_g that is smaller or equal to the track length. There are also n^{mix} number of mixing tracks with a combined total length of l^{mix} .

The shunting problem consists of sorting wagons from arriving trains, $i \in \mathcal{A}$, to departing trains, $j \in \mathcal{D}$. The set of arriving trains that a departing train j receives wagons from is given by $\mathcal{A}(j)$, and we call the set of wagons from i to j a *wagon association*, denoted s_{ij} . The length of a wagon association is given by l_{ij} and the number of wagons by $|s_{ij}|$.

Let l_j denote the length of a departing train j . A departing train has to be scheduled on a train formation track that is long enough to accommodate the train, i.e. a departing train j can be scheduled in a train formation track group g if $l_j \leq l_g$. Let $\mathcal{G}(j)$ be the set of track groups that train j fits on.

A train pair (j, k) is trivially schedulable if $t_{d_j} + \Delta_{dd} \leq t_{d_k}$, where t_{d_q} is the departure time of train q and Δ_{dd} is the time required between two departures. The set of all trivially schedulable train pairs is denoted \mathcal{C} . Let $E(j)$ and $F(j)$ be the set of trains that are trivially schedulable before respectively after departing train j . Note that if a train exists in $E(j)$ it can not exist in $F(j)$ and vice versa, and any set of trains that can be scheduled on a track will be totally ordered. For a train j to be scheduled before a train k in the final solution trivial schedulability is only the base requirement, and more requirements on the roll-in times and/or pull-back times must also be fulfilled (see Section 3.3).

3 Optimisation model, E-IP

The optimisation model E-IP is an extension of the model A-IP previously presented in Bohlin et al. (2015). E-IP has three cores: the ordering of events, the allocation of train formation tracks, and the mixing and pull-back scheduling. Section 3.1 presents the event ordering model, including how the order variables are used to ensure that the arrival yard capacity is respected. The modelling of formation track allocation is then presented in Section 3.2. Finally, Section 3.3 presents the mixing and pull-back model which interacts with both the event ordering and the classification bowl track allocation.

3.1 Ordering of events

The events included in E-IP are arrivals $a \in \mathcal{A}$, roll-ins $r \in \mathcal{R}$, pull-backs $p \in \mathcal{P}$, builds $b \in \mathcal{B}$ and departures $d \in \mathcal{D}$. Build events are only necessary for detailed mixing track modelling and will be discussed in Section 3.3 rather than here. Likewise, the ordering of pull-backs is intrinsically connected with the pull-back modelling and will therefore be handled in Section 3.3. This leaves arrivals, roll-ins and departures to be considered in this section.

Roll-ins are *variable events* that can take place any time within a predefined time range. For example, the earliest possible time for a roll-in of arriving train i is $t_{a_i} + \Delta_{ar}$, where t_{a_i} is the arrival time of train i and Δ_{ar} is the time required for preparations before roll-in (inspection and decoupling). Further, if train j is the train with the earliest departure time that train i is associated with, train i must be rolled in at latest at time $t_{d_j} - \Delta_{rd}$ where t_{d_j} is the departure time of train j and Δ_{rd} the minimum time required between roll-in and departure (time for rolling, coupling, and brake preparation and test). The earliest and latest time of a variable event v is denoted T_v^e and T_v^l respectively. Arrivals and departures are fixed events with fixed times. Let T_v denote the fixed time of event v . The fixed events are totally ordered, while the roll-ins are partially ordered among themselves and also partially ordered with fixed events. Due to the partial ordering some events have a predefined order. The set of unordered event pairs is given by $(v, q) \in \mathcal{U}$, and the set of ordered event pairs by

$(v, q) \in \mathcal{O}$. Note that (v, q) is an ordered pair. That is, for unordered events both $(v, q) \in \mathcal{U}$ and $(q, v) \in \mathcal{U}$.

Now, the work time requirements from Figure 2 are ensured by timing limits and timing constraints, and if required a binary order variable and big-M modelling is used to activate the appropriate constraint. In big-M modelling a large constant that dominates the constraint, denoted M , is used together with decision variables to active and de-active constraints. The timing constraints are as follows,

$$t_v + \Delta_{vq} - Mx_{qv} \leq t_q \quad (v, q) \in \mathcal{U} \quad (1)$$

$$x_{qv} + x_{vq} = 1 \quad (v, q) \in \mathcal{U} \quad (2)$$

$$t_v + \Delta_{vq} \leq t_q \quad (v, q) \in \mathcal{O} \quad (3)$$

$$t_v = T_v \quad v \in \mathcal{A} \cup \mathcal{D} \quad (4)$$

$$T_v^e \leq t_v \leq T_v^l \quad v \in \mathcal{R}. \quad (5)$$

Arrival yard capacity

All tracks on the arrival yard are assumed to be long enough to accommodate any train. Therefore the arrival yard capacity constraint is reduced to ensuring that no more than T^a trains are parked on the arrival yard at any point in time. Recall that T^a was the number of arrival yard tracks. Let $N_{arr}^{p+d}(i)$ be the number of trains that might be on the arrival yard when train i arrives. Some trains will definitely be on the arrival yard as their arrival times are so close in time to the arrival of i that their roll-in preparations are not yet finished. For other trains the roll-in preparations are finished but they may still be on the arrival yard as their latest roll-in times have not yet passed. Call this second type of trains *overlapping ready trains*. An overlapping ready train will have a roll-in event that is unordered with the arrival event of i . Given that the arrival event of i is denoted a_i let the set of roll-ins that are unordered with a_i be denoted $\mathcal{U}(a_i)$. Now, if train i is to fit on the arrival yard there must be at least one available arrival yard track, i.e. at least $N_{arr}^{p+d}(a) - T^a + 1$ of the overlapping ready trains must have been rolled in before the arrival of train i ,

$$N_{arr}^{p+d}(i) - T^a + 1 \leq \sum_{r \in \mathcal{U}(a_i)} x_{ra_i} \quad i \in \mathcal{A}. \quad (6)$$

3.2 Train formation track allocation

The core of the train formation track allocation model is the AI-IP model presented in Bohlin et al. (2015) but with track groups. The model is based on binary variables that define direct successors on tracks. The binary variable x_{jk}^g encodes the decision of putting departing train k directly after departing train j on a train formation track in track group g , i.e. $x_{jk}^g = 1$ if train k directly follows train j on a track in g . A train must directly follow one and only one train, and may be directly followed by one and only one train. This will sort the departing trains into trivially schedulable chains that are allocated to specific tracks. Note that the total ordering provided by the trivial schedulability requirement will prevent cycles. Dummy trains u are used to mark the start of a chain. At most N_g train chains can be allocated to a track group g , where N_g is the number of tracks in the track group. That is, at most N_g dummy start trains can be used for a track group g . Further, a dummy train representing the end of a chain is included in the set of trains that can follow a departing

train. The constraints are as follows,

$$\sum_{g \in \mathcal{G}(j)} \sum_{k \in \{E(j): l_k \leq l_g\}} x_{kj}^g = 1 \quad j \in \mathcal{D} \quad (7)$$

$$\sum_{g \in \mathcal{G}(j)} \sum_{k \in \{F(j): l_k \leq l_g\}} x_{jk}^g = 1 \quad j \in \mathcal{D} \quad (8)$$

$$\sum_{j \in \{\mathcal{D}: l_j \leq l_g\}} x_{uj}^g \leq N_g \quad g \in \mathcal{G} \quad (9)$$

$$\sum_{(k,j) \in \{\mathcal{C}: l_k \leq l_g\}} x_{kj}^g = \sum_{(j,k) \in \{\mathcal{C}: l_k \leq l_g\}} x_{jk}^g \quad j \in \mathcal{D}, g \in \mathcal{G}(j). \quad (10)$$

Constraints (7) ensure that one and only one train is scheduled before departing train j . Note that this may be a dummy train if train j is the first train on a track. Constraints (8) ensure that one and only one train is scheduled after departing train j . Once again, this may be a dummy train if train j is the last train in a chain. Constraints (9) ensure that the number of train chains built in each track group are fewer or equal to the number of tracks available in the track group. Finally, constraints (10) ensure that the number of trains preceding a train j in a track group g is the same as the number of trains following it. That is, if train j follows a train on a track in track group g , then there will also be a train following train j in the same track group g . Contrarily, if train j does not follow any train on a track in track group g , then no train will follow j in track group g . This ensures that all trains in a chain are allocated to one and only one track group.

3.3 Modelling mixing and pull-backs

Knowing which wagons that are mixed

If a departing train a is scheduled directly before a train j on a train formation track, then wagons to train j will need to be mixed if they are rolled in before the departure of train a . Let z_{ij} be a binary variable that takes value 1 if wagons from arriving train i to departing train j are mixed. Then,

$$\sum_{g \in \mathcal{G}(j) \cap \mathcal{G}(a)} x_{aj}^g \leq x_{d_a r_i} + z_{ij} \quad (a, j) \in \mathcal{C}, i \in \mathcal{A}(j). \quad (11)$$

Note that z_{ij} can be used in an objective function together with $|s_{ij}|$ to minimise the number of mixed wagons.

Scheduling pull-backs

The time-feasibility of the pull-back schedule has to be ensured. That is, roll-ins and pull-backs must be scheduled without time overlap on the hump.

A total ordering of pull-backs can be assumed without loss of generality, and integer variables for counting the number of executed pull-backs are used to represent the pull-back schedule. Every roll-in and departure is assigned an integer variable, $N_v, v \in \mathcal{R} \cup \mathcal{D}$. The variable defines the number of pull-backs that have started **before** the event v . The N_v variables are constrained to be less than or equal to the maximum number of allowed pull-backs, N^M . Whenever the pull-back variable count value increases between two subsequent events a pull-back must be scheduled to start after the first event start but before the second.

Big-M modelling and the same ordering binaries as in Constraints (1) are used to correctly order and constrain the pull-back count variables,

$$N_v - Mx_{qv} \leq N_q \quad (v, q) \in \mathcal{U}, v, q \in \mathcal{R} \cup \mathcal{D} \quad (12)$$

$$N_v \leq N_q \quad (v, q) \in \mathcal{O}, v, q \in \mathcal{R} \cup \mathcal{D}. \quad (13)$$

Now, both roll-ins and pull-backs use the hump, so a pull-back must be scheduled at least Δ_{rp} time after a roll-in, and a roll-in must be scheduled at least Δ_{pr} time after a pull-back. If two roll-ins have a pull-back between them, there must be at least $\Delta_{rp} + \Delta_{pr}$ time between them. Further, for any additional pull between the two roll-ins Δ_{pp} time is needed. In our case, $\Delta_{pp} < \Delta_{rp} + \Delta_{pr}$. Let $d_1 = \Delta_{rp} + \Delta_{pr} - \Delta_{pp}$. Then,

$$t_{r_1} + (\Delta_{rp} + \Delta_{pr})(N_{r_2} - N_{r_1}) - d_1(N_{r_2} - N_{r_1} - 1) \leq t_{r_2} + Mx_{r_2r_1} \quad (14)$$

$$(r_2, r_1) \in \mathcal{U}, r_1, r_2 \in \mathcal{R}$$

$$t_{r_1} + (\Delta_{rp} + \Delta_{pr})(N_{r_2} - N_{r_1}) - d_1(N_{r_2} - N_{r_1} - 1) \leq t_{r_2} \quad (15)$$

$$(r_1, r_2) \in \mathcal{O}, r_1, r_2 \in \mathcal{R}.$$

Note that the constraints assume that $d_1 \leq \Delta_{rr}$, and that big-M modelling is used for the unordered events in Constraints (14).

If there is a pull-back between a roll-in and a departure, or between a departure and a roll-in, timing constraints for ensuring hump-time feasibility are also needed. Note that in this case the time required for any extra pulls is larger than the time required for the first pull. This is because the time required between a pull-back and a roll-in, or between a roll-in and a pull-back is smaller than the time required between two pull-backs. For example, let Δ_{rp} be the time required between a roll-in and a pull-back, and let $d_2 = \Delta_{pp} - \Delta_{rp}$. Then,

$$t_r + \Delta_{rp}(N_d - N_r) + d_2(N_d - N_r - 1) \leq t_d + Mx_{dr} \quad (16)$$

$$(d, r) \in \mathcal{U}, r \in \mathcal{R}, d \in \mathcal{D}$$

$$t_r + \Delta_{rp}(N_d - N_r) + d_2(N_d - N_r - 1) \leq t_d \quad (17)$$

$$(r, d) \in \mathcal{O}, r \in \mathcal{R}, d \in \mathcal{D}.$$

Similar constraints are included for pull-backs between departures and roll-ins.

Finally, a pull-back start between two departures don't affect the departure times. However, if n pull-backs are scheduled between two departures, and $n > 1$, then there must be time for at least $n - 1$ full pull-backs between the departures as only one pull-back can be carried out at a time. Let Δ_{pp} be the time required to perform a full pull-back. Then,

$$t_v + \Delta_{pp}(N_q - N_v - 1) \leq t_q \quad (v, q) \in \mathcal{O}, v, q \in \mathcal{D}. \quad (18)$$

But from being internally time feasible the pull-back schedule should also provide a feasible pull-back for all mixed wagons. A mixed wagon must be moved from the mixing track to its assigned train formation track at least a time Δ_{pd} before its departure time to allow time for the pull-back and departure preparations, including rolling the train from the train formation track to the line.

Assume that train m has mixed wagons, then a pull must be scheduled after a formation track has been made available for train m . Assume that train a directly precedes train m on its allocated train formation track, i.e. the pair (a, m) is trivially schedulable, $(a, m) \in \mathcal{C}$,

and has been scheduled $x_{am}^g = 1$ on some track group $g \in \mathcal{G}(a) \cap \mathcal{G}(m)$. Then a pull-back has to be scheduled after the departure of train a but before the departure of train m . In other words, the number of pulls scheduled before the departure of m must be one more than the number of pulls scheduled before the departure of a . Let \bar{x}_{am} be a variable that takes value 1 if a is **not** allocated directly before m and let \bar{z}_m be a variable that takes value 1 if m is **not** mixed. The variables \bar{x}_{am} and \bar{z}_m are constrained using the allocation variables from Constraints (7) to (10) and the mixing variables from Constraints (11) respectively. Now, once again using big-M modelling,

$$N_{d_a} + 1 \leq N_{d_m} + M\bar{x}_{am} + M\bar{z}_m \quad (a, m) \in \mathcal{C}. \quad (19)$$

The pull-back that moves wagons to m must be early enough to move the wagons to the train formation track before the departure preparations of m starts. That is, the pull-back time must be earlier than $t_{d_m} - \Delta_{pd}$. A trivial case for when train m can not have mixed wagons is when the departure time of train a is too late for a pull to be carried out, $t_{d_m} - \Delta_{pd} < t_{d_a}$. If this is the case the allocation of the train combination (a, m) will forbid wagons of train m to be mixed,

$$\sum_{g \in \mathcal{G}(a) \cap \mathcal{G}(m)} x_{am}^g = \bar{z}_m \quad (a, m) \in \{\mathcal{C} : t_{d_m} - \Delta_{pd} < t_{d_a}\}. \quad (20)$$

On the other hand, if wagons for train m are indeed mixed the pull moving wagons to m will be pull $N_{d_a} + 1$. Any event v before pull $N_{d_a} + 1$ must therefore be scheduled earlier than $t_{d_m} - \Delta_{pd} - \Delta_{vp}$, where Δ_{vp} is the time required between event v and a pull. The model proposed in this paper will ensure that all events that happen after pull N_{d_a} but before pull $N_{d_a} + 1$ are scheduled early enough to allow pull $N_{d_a} + 1$ to move the mixed wagons. However, if pulls are scheduled back-to-back and the pull time of pull N_{d_a} restricts the start time of pull $N_{d_a} + 1$ rather than any event, then it may not be enough to only consider events between pull N_{d_a} and $N_{d_a} + 1$. However, as this is rarely the case this problem is ignored and the final schedules checked to ensure there are no trains with problematic pulls. In our problem instances there were no problematic pulls, but if there are, further constraints are required to handle pulls before pull N_{d_a} as well.

Now, let $v \in S(a, m)$ be the set of events that may have $N_v = N_{d_a}$ and a time that may be late enough to be problematic; $T_v^l > t_{d_m} - \Delta_{pd} - \Delta_{vp}$. Such potentially problematic events are dealt with in two different constraint sets. First of all, potentially problematic events that happens after the departure of a are considered. To deal with this situation timing constraints are included. The constraints use big-M modelling to cancel the constraint if 1) m is not allocated after a , 2) m is not mixed, 3) $N_v > N_{d_a}$ or 4) event v is before the departure of a . In cases 1-2 the constraint does not need to be respected as the event is not problematic for the combination (a, m) , in case 3 the event happens after the pull moving the mixed wagons, and finally in case 4 the event happens before the departure of a and is dealt with in a different way. This results in the constraints,

$$t_v + \Delta_{vp} + \Delta_{pd} \leq t_{d_m} + M(\bar{x}_{am} + \bar{z}_m + (N_v - N_{d_a}) + (N^M + 1)x_{vd_a}) \quad (21)$$

$$(a, m) \in \mathcal{C}, v \in S(a, m), (v, d_a) \in \mathcal{U},$$

if d_a and v are unordered. Equivalent constraints are used if they are ordered.

The constraints above only handle potentially problematic events that happen after the departure of a . However, roll-ins that happen before d_a may also be problematic as it takes

more time to perform a roll-in and a pull-back than a departure and a pull-back. Therefore constraints for potentially problematic roll-ins that take place before the departure of a are included,

$$t_r + \Delta_{rp} + \Delta_{pd} \leq t_{dm} + M (\bar{x}_{am} + \bar{z}_m + (N_{da} - N_r) + (N^M + 1) x_{da,r}) \quad (22)$$

$$(a, m) \in \mathcal{C}, r \in S(a, m) \cap \mathcal{R}, (d_a, r) \in \mathcal{U},$$

if d_a and r are unordered. Equivalent constraints are used if they are ordered.

Which wagons are in which pull-back

The pull-back modelling presented so far will generate a schedule for infinite mixing track capacity. However, if the mixing track capacity is likely to be constraining, or if the number of wagon pull-backs is to be minimised, further modelling for assigning wagons to pulls is required.

Assume that train m may have mixed wagons from arriving train i . That is, the association s_{im} may be mixed. Recall that the pulls are totally ordered. Binary variables z_i^p are used to determine the first pull p that s_{im} may take part in, i.e. the first pull after the roll-in of i . Let n_p be the index number of pull $p \in \mathcal{P}$, i.e. the first pull has $n_1 = 1$. Then,

$$\sum_{p \in \mathcal{P}} n_p z_i^p = N_{r_i} + 1 \quad i \in \mathcal{A} \quad (23)$$

$$\sum_{p \in \mathcal{P}} z_i^p = 1 \quad i \in \mathcal{A}. \quad (24)$$

Another binary variable, z_m^p , is included to determine the first pull that association s_{im} can not be part of, i.e. the second pull after the build of m . Note that the first pull after the build of m will move all wagons from the mixing track to the train formation track. Let N_{b_m} be an integer pull-back count variable for the build of train m equivalent to the ones used for roll-ins and departures in Constraints (12) - (13). Then, big-M modelling can be used to set N_{b_m} ,

$$N_{da} - M \bar{x}_{am} \leq N_{b_m} \quad (a, m) \in \mathcal{C} \quad (25)$$

$$N_{b_m} - M \bar{x}_{am} \leq N_{da} \quad (a, m) \in \mathcal{C}, \quad (26)$$

and using the same type of constraints as in (23)-(24) the first pull that association s_{im} may not be present in, i.e. the second pull after the N_{b_m} pull, can be identified,

$$\sum_{p \in \mathcal{P}} n_p z_m^p = N_{b_m} + 2 \quad m \in \mathcal{D} \quad (27)$$

$$\sum_{p \in \mathcal{P}} z_m^p = 1 \quad m \in \mathcal{D}. \quad (28)$$

Note that we require two more pulls in \mathcal{P} than the maximum number of allowed pulls, but these extra pulls will never be used.

Now, for an association to be present in pull p three conditions need to be fulfilled. First of all the association must be mixed. Secondly, pull p must be after the roll-in of i , i.e. some $z_i^q = 1, q \leq p$. If this is not true the wagons would still be on the arrival yard during pull p .

Finally, pull p must be before the second pull after the build of m , i.e. $z_m^q = 0, \forall q \leq p$. Let z_{im}^p be a binary variable that takes value 1 if association s_{im} is present in pull p , then,

$$\sum_{q=1}^p z_i^q - \sum_{q=1}^p z_m^q + z_{im} - 1 \leq z_{im}^p \quad m \in \mathcal{D}, i \in \mathcal{A}(m), p \in \mathcal{P}.$$

The capacity constraint of the mixing tracks can now be included by using the binary variables z_{im}^p and the length of the association l_{im} . As the number of wagons on mixing tracks increases until a pull-back is carried out it is sufficient to constrain the length of associations taking part in each pull-back. Let l^{mix} be the total length of the mixing tracks, then,

$$\sum_{m \in \mathcal{D}} \sum_{i \in \mathcal{A}(m)} l_{im} z_{im}^p \leq l^{mix} \quad p \in \mathcal{P}.$$

Similarly, if $|s_{im}|$ is the number of wagons in an association, the objective function for minimising the number of wagon pull-backs is,

$$\min \sum_{p \in \mathcal{P}} \sum_{m \in \mathcal{D}} \sum_{i \in \mathcal{A}(m)} |s_{im}| z_{im}^p.$$

3.4 Evaluating track capacity requirements

The number, or cost, of tracks required in the arrival yard and the classification bowl could be minimised. The number of arrivals tracks can be included in the objective function by making T^a a variable. To minimise the number of classification bowl tracks the number of tracks in each track group, N_g , can be made variable and included in the objective function. Assume that the relative cost for building and maintaining an arrival track (classification bowl track) is c_{arr} (c_{class}^g). The objective function that minimises the track cost is then,

$$\min \quad c_{arr} T^a + \sum_{g \in \mathcal{G}} c_{class}^g N_g. \quad (29)$$

Note that different classification bowl track groups can have different costs. Further, as the number of pull-backs is limited the shunting work effort will be limited. However, it may be sensible to include additional constraints to limit the increase in work when minimising the track cost.

4 Experiments

4.1 Data

The experiments are based on data provided by Green Cargo AB, the operator of the Sävsnäs marshalling yard. All wagon assignments for a year in combination with the yearly timetable were used as input data. However, the data had some inconsistencies caused by e.g. multiple bookings of the same wagon and wagons being re-booked on trains that were running late or early in real operation. Therefore the data had to be cleaned.

The problem instances consist of 4 days each, and the wagons of all departing trains whose first wagon arrives to the marshalling yard during that time period are included.

The yard is assumed to be empty at the start of each problem instance. For minimising the number of wagon pull-backs the entire year is divided into 4 day instances, while for minimising the track costs (which takes longer time) a four day instance starting every 27 days was included. A step size of 27 days was used as this returns instances for different week-day combinations approximately once per month. The number of allowed pull-backs was set to 8, i.e. on average two per day.

The classification bowl track lengths varied from 360m to 829m. If each track is represented by its length, the track groups were as follows: {360, 362, 394}, {434, 472, 479, 481}, {522, 593, 597}, {601, 603, 609, 611, 628, 658, 685, 687, 695, 695, 695, 697}, {711, 712, 741, 741}, {806} and {829}. Two tracks of length 541m each were used as the mixing tracks.

Finally, timing estimates were taken from Alzén (2006), where the timings of shunting tasks have been measured for another Swedish marshalling yard, Hallsberg marshalling yard.

4.2 Optimisation Set-up

CPLEX 12.2 was used to solve the optimisation problems. The problems consisted of the constraints presented in this paper and some extra valid inequality constraints. The Network optimiser was used for solving the root node, while the automatic method option of CPLEX was used for the other nodes. For minimising the number of wagon pull-backs an execution time of 30 minutes was set for E-IP, and when minimising the track cost 60 minutes was used. All classification bowl tracks were assumed to have the same cost c_{class} , and experimental runs were made for three different combinations of arrival and classification track costs: $\{c_{arr} = 10, c_{class} = 20\}$, $\{c_{arr} = 10, c_{class} = 0\}$, and finally $\{c_{arr} = 0, c_{class} = 10\}$.

4.3 Generating a hot-start solution

As it sometimes takes over an hour for CPLEX to find the first feasible solution to E-IP, a routine for finding a hot-start solution was devised. The routine starts by solving a reduced problem, *R-IP*, that consists of Constraints (7) to (10), and where a fixed cost $c_{(k,j)}$ is used for scheduling a train k directly before train j . The aim of R-IP is to find a classification bowl allocation with few mixed wagons, and the cost used was the smallest number of wagons that have to be mixed if the train pair is scheduled. That is, all trains were assumed to be rolled in as late as possible. This problem is very similar to the one solved in Bohlin et al. (2015), but with track groups and unrestricted mixing track capacity. Further, the time limit used to decide if a train pair is schedulable or not was increased from 44 minutes to 156 minutes to improve the chance of finding a feasible solution to E-IP using the R-IP classification bowl track allocation.

The classification bowl track allocation from the solution of R-IP may not be feasible when the complete E-IP problem is considered. To find a feasible hot-start solution to E-IP the track allocations from R-IP were softly constrained by penalising track changes, and the number of track changes minimised. Call this problem SE-IP. CPLEX was set to return the first feasible solution found, and the optimisation was then terminated.

If a feasible classification bowl track allocation was found the hot-start solution can be further improved. This is done by fixing the feasible track allocation and starting an

optimisation with the original objective function. In this paper, hot-start solutions were improved for minimising the number of wagon pull-backs, but not for minimising the track cost, as minimising the track cost requires the track allocation to change. A maximum time limit of 500s was set for improving the hot-start solution.

4.4 First-come first-served roll-in order

In order to gauge the effects of comprehensive marshalling yard planning the results from solving E-IP are compared to a marshalling yard schedule where the roll-in order is the same as the arrival order. The exact times are not fixed, but rather the event order variables. Hot-start solutions were generated using the procedure in Section 4.3, but an execution time limit of an hour was set for SE-IP. If no feasible solution to SE-IP had been returned within an hour E-IP was started without a hot-start.

4.5 Results

Finding a hot-start solution

Table 1 shows the number of trains, wagons and associations of each problem instance, as well as how many CPU seconds it took to solve R-IP and SE-IP for an unrestricted roll-in order. As can be seen R-IP was solved in less than three seconds for all problem instances. Further, solving SE-IP took less than 54 CPU seconds for all instances but from one that took 1296 CPU seconds.

Minimising the work effort

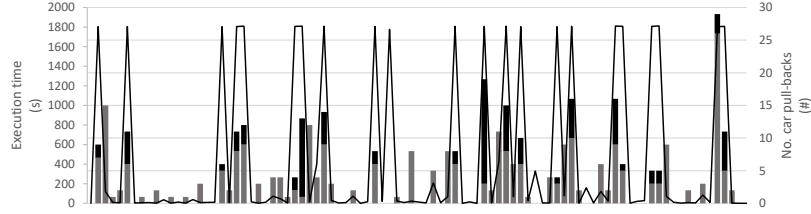
E-IP returned a proven optimal solution within 30 minutes in 70 out of 91 instances when minimising the number of wagon pull-backs. Figure 3a shows the number of wagon pull-backs in the returned solutions, and as can be seen E-IP is generally solved very quickly for problem instances that require no or little mixing. This is often due to the hot-start solution being optimal. The hot-start solution was optimal in 43 instances, and it took 7.3 CPU seconds on average for CPLEX to prove that a hot-start solution was optimal. For the remaining 48 instances E-IP failed to return a proven optimal solution within 30 minutes for 21 instances, and for the other 27 instances the average execution time was 160 CPU seconds. In general, the lower bound of the model seems to be tight with respect to the number of wagon pull-backs but it takes too long to solve the problem in the nodes of the branch-and-bound tree, and few nodes are solved before the time limit is reached. Therefore we conclude that E-IP might be effective for proving optimality, but is too slow to find solutions on its own. Further, the procedure for generating a hot-start solution is very effective for producing schedules of high quality with respect to the number of wagon pull-backs.

The results from minimising the number of wagon pull-backs given a first-come first-served roll-in order are shown in Figure 3b. The results are remarkably worse than for an unrestricted roll-in order. Note that the scale on the right hand axis measuring the number of wagon pull-backs is 10 times larger for the first-come first-served bar chart than for the unrestricted case. Also, no solution was found for 14 instances when restricting the roll-in order. The poor solution quality can be explained by inadequate hot-start solutions, but even the lower bounds of the restricted roll-in order problem was higher than the solution returned

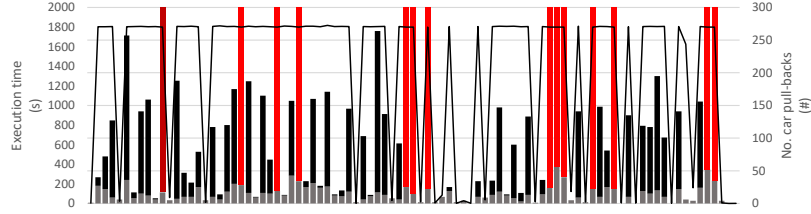
Table 1: The number of arriving trains (Arr.), departing trains (Dep.), wagons (W.) and associations (A.) in each problem instance, and also much time it took to solve R-IP and SE-IP. The instance name is the start day of the instance, counted from 1 January 2014.

Inst.	Arr. (#)	Dep. (#)	W. (#)	A. (#)	R-IP (CPU s)	SE-IP (CPU s)	Inst.	Arr. (#)	Dep. (#)	W. (#)	A. (#)	R-IP (CPU s)	SE-IP (CPU s)
1	42	62	845	217	0.5	1.3	185	59	77	1048	295	0.4	8.0
5	62	85	1292	331	0.5	6.2	189	94	117	1834	468	1.5	8.3
9	69	92	1434	404	0.7	17.3	193	55	70	999	260	0.4	4.9
13	118	123	2051	566	2.0	7.2	197	73	98	1342	354	0.6	7.5
17	61	80	1089	323	0.3	9.0	201	56	88	1136	280	0.6	8.3
21	119	137	2276	635	3.6	25.2	205	51	65	992	243	0.3	2.7
25	66	87	1306	356	0.5	4.8	209	90	109	1552	408	1.2	12.1
29	93	109	1726	498	1.3	7.6	213	50	63	833	211	0.2	2.5
33	94	111	1689	476	1.4	6.2	217	93	121	1830	497	1.6	1295.6
37	66	83	1303	367	0.5	3.1	221	62	90	1204	310	0.4	4.1
41	117	125	2290	601	1.4	15.1	225	91	115	1841	483	1.4	12.3
45	60	77	1226	330	0.5	3.5	229	89	114	1761	480	1.4	26.0
49	118	131	2183	559	2.2	8.1	233	70	87	1199	350	0.6	6.4
53	67	87	1153	343	0.5	5.3	237	108	128	1904	532	2.2	16.0
57	94	106	1871	493	0.9	11.3	241	62	77	1186	337	0.3	6.9
61	89	112	1753	494	1.4	7.9	245	115	132	2257	632	2.0	21.3
65	62	87	1383	362	0.6	9.1	249	66	86	1301	344	0.5	4.1
69	113	123	2216	576	1.5	7.7	253	92	110	1831	519	1.2	8.1
73	58	81	1223	327	0.4	4.0	257	92	107	1731	481	1.1	9.4
77	111	133	2179	592	2.3	7.8	261	65	92	1302	380	0.7	6.7
81	67	94	1347	387	0.8	5.6	265	114	132	2205	613	2.3	53.1
85	93	114	1871	514	1.7	11.5	269	63	82	1157	331	0.5	4.0
89	87	111	1745	465	1.0	13.4	273	106	127	1946	551	2.1	26.1
93	63	87	1266	354	0.7	6.1	277	54	79	891	247	0.4	3.6
97	112	122	2207	589	1.4	8.5	281	87	111	1745	468	1.7	12.2
101	60	81	1252	349	0.4	13.1	285	86	115	1636	447	2.0	6.4
105	92	112	1856	537	1.1	20.1	289	65	89	1258	348	0.6	6.1
109	29	53	569	152	0.1	18.8	293	115	128	2125	571	1.7	10.9
113	92	113	2039	570	1.3	11.2	297	61	77	1107	316	0.4	3.6
117	89	116	1774	496	1.4	26.6	301	109	128	1954	553	1.9	16.3
121	37	65	818	237	0.3	5.2	305	63	79	1109	313	0.4	4.0
125	119	127	2293	634	2.3	45.3	309	88	109	1545	465	1.3	9.6
129	60	80	1298	358	0.5	8.5	313	88	112	1621	456	1.5	11.3
133	113	134	2131	602	2.4	11.7	317	65	88	1221	346	0.6	5.5
137	65	79	1225	356	0.7	12.2	321	115	125	2175	566	1.7	15.6
141	89	105	1798	486	1.0	8.5	325	58	69	1058	303	0.3	3.1
145	82	109	1563	450	1.0	8.1	329	115	133	2076	597	2.2	17.0
149	36	56	787	220	0.1	4.2	333	66	83	1156	322	0.4	8.0
153	114	121	2154	590	1.4	7.5	337	91	104	1608	468	0.9	9.6
157	38	61	728	215	0.1	2.7	341	87	111	1745	475	1.3	11.2
161	113	132	2264	608	2.0	20.2	345	70	91	1376	404	0.8	12.0
165	67	86	1265	355	0.4	8.3	349	112	132	2047	591	2.3	39.2
169	60	82	1254	360	0.5	7.8	353	59	72	1064	304	0.2	2.8
173	85	106	1629	433	0.8	13.8	357	36	48	604	173	0.2	1.3
177	65	89	1401	379	0.6	6.0	361	48	48	698	160	0.1	2.3
181	109	123	2191	601	1.4	5.5							

by E-IP with no roll-in restriction for all instances but from two. The lower bound when enforcing a first-come first-served roll-in order is on average 11.5 wagon pull-backs higher than the one for an unrestricted case. This indicates that it is beneficial to comprehensively plan marshalling yards.



(a) Results for minimising the number of wagon pull-backs with unrestricted roll-in order.



(b) Results for minimising the number of wagon pull-backs given a first-come first-served roll-in order. Red bars reaching the 300 mark indicates that no solution was found.

Figure 3: Results for minimising the number of wagon pull-backs. The line is the execution time (left axis) and the bars show the number of wagon pull-backs (right axis). The grey bar shows the lower bound after 30 minutes of optimisation and the black bar the returned solution. If the optimal solution is returned the black bar is not visible. Note that the right hand axis scale is 10 times larger for the (b) chart. The time for finding a hot-start solution is not included.

Evaluating track costs

When minimising the track cost with costs $\{c_{arr} = 10, c_{class} = 20\}$ the optimal solution was never returned within the time limit of 60 minutes. The average gap was 66%, with the smallest being 57% and the largest 69%.

In Figure 4 the number of arrival tracks (dark grey bar), the number of classification bowl tracks (light grey bar) and the execution times (black line) are shown for the different problem instances. The dashed red lines show the number of available tracks. In 9 instances all tracks are used in the returned schedule, while in 5 instances the schedule uses fewer arrival and/or classification bowl tracks. Note that as the solution was never proven optimal, the execution time is always the time limit of 60 minutes. Experimental runs were also made for minimising either the number of arrival tracks (classification bowl tracks have a cost of 0) or the number of classification bowl tracks (arrival tracks have a cost of 0).

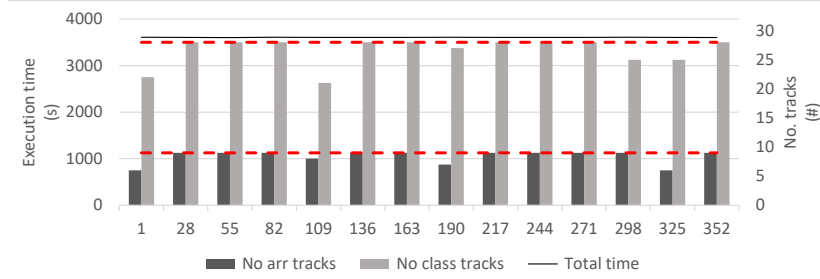


Figure 4: Results from minimising the tracks cost. The light grey bar represents the classification bowl tracks (right axis) and the dark grey the number of arrival yard tracks (right axis). The black line shows the execution time (left axis). The dotted red lines show the number of available arrival and classification bowl tracks.

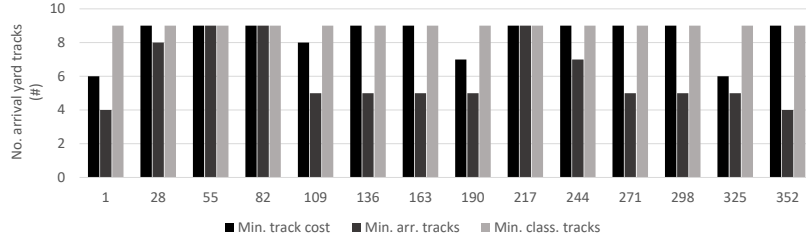


Figure 5: Number of arrival tracks required when minimising different objective functions.

When minimising the number of arrival tracks the optimal solution was returned within 60 minutes for 10 out of 14 problem instances. When minimising the number of classification bowl tracks no optimal solution was returned, and the instances had an optimality gap of on average 70%, with the smallest gap being 63% and the largest 75%.

Figure 5 shows the number of arrival tracks utilised in the final schedules when minimising the track cost (black), the number of arrival tracks (dark grey) and the number of classification bowl tracks (light grey). Figure 6 shows the number of classification bowl tracks for the same objective functions. As can be seen, the choice of objective function does affect the number of tracks, and especially minimising the number of arrival yard tracks yields schedules with different track usage. Although all tracks are often required in the schedules returned, there are some time periods when the traffic can be shunted using fewer tracks. Identifying such periods may be useful for e.g. maintenance planning.

The inability of E-IP to return an optimal solution when minimising the track cost is in line with the reasoning above regarding the model's effectiveness. Very few nodes are solved before the execution time limit, so either a good hot-start solution must be provided or the remaining problem must be tractable for the heuristics in CPLEX. As the hot-start solutions were generated with the aim of limiting the number of mixed wagons, and no improvements were done before submitting them to E-IP, they are poor with respect to the track cost. For minimising the number of arrival tracks the optimal solution is often returned in very few iterations despite the poor hot-start. However, the solution is often returned by the heuristics in CPLEX rather than the branch-and-bound optimisation.

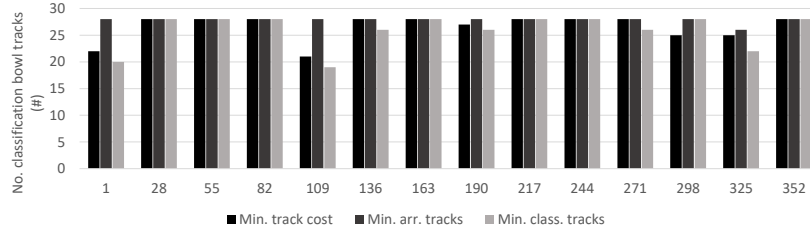


Figure 6: Number of classification bowl tracks when minimising different objective functions.

5 Conclusion and future work

This paper introduced an integer programming model (E-IP) for multi-stage train formation with mixing tracks in a marshalling yard without departure yard. The model schedules all shunting tasks and allocates arrival and classification bowl tracks to arriving and departing trains. The paper considered two different objective functions: one for minimising the work effort in terms of wagon pull-backs and one for minimising the track costs.

Comprehensively planning a marshalling yard is hard but beneficial. If all aspects of the marshalling yard planning can be handled in unity, the schedule quality and consistency can be improved. Allowing any roll-in order rather than enforcing a first-come first-served order decreased the lower bound from 14.6 to 3.1 wagon pull-backs, indicating that comprehensive planning is indeed beneficial.

Further, the results showed that while E-IP seems to provide a tight lower bound for minimising the number of wagon pull-backs, it takes a very long time to solve and hot-starting is favourable. Therefore, a procedure for generating a hot-start solution with few mixed wagons was designed and presented. For a schedule with unrestricted roll-in order, the hot-start solution was optimal in 43 out of 91 cases when minimising the number of wagon pull-backs. In 27 of the remaining 48 cases an optimal solution could still be returned within 30 minutes.

When it comes to minimising the track cost, the hot-start solutions were poor. For minimising the number of arrival tracks the problem was tractable for the heuristics in CPLEX, and an optimal solution was returned within 60 minutes for 10 out of 14 problem instances. However, when minimising the cost for classification bowl tracks the optimal solution was never returned within 60 minutes.

There are a few remaining issues that would be interesting to pursue. First of all, extending the model such that back-to-back pull-backs are handled is an important step towards completeness. Further, investigating the extent to which the track-group simplification decreases schedule effectiveness would be interesting. More substantial extensions would be to include a departure yard and more detailed mixing track modelling, and also to allow for wagon-orderings in inbound and outbound trains.

A different area that calls for work is the question of how to generate high quality hot-start solutions for different objective functions, as well as designing other solution procedures for evaluating the number of tracks required. For example, minimising the number of wagon pull-backs for hypothetical yards with different but fixed track set-ups might be a more apt way to gauge track requirements.

References

- Alzén, C., 2006. *Handbok BRÖH 313.00700: Trafikeringsplan Hallsbergs rangerbangård*. Banverket.
- Bohlin, M., Gestrelus, S., Dahms, F., Mihalák, M., and Flier, H. 2015. "Optimization methods for multistage freight train formation.", *Transportation Science*, 50(3):823–840.
- Boysen, N., Fliedner, M., Jaehn, F., and Pesch, E., 2012. "Shunting yard operations: Theoretical aspects and applications." *European Journal of Operational Research*, 220(1):1–14.
- Dirnberger, J. and Barkan, C., 2007. "Lean railroading for improving railroad classification terminal performance: bottleneck management methods." *Transportation Research Record: Journal of the Transportation Research Board*, (1995):52–61.
- Gatto, M., Maue, J., Mihalák, M., and Widmayer, P., 2009. "Shunting for dummies: An introductory algorithmic survey." In: R. Ahuja, R. Mhring, and C. Zaroliagis, editors, *Robust and Online Large-Scale Optimization*, volume 5868 of *Lecture Notes in Computer Science*, pages 310–337. Springer Berlin / Heidelberg.
- Gestrelus, S., Dahms, F., and Bohlin, M., 2013. "Optimisation of simultaneous train formation and car sorting at marshalling yards." In *Proceedings of the 5th International Seminar on Railway Operations Modelling and Analysis (RailCopenhagen)*, Copenhagen, Denmark.
- Haahr, J. T. and Lusby, R. M., 2016. "A matheuristic approach to integrate humping and pullout sequencing operations at railroad hump yards." *Networks*, 67(2):126–138.
- Hansmann, R. S. and Zimmermann, U. T., 2008. "Optimal sorting of rolling stock at hump yards." In: H.-J. Krebs and W. Jager, editors, *Mathematics Key Technology for the Future*, pages 189–203. Springer Berlin Heidelberg.
- Jacob, R., Márton, P., Maue, J., and Nunkesser, M., 2011. "Multistage methods for freight train classification." *Networks*, 57(1):87–105.
- Jaehn, F., Rieder, J., and Wiehl, A., 2015. "Minimizing delays in a shunting yard." *OR Spectrum*, 37(2):407–429.
- Márton, P., Maue, J., and Nunkesser, M., 2009. "An Improved Train Classification Procedure for the Hump Yard Lausanne Triage". In: J. Clausen and G. D. Stefano, editors, *9th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems (ATMOS'09)*, volume 12 of *OpenAccess Series in Informatics (OASICS)*, Dagstuhl, Germany. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
- Maue, J. and Nunkesser, M., 2009. *Evaluation of computational methods for freight train classification schedules*. Technical report, ARRIVAL Project.
- Stefano, G. D., Maue, J., Modelski, M., Navarra, A., Nunkesser, M., and van den Broek, J., 2007. *Models for rearranging train cars*. Technical Report ARRIVAL-TR-0089, ARRIVAL Project.