Received: 16 February 2022 | Revised: 4 May 2022 | Accepted: 19 May 2022

DOI: 10.1049/itr2.12216

IET Intelligent Transport Systems

The Institution of Engineering and Technology WILEY

REVIEW

Train management in freight shunting yards: Formalisation and literature review

Samuel Deleplanque[1] | **Pierre Hosteins**[2] | **Paola Pellegrini**[3] | **Joaquin Rodriguez**[2]

[1]CNRS, Centrale Lille, JUNIA, Univ. Lille, Univ. Valenciennes, Lille Cedex, France

[2]Univ Gustave Eiffel - ESTAS-COSYS, Lille, France

[3]Univ Gustave Eiffel - LEOST-COSYS, Lille, France

Correspondence

Samuel Deleplanque, CNRS, Centrale Lille, JUNIA, Univ. Lille, Univ. Valenciennes, UMR 8520 IEMN, 41 boulevard Vauban, Lille Cedex 59046, France.
Email: samuel.deleplanque@junia.com

Pierre Hosteins, Univ Gustave Eiffel - ESTAS-COSYS, Lille, France.
Email: pierre.hosteins@univ-eiffel.fr

Funding information

European Union Shift2Rail Joint Undertaking, OptiYard project, Grant/Award Number: 777594

Abstract

This paper treats theoretical and practical aspects of train management in freight shunting yards. It is a literature survey extending the previous ones and presenting the new important papers published in the last decade. The operations realised in such yards are formalised by modelling them in unified modelling language diagrams and the algorithms implemented to solve the optimisation problems of operations management for the different types of freight shunting yards and for the existing models are presented. The details on real yards that can be found in the literature are also reported, to allow the reader understanding what case studies can be actually tackled with the existing methods.

1 | INTRODUCTION

The literature on railway system optimisation is vast. Some of the main problems are called line planning, timetabling, platforming, rolling stock, and crew scheduling [1]. These problems typically consider passenger railway systems. However many studies on less known problems exist. In this work, we propose a review on recent literature dealing with freight railway systems.

European rail freight traffic has been in steady decline for about 25 years [2]. However, political authorities as the European Union continue to push for the evolution of this mode, mainly for improving social and environmental aspects. For example, new technologies are to be introduced in freight shunting yards (named simply yards in the rest of the paper), which are areas where operations are performed to create and recombine freight trains. Nowadays these yards automate some of their operations (e.g. through automatic switches and automatic brakes) but the integration of optimisation tools into their management is still lacking.

We focus on a type of yard operating Car Load services, as opposed to full train load ones [3]. In the former, trains must be split and cars recombined to form new trains since they do not share the same origin and/or destination. In the latter, trains keep their structure throughout the whole journey since all cars have the same origin and the same destination. Indeed, when full train load services need to stop in yards, the operations to which they are submitted can be seen as a subset of those performed on car load ones.

Even though the full chain of operations in a shunting yard is a complex process, it can be generally summed up by the following basic operations: an incoming train is stored on a receiving track; after inspection its cars and locomotive are decoupled; the cars are pushed over an artificial hill named hump and roll towards classification tracks where they are coupled again; finally a new train is formed and stored on a set of departure tracks waiting for its departure time.

The efficiency of yards has an impact on the fluidity of the global rail system and on trains travel time. In ref. [4], it is stated that the percentage of yard time compared to total travel time of a train may be from 10% to 50%. Efficiently managing yards may allow to decrease this percentage. Today, this efficiency is sought mostly manually by operators, and practically no intelligent decision-support tool is used in real cases.

This paper reviews the state of the art on yard management, following refs. [5] and [6]. Many of the existing algorithms are interesting mostly from a theoretical perspective but are not necessarily practical in real-world yards. This is due to the fact that only very specific sub-problems are typically tackled, neglecting the complexity of the overall operations [7]. Nevertheless, a few studies consider more comprehensive models (i.e. models covering additional operations performed in yards), thus moving a step closer to real operational conditions. We report the analysis of these studies after detailing the ones on the sub-problems they partially build on.

In the presented literature, the problems tackled are often presented with little description of their role in the overall yard management process. Moreover, various terms are used to designate the same concepts. To provide a systematic and comprehensible review of the state of the art, we first propose a formal description of yard management elements and processes. To the best of our knowledge, such formalisation has never been presented before, and it constitutes an original contribution of this paper. In particular, we start our analysis with the description of the layout, rolling stock and operations through unified modelling language (UML) diagrams. UML is a standard language for specifying, visualising, constructing, and documenting systems and processes. As further contributions, we summarise the main real case studies tackled in the literature to allow the reader picturing the difficulty of solvable instances. Moreover, we propose a glossary which aims at bringing together the main terms used in the papers, underlying the many synonyms often used in the literature.

The rest of this paper is organised as follows. In Section 2, we formalise the main elements and operations which typically characterise yards and can differ from one yard to another. In Section 3, we highlight the differences between the numerous models and optimisation methods for the management of typical yards. In Sections 4 and 5, we review the recent contributions on yards management while in Section 6, we briefly review important recent papers on other types of freight yards. In Section 7, we gather data found in the literature on real yards (e.g. number of tracks, number of trains treated per day etc.) while we report in an appendix the general glossary used in the cited literature.

2 | LAYOUTS, ROLLING STOCK, AND OPERATIONS IN YARDS

Although many variants exist, the general sequence of operations in a yard can be described as follows. When a train arrives, it is stored on a first set of tracks where its cars are inspected. Then, the locomotive is stored to be coupled later with a new train and the cars are detached from each other. A shunting locomotive arriving behind the cars pushes them over a hump. Rolling down from the hump with the aid of gravity, the cars are classified on tracks dedicated to forming the new trains they will leave the yard on. The route setting is typically automated through a car identification system. Once a new train is formed, it is stored on a last set of tracks to wait for the right time to enter the rail network.

In addition to the way operations are organised, the efficiency of yard management is directly linked to the particular layout of the infrastructure [8]. Even if the layout considered in the literature is often the same, there is a great variety of yards in reality. Figure 1 represents a typical yard layout where the tracks can be divided into four different parts: receiving tracks, hump, classification tracks, and departure tracks. Moreover, yards can have a return track which connects the classification to the receiving tracks. The presence and characteristics of these parts differentiate yard layouts.

Trains entering a yard are called inbound trains. They are stored in the receiving tracks, and once the cars have been uncoupled from their locomotive and are detached from each other, the series of cars on a receiving track is called a cut. Train locomotives are mainly used to enter and leave the yard, while shunting locomotives are used to push and pull cars along yard tracks. The new trains which are built on the classification tracks and which will leave the yard through the departure tracks are called outbound trains.

To formalise the yard features that are relevant for the analysis of the literature, we use UML. It is based on the description of a system through a conceptual model, made of concepts and their relationships. On the one hand, entities are statically described through classes. They are organised thanks to inheritance relations: specific classes incorporate the structure and behaviour of the more general ones. Classes and inheritances are grouped into class diagrams. On the other hand, activity diagrams, are dynamic views of the system showing the characteristics of the processes that are performed. Through this pictorial language, we can achieve a high level of formalisation without the need of a very technical codification.

Figures 2 and 3 collect the rolling stock and the layout into two class diagrams. Trains and tracks are specified according to their functions in the yard. Figures 7 and 8 represent Activity diagrams formalising yard operations. The former considers a full yard and the latter specifies the composite activity related to the creation of an outbound train. Based on these diagrams, we start describing the rolling stock and layout classes in Sections 2.1 and 2.2. Then, we focus on operations in Section 2.3.

2.1 | Rolling stock

The main classes related to rolling stock are depicted in Figure 2. Both inbound and outbound classes are specifications of the generic Train class, as indicated by the generalisation paths (empty-head arrow). A train may aggregate one or more locomotives and one or more cars and this is shown through aggregation connectors, i.e. lines terminating with an empty diamond. On each of them, we indicate the cardinality windows to show the number of elements of each class which needs to be included in a train. Similarly to the train class, the cut class aggregates cars and possibly a shunting locomotive.

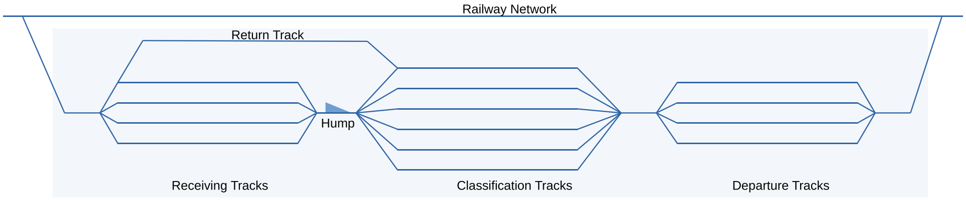

FIGURE 1 Schematic representation of a typical yard

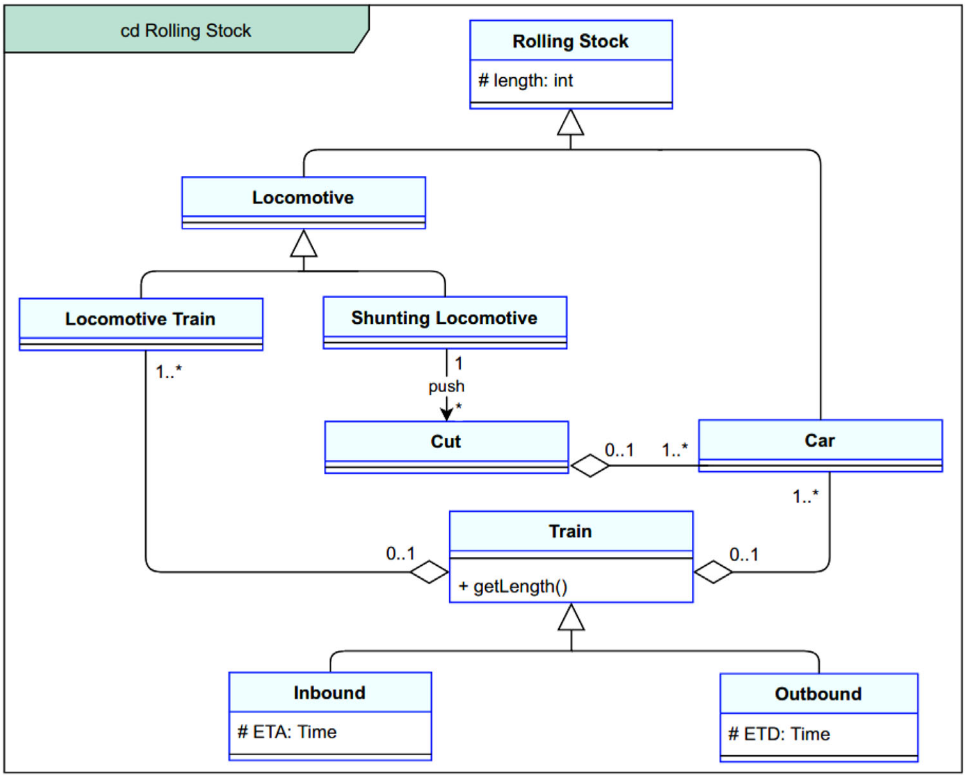

FIGURE 2 UML class diagram: rolling stock

2.2 | Layout

Track is the main class of the layout class diagram shown in Figure 3. It specifies into the receiving, hump, classification, and departure classes. Composite association paths, i.e. solid diamonds connectors, link the yard class and the specified tracks. They define the composition of a yard: a yard is composed by zero or more elements of each specified track class, apart from classification tracks which are necessarily existing. Moreover, the diagram points out how the specified tracks are connected to each other. The simple arrows, also called "association arrows," are used to represent these connections. Specifically, if all the classes of tracks are present, the receiving tracks are connected to the hump, which in turn is connected to classification tracks. Classification tracks are connected to departure and receiving tracks through the return track. The simple tracks through which these connections are ensured also participate in the composition of the yard. Every class but the yard has one or more attributes whose names are preceded by a # character. Receiving tracks are characterised by cuts and inbound trains which can occupy them, while hump and classification tracks are characterised by cars. Outbound trains are created on classification tracks and then occupy departure tracks. Finally, methods are defined for each specified track class and their names are preceded by the character +. They represent the operations that take place on such tracks and will be discussed in Section 2.3.

Following the depicted schema, the layout can differ from one yard to another. Its particular characteristics can make the system and the management processes more or less complex. We partition these characteristics depending on the four specified track classes: receiving tracks, hump, classification tracks, and departure tracks. For each of these classes, we list the main aspects impacting complexity as follows:

- Receiving tracks: In yards with long receiving tracks, several inbound trains can be sent to one track in order to form a cut. Once all the locomotives have been detached and moved to specific tracks along the receiving tracks, the full cut is pushed

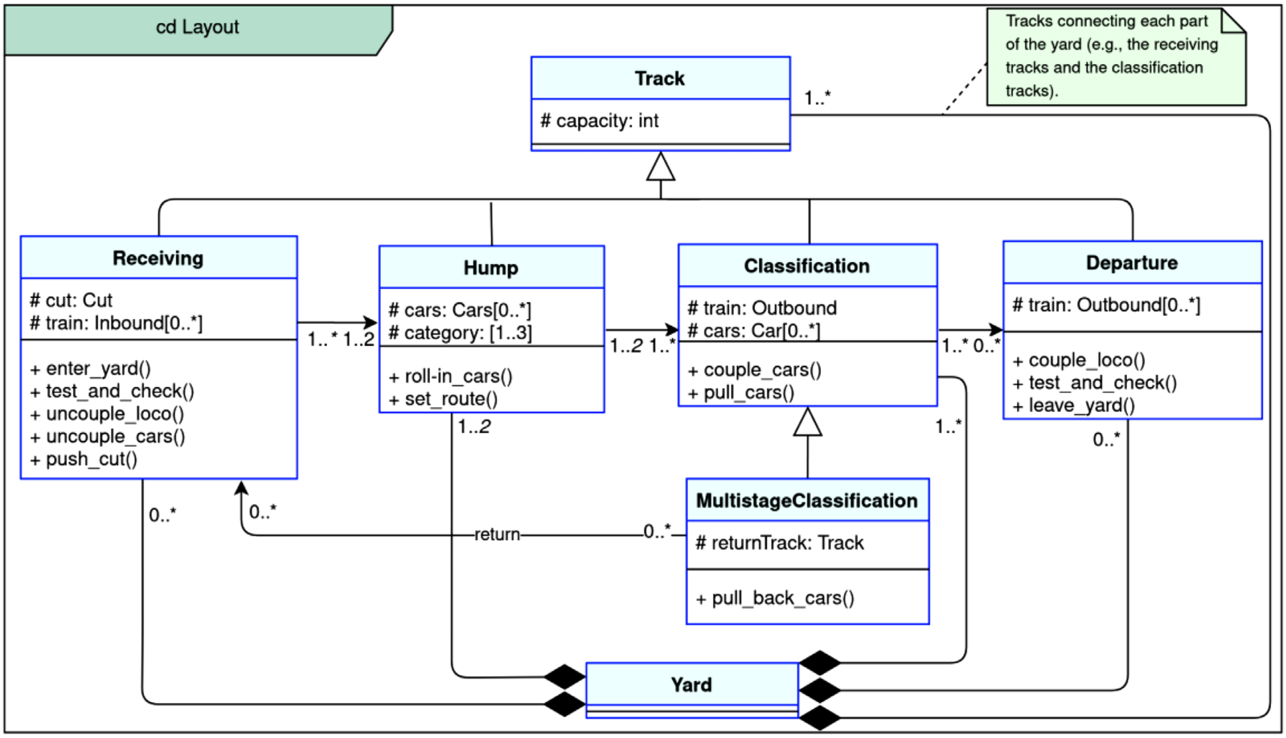

FIGURE 3 UML class diagram: layout

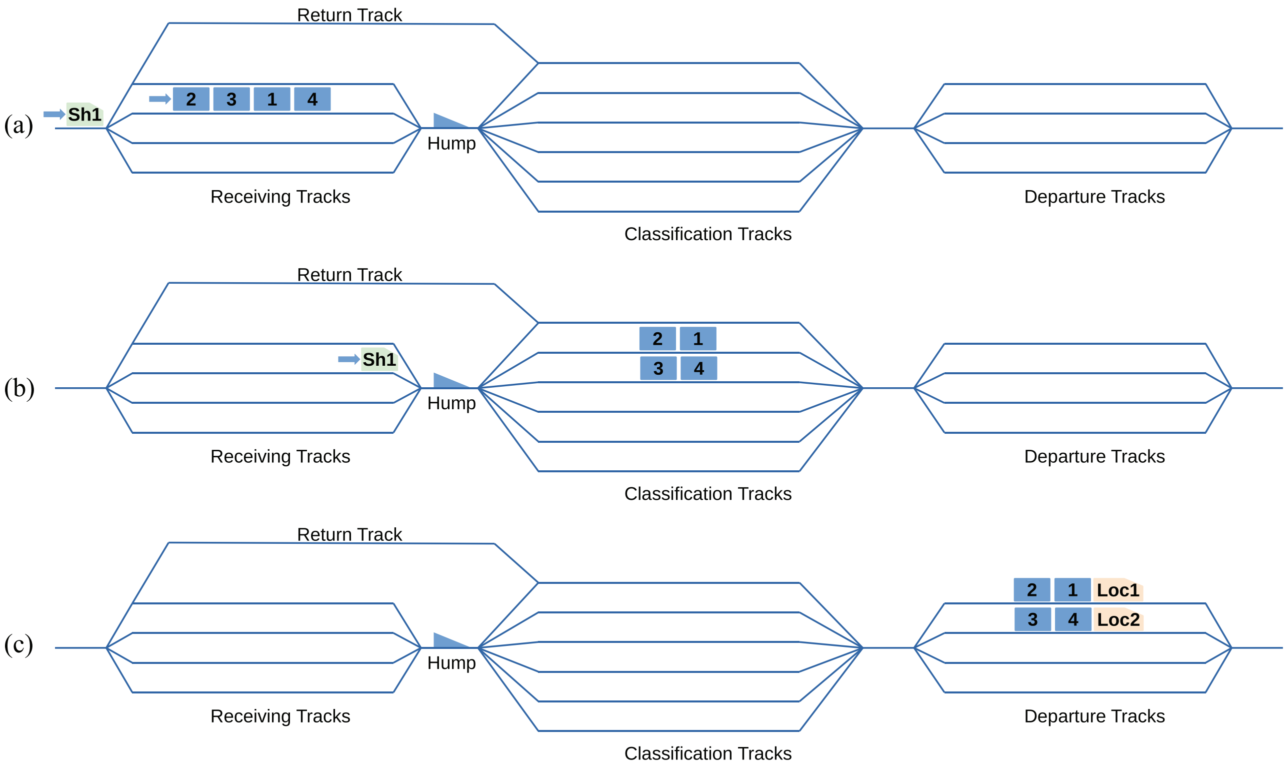

FIGURE 4 Example: single-stage classification

over the hump. In this case, the combination of the different series of cars on receiving tracks can be strategic for the effectiveness of the yard management.

Moreover, some yards do not possess receiving tracks. Hence, in Figure 3, for such tracks, we indicate the cardinality window [0..*] starting at 0 on the composite connector linking receiving and yard classes. If there are no receiving tracks, inbound trains are stored directly on classification tracks.

Finally, the cardinality windows ([1..*] and [1..2]) of the simple arrow connecting receiving and hump classes in Figure 3 show that each receiving track is connected to at least one of the two humps for the case of a 2-hump yard.

- Hump track: There is not necessarily only one hump track connecting the receiving and the classification tracks, even if this is the most common case. In Figure 3, we specify a number of humps equal to 1 or 2.

  Depending on the presence and the characteristics of hump tracks, there are three categories of yard:
  1. flat-shunted yard,
  2. gravity yard,
  3. hump yard.
- The flat-shunted yard has neither hump nor hill: the shunting locomotive pushes the cars up to the classification tracks. In this case, more energy is in general consumed by locomotives than in the case of classic humps, and this can have an impact on the strategy employed by operators. The gravity yard does not have any hump as well but trains are going downhill from receiving to departure tracks. It is considered as the most efficient process and is very often used on large systems [9, 10].

  In the general case, the hump is connected to all classification tracks or a subset of them (cf. cardinality window [1..*] on the simple arrow connecting the two classes).
- Classification tracks: The length of classification tracks can have a remarkable impact on the shunting process. For example, if these tracks are too short to build a full outbound train, the train is divided into sub-trains on several tracks, which are then merged on a longer departure track.

  The number of classification tracks is also important. If there are fewer tracks than the number of outbound trains to build from a set of inbound trains, the shunting process must be adapted. In this case, some classification tracks called mixing tracks can be used to store cars while waiting for a typical classification track on which the outbound train can be assembled. The typical classification tracks are then called formation tracks, and are used to build outbound trains. We will go back to the description of this process in Section 2.3.

  In Figure 3, we specify if there is a return track with a simple arrow connecting multistage classification and receiving classes and with cardinality [0..*]. Moreover, the diagram shows that classification tracks are connected to departure tracks, if they exist, through the cardinality windows [1..*] and [0..*] on the simple arrow between the two classes.
- Departure tracks: Some yards do not have departure tracks. It is represented in Figure 3 with the cardinality window [0..*] on the composite association path of the departure tracks. Similar to the case in which no receiving tracks are present, in case of no departure tracks, outbound trains which are ready to leave the yard stay on a classification track until their departure. In such a case, the number of available classification tracks is more critical.

In addition to the presence and characteristics of the specified tracks, other peculiarities can differentiate yard layouts. In some yards, trains enter and leave from the same side (they are called "yards with one end"). Moreover, tracks can have several functions concurrently (e.g. receiving and classification tracks). This case appears generally in yards which deal with a small number of trains since capacity is significantly reduced. Finally, in some cases there are specific tracks, like transit tracks used to bypass the classification area, or tracks used to store the locomotives decoupled from inbound trains. If present, the latter typically start between the receiving tracks and the hump and end between the classification and the departure tracks. A storage area can then be accessible in order to store locomotives. In this area, locomotives can wait the outbound trains they will be attached to.

2.3 | Operations

Operations tend to differ from yard to yard. A main difference is linked, first of all, to the number of classification stages. A stage is a sequence of classification operations which concretise in a series of cars being moved once from a receiving to a classification track. If this movement does not bring the cars in a suitable configuration for building an outbound train, for example if they are stored in a mixing track as mentioned in Section 2.2, a second classification stage is necessary, i.e. the cars need to be brought back to the receiving tracks. To realise this operation, which is called a pullback, a shunting locomotive pulls the cars through the return track.

Depending on the number of stages which characterise operations, we can classify the types of classification of a yard into three categories:

- single-stage classification,
- multistage classification with mixing tracks,
- multistage classification with car ordering.

In single-stage classification yards, cars are moved only once from receiving to classification tracks. Once they reach the latter, they start composing the outbound train to which they should belong. Figure 4 shows an example of operations in a single-stage classification yard.

Here, an inbound train is stored on a receiving track. There are two planned outbound trains which have to be built from the inbound train with composition (4,1,3,2). Two pairs of cars share the same destination: (1,2) and (4,3). No order is required for the cars of each outbound train.

In Figure 4(A), the inbound train becomes a cut when its locomotive and cars have been uncoupled. The cut is pushed by the shunting locomotive (Sh1) over the hump and the cars go downhill to the selected classification track for each car. Then,

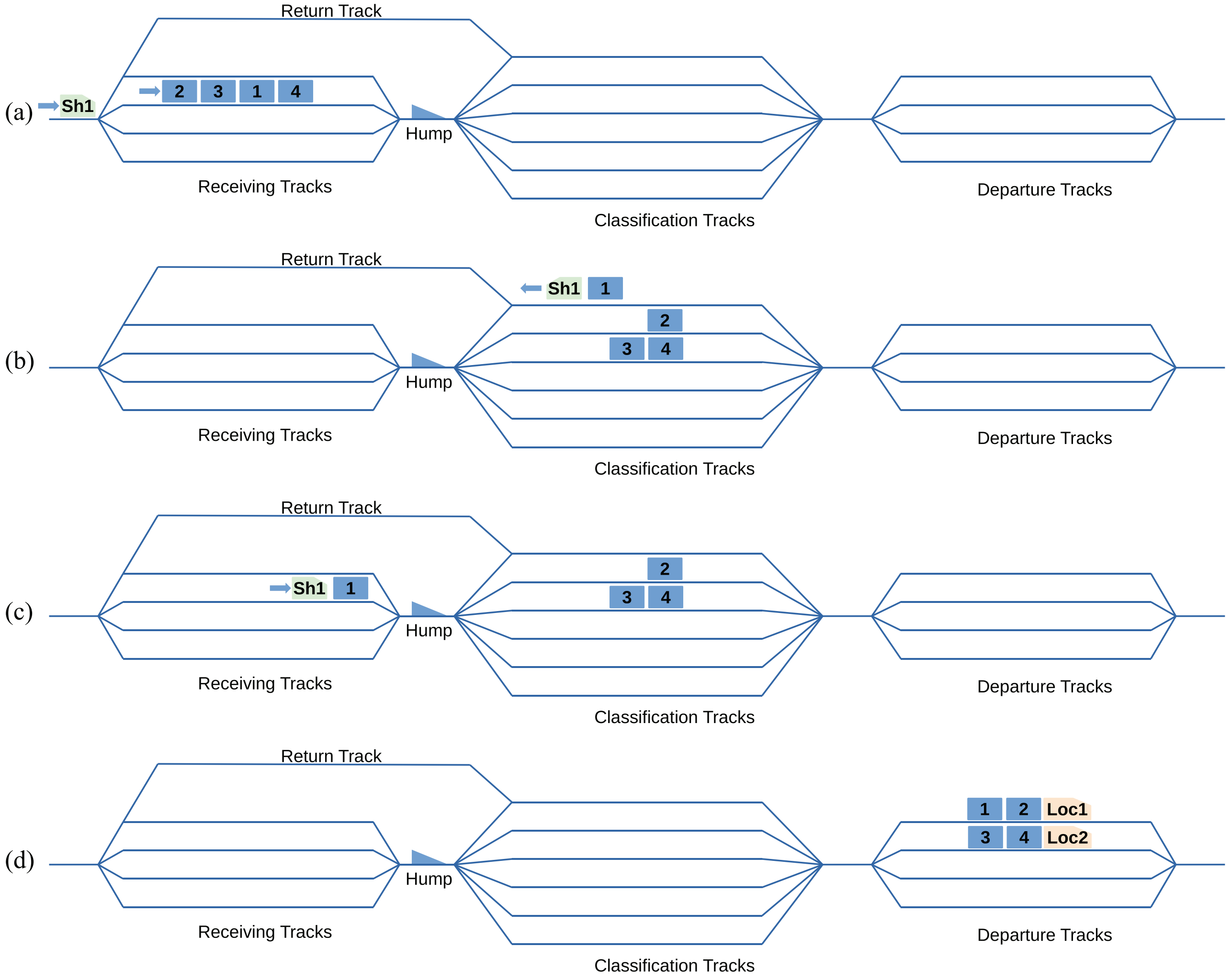

FIGURE 5 Example: multi-stage classification with cars ordering

in Figure 4(B), (1,2) and (4,3) are coupled. Finally, in Figure 4(C), once the outbound trains have been pulled to be stored on the departure tracks and attached to their respective locomotive (Loc1 and Loc2), they wait the right moment to leave the yard.

In multistage classification with mixing tracks, there are not enough formation tracks to start building all outbound trains. For trains which must wait an available formation track, cars are stored in one or more mixing tracks. These cars will need to be pulled back to the receiving tracks at least once.

Finally, in multistage classification with car ordering, cars can be immediately used for composing the outbound train they are aimed to, but they must be placed in a specific order. Hence, it may be necessary to pull them back to the receiving tracks once or several times.

In Figure 5, we propose an example of a multistage classification with car ordering. As in the single-stage example, there are two expected outbound trains and two pairs of cars share the same destination with a specific order: (2,1) and (4,3).

In Figure 5(A) the cut is pushed to the hump, but differently from the single-stage case, car 1 is routed to a different classification track than car 2 since they would not be in the correct order. In Figure 5(B), car 1 is pulled back through the return track to be re-rolled-in in Figure 5(C), to end up on the same track as car 2. Then, the cars of both outbound trains are in the correct order: (2,1) and (4,3). Finally, the trains are pulled to the departure tracks, waiting to leave once their locomotive is attached as we can see in Figure 5(D).

Figure 6 shows the different actions of pushing, pulling, and exploiting gravity force which are used in a yard with a multistage classification. We consider here the general case of a hump yard.

Disregarding the number of stages characterising operations, the main activities carried out in a yard are represented in the UML activity diagram of Figure 7. In such diagrams, an action represents a discrete unit of functionality in an activity. We articulate them depending on their object: inbound train, cut and cars, and outbound train.

The activity diagram starts with the opening of the yard, when the (e.g. daily) operations start according to a predetermined schedule, and ends with its closure, when all

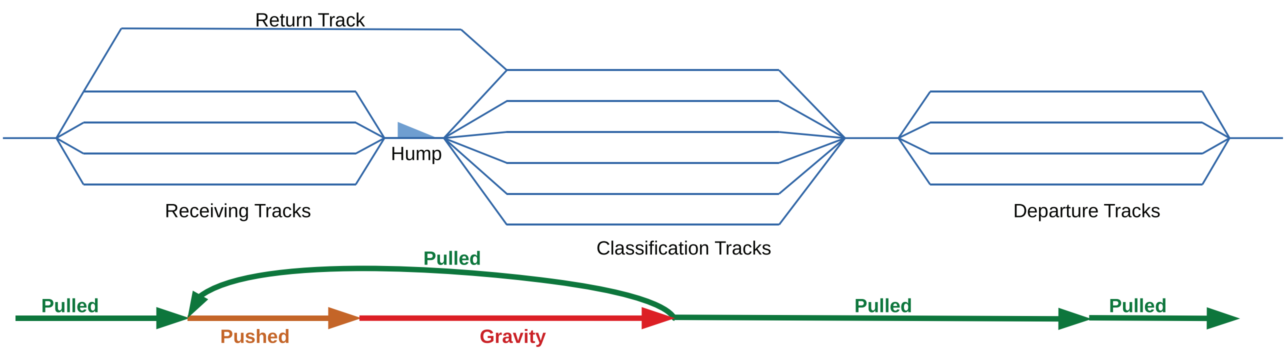

FIGURE 6 Schematic representation of the actions/forces moving cars in the different part of the layout

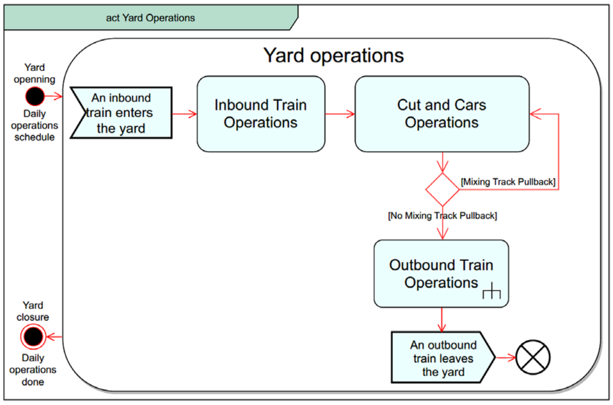

FIGURE 7 UML activity diagram (act) for the operations in the yard (the composite activity concerning the outbound train operations is given in Figure 8)

operations are completed. We identify three main groups of operations:

- Inbound train operations: The signal action an inbound train enters the yard introduces an inbound train in the system, which is handled by the inbound train operations action. This starts with the selection of the receiving track where the train is stored. Then, a specific crew working on the yard inspects and uncouples the locomotive and the cars going to different destinations. In the receiving class of Figure 3, these operations are represented by the methods *test_and_check*, *uncouple_loco*(), and *uncouple_cars*(), respectively.
- Cut and cars operations: A cut is considered by the system once the cars and the train locomotive are detached (cf. Cut and cars operations action in Figure 7). A shunting locomotive is then brought behind the cut (i.e. on the left side of the receiving track in the typical representation of a yard in Figure 1) to push it. The cut is pushed over the hump and the cars roll down to the classification tracks according to the route setting sequence defined. This operation is called roll-in. This sequence is typically automatic, i.e. the system recognises the cars and sets automatically the right route sequence to the selected classification track. In Figure 3, this corresponds to the methods *push_cut*() in the receiving class, *roll_down_cars*(), and *set_route*() in the hump class. In a multistage classification with mixing tracks, if some cars are not classified at the current stage, they are routed on the mixing tracks and pulled back to a receiving track at the end of the stage (method *pull_back_cars*() in the classification class). For the cars which have been pulled back, the action then starts again for a new stage.
- Outbound train operations: The system considers an outbound train being built once a classification track is chosen and the first car has reached it. The outbound train operations action is then executed. This action is signalled with a trident or "rake" symbol in Figure 7, which means that it is represented by the sub-activity diagram in Figure 8. The first operations represent the classification process while cars are accumulated on the classification track. They can be pulled back (method *pull_back_cars*() in the classification class in Figure 3), once or more than once, from the classification to the receiving tracks if we are in a case of multistage classification with car ordering. Once classification is done, cars are coupled (method *couple_cars*() from the classification class). Then, either a shunting locomotive or the outbound train locomotive pulls the outbound train to the selected departure track (method *pull_cars*()). If a shunting locomotive is used, the outbound train locomotive will be attached to the train on the departure track. After some checks (method *test_and_check*() in the departure class), the outbound train waits until its departure time to leave the yard. In Figure 7, the signal action an outbound train is leaving the yard terminates the operations of the current outbound train.

Remark that we only mention operations directly touching trains and cars shunting in the activity diagrams. Indeed, yard crews may have to perform other operations, such as cars registration and maintenance. However, we do not explicitly consider them in this paper since they can be included in the test and check activities if necessary. In the same way, we do not focus on the movement of the shunting locomotive. We made this choice to simplify the diagrams, and it is not restrictive for our literature review since no existing approach deals with these specific operations and movements.

The yard operations described in the activity diagrams above are the ingredients used by the optimisation problems that we survey in the next sections. Therefore, they can illustrate the interactions between the different parts of a yard management problem and consequently the possible articulations between

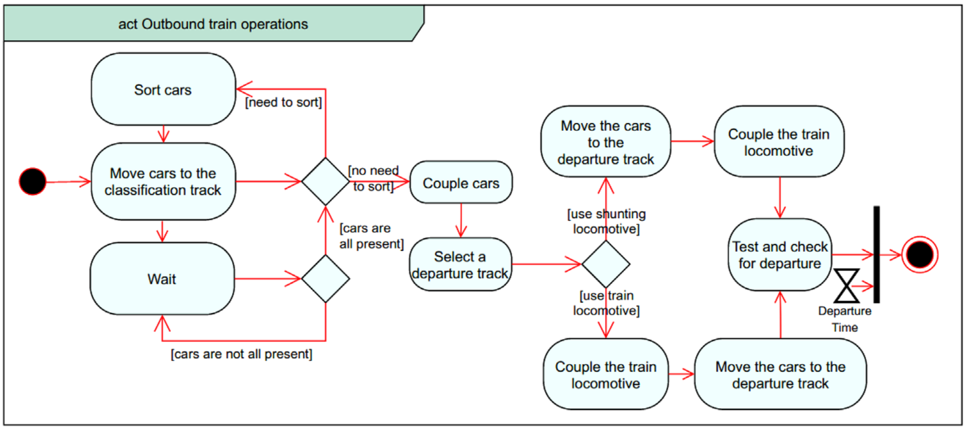

FIGURE 8 UML activity diagrams for the composite activity concerning outbound train operations

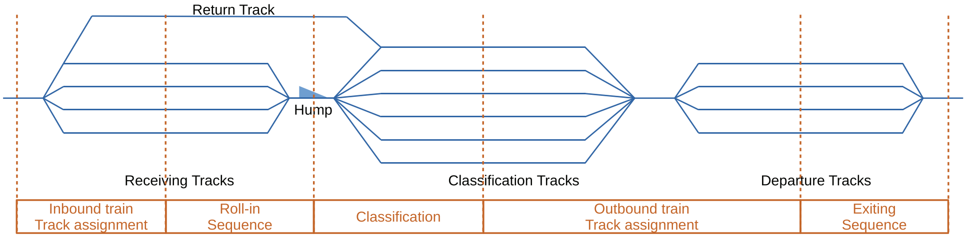

FIGURE 9 Optimisation problems emerging from yard operations

the different problems tackled in the literature. Any additional action to be performed can theoretically be formalised in very similar way using UML diagrams in order to make the yard description more precise, when additional aspects are added to a yard management problem.

3 | PROBLEMS TACKLED IN THE LITERATURE

Decisions related to the operations described in Section 2.3 can be formalised into optimisation problems, which are schematised in Figure 9.

The inbound train track assignment problem consists in deciding on which receiving tracks inbound trains are stored. This may be particularly relevant if receiving tracks are long enough to host more than one train. Another similar situation is when departure tracks are very long and several trains can be placed on the same track: this gives rise to the outbound train track assignment problem. In both problems, different solutions may have different impacts on the efficiency of the yard. In some cases, the network infrastructure manager (IM) decides these assignments. The roll-in sequence problem considers the order of cuts pushed over the hump and the classification problem corresponds to the shunting process, solved through single-stage, or multistage classification.

Presently, the literature almost exclusively focuses on the classification problem, as we will detail in Section 4. Some works consider that managing the arrival of trains in receiving tracks or the departures of new trains from the departure tracks is not relevant.

Whichever problem is considered, however, input and output definitions are quite constant. Specifically, most problems consider (at least) the following input:

- Estimated time of arrival (ETA) of inbound trains,
- estimated time of departure (ETD) of outbound trains,
- composition (i.e. sequence of cars) of inbound and outbound trains,
- layout and rolling stock characteristics (e.g. number of tracks, track lengths, type of cars, car lengths etc.),
- duration of each shunting operations.

The solution returned typically includes (at least), the following output:

- Schedule of the operations (time to leave the receiving track, time to enter the classification track, time to operate a pullback etc.),
- schedule of locomotives,
- planned departure time of all outbound trains.

The current schedule of operations and locomotives can also be part of the input in case the optimisation problem is not starting from scratch but is proposing modifications to an existing plan, for example to respond to a perturbation.

For optimising yard operations, the models use a given input (defining the problem instance) to obtain an output (feasible solution of the instance) which respects a set of constraints. The goal is to find the best output (optimal solution) following the direction of the optimisation which is given by an objective function, to be maximised or minimised. The most commonly employed objective function, for both single and multistage classifications, is the minimisation of the number of classification tracks used. For the multistage case, minimising the number of pullbacks is also frequent. Another typical objective function is the minimisation of delays of outbound trains. According to the type of yards and to the models evolution in the literature over time, several other objective functions have been considered. For example, some papers focus on the minimisation of locomotive energy consumption. There are also multi-criteria objective functions characterising some problems, where criteria are often concurrent (e.g. maximising the efficiency of the system while reducing the use of resources).

Even if yard management problems can often be seen as well known optimisation problems, their complexity was not deeply studied in the literature until recently. Specifically, refs. [11], [12], and [13] propose an overview on some problems complexity, and, more recently, ref. [8] propose an elaborate study of complexities for the general single-stage and multistage classifications. All these problems are sorted by complexity from polynomial to NP-Hard.

4 | ROLL-IN SEQUENCE AND CLASSIFICATION PROBLEMS

In this section, we report a review of the state of the art on classification problems. We start with single-stage classification, and we continue with multistage, with mixing tracks and car ordering. The focus of classification problems is very much on the humping sequence of inbound trains and the distribution of the cars on the classification tracks, with a more or less detailed accounting of the shunting and pulling locomotive schedules. However, a few works try to extend the analysis to other aspects of the yard, such as the schedule of car inspection or the allocation of receiving and departure tracks. We will discuss these works in the next section. The works presented in Sections 4 and 5 are summed up in Table 1, where we highlight the operational decisions on which each work concentrates, as well as the possible use of simulation techniques and the main algorithmic methods devised for tackling each problem. For the sake of generality and simplicity, we group the problems of assigning a classification track and an outbound train (train makeup problem) for each car under the label classification track allocation (see, e.g. ref. [14] for more details on the train makeup problem).

4.1 | Methods for single-stage classification

Following the definition of single-stage classification discussed in Section 2.3, the main optimisation problem tackled most often in the literature consists in selecting the sequence of trains and cuts to be sent over the hump, and to determine the classification track for each car. These operations are those referred to as cut and cars operations in Section 2.3.

An early contribution for this problem is the one by ref. [15], who propose an optimisation-simulation framework called HSS (hump sequencing system) to build the sequences of cuts ready to be rolled-in. They consider a system where inbound trains can enter the yard with some delay, and their cars are to be coupled to a later outbound train. HSS optimises the average yard throughput costs, represented by the cars' idle time. The main method behind HSS is based on dynamic programming (DP) which finds good solutions but requires quite a long computation time (a computer from the early 80's required 1 day of calculation for an instance with 20 trains). If the number of trains is too high, some trains to be rolled-in will be filtered (delayed) by a screening procedure according to a specific priority factor provided by the authors. The system simulates the state of the yard after each operation and can modify the solution dynamically or modify some parameters such as the arrival date of inbound trains.

Ref. [16] propose an event-based model for the same problem. The model runs on a rolling horizon basis: at the occurrence of each event, the system is modified and optimised based on the new status. A similar problem, although applied to passenger trains, is the one treated by ref. [11]. The authors optimise for two objectives: first, the minimisation of the number of tracks used and, second, the delays of trains. They study a case where the composition of outbound trains is modifiable. The model considers possible late inbound trains whose cars originally expected in an outbound train can be delayed and be part of the next one with the same destination. Delays are weighted according to the priority of cars. To find the best assignment of trains to classification tracks the problem is transformed in the so called chromatic number problem, which consists in finding the smallest number of colours needed to colour the vertices of a graph such that no edge has two vertices of the same colour. Here, the authors consider a permutation graph.

Concurrently, refs. [17, 18] focus on another variant of the problem, where multiple outbound trains have the same destination and one needs to decide to which outbound train the cars must be assigned. An optimal solution of the problem minimises the weighted tardiness of all outbound trains. The authors introduce the single stage sequencing problem with weighted tardiness (SSSWT), where the decisions concern the roll-in sequence of inbound trains. The problem, which is proven to be NP-hard, is modelled mathematically as an integer program (IP). Two heuristics are also proposed, in addition to two branch-and-bound procedures tackling the IP. The best of these heuristics is a tabu Search (for more details on the tabu search, we refer the interested reader to ref. [19]).

Inspired by the shunting problem emerging in Chinese yards, [20] deal with what they call the train marshalling problem (TMP). To optimise the cut and cars operations, the authors aim to minimise the number of tracks necessary to build

TABLE 1 Decision processes that each single-stage paper described in Sections 4 and 5 tackles, along with the main algorithmic approaches of each work. (B&B) Branch-and-bound; (B&P) branch-and-price; (DCO) distributed constraint optimisation; (DP) dynamic programming; (GA) genetic algorithm; (Heur) heuristic; (MILP) mixed integer linear program; (MINLP) mixed integer non linear program; (TS) tabu search; (VNS) variable neighborhood search; (ALNS) adaptive large neighborhood search

	Inspector scheduling	**Locomotive scheduling**	**Humping sequence**	**Receiving track allocation**	**Classification track allocation**	**Departure track allocation**	**Simulation**	**Algorithms and/or Results**
Single-stage								
[15]	–	–	✓	–	✓	–	✓	(DP)
[20]	–	–	–	–	✓	–	–	NP-completeness proof and bound
[47]	✓	✓	✓	–	✓	–	–	(GA) + *Fuzzy* (MILP)
[49]	–	✓	✓	✓	✓	✓	–	(MILP)
[11]	–	✓	–	–	✓	–	–	(Heur)
[17]	–	–	✓	–	–	–	–	(MILP) + (B&B) + (TS)
[18]	–	–	✓	–	–	–	–	(MILP) + (Heur)
[46]	✓	✓	✓	–	–	–	–	(MILP)
[16]	–	–	✓	–	–	–	✓	Event-based model
[28]	–	✓	✓	–	✓	–	–	MILP with valid inequalities
[23]	–	–	✓	–	–	–	–	(DP) + (DP)-based heuristic
[24]	–	✓	✓	–	✓	–	–	(MILP) + (Heur)
[26]	–	✓	✓	–	✓	–	✓	Rule-based framework
[21]	–	–	–	–	✓	–	–	(DP)
[27]	–	✓	✓	–	✓	–	–	High-level (MILP) + low-level (Heur)
[22]	–	–	–	–	✓	–	–	(DP)
[14]	✓	✓	✓	✓	✓	✓	–	(MILP) + (ALNS)
Multistage classification w/ mixing tracks								
[29–31]	–	–	✓	–	✓	–	–	(MINLP)
[3]	–	–	–	–	✓	–	–	(MILP) + (B&P)
[7]	–	✓	✓	✓	✓	✓	–	(MILP)
Multistage classification w/ car ordering								
[34]	–	–	–	–	✓	–	–	Approximate formulae
[35]	–	–	–	–	✓	–	–	By-train and triangular sorting
[44]	–	–	–	–	✓	–	–	Robust Algorithms
[37]	–	✓	✓	–	✓	–	✓	(MILP)
[42]	–	–	–	–	✓	–	–	Robust Algorithms
[13]	–	–	–	–	✓	–	–	(MILP) + Bounds + Greedy
[39]	–	–	–	–	✓	–	–	(DCO)
[41]	–	–	–	–	✓	–	–	(MILP) + (VNS)
[40]	–	–	–	–	✓	–	–	(MILP) + (TS)

the outbound trains. All cars are sorted by defining a partition, in which each set includes the cars assigned to a track for a future outbound train. A DP algorithm based on the inclusion–exclusion principle, which solves a graph theoretical model of the decision version of the TMP, is proposed in ref. [21]. The algorithm has a computational time complexity of $O(nkt^2 2^t)$, where n is the number of trains to sort out, k the number of available of auxiliary tracks and t the number of destinations, which demonstrates that the problem is fixed-parameter tractable when the number of destinations is fixed. An improved DP algorithm is proposed in ref. [22] that directly solves the optimisation problem instead of its decision version, with a worst-case time complexity of $O(nt2^t)$. This is achieved by grouping together subinstances that have the same optimal solution on the one hand, and by using the memorisation technique on the other, i.e. solving only the subinstances whose optimal solutions are needed to solve the original instance.

Focusing on another modelling of single-stage classification, ref. [23] study the assignment of each car from an inbound train to an outbound train. The authors assume that the inbound trains are pre-allocated to receiving tracks. The problem does not detail the sequence of cars over the hump nor the working time and the routing of the shunting locomotive. The authors propose an exact and a heuristic method, both based on DP and both considering priority values that depend on the urgency and importance of the cars loads. For example, they handle empty cars with a very low priority. The time needed to optimally solve the problem grows exponentially with the size of the instances, even with the use of the sophisticated upper bound proposed in the paper. The heuristic is based on the DP algorithm but it does not explore the whole set of dynamic states. It returns feasible solutions in a short time for small and medium sized instances.

Ref. [24] tackle the hump yard block-to-track assignment (HYBA) problem. This problem was proposed as a challenge in Railway Applications Section (RAS) of the Institute for Operations Research and the Management Sciences in 2014, see ref. [25]. Given a planning horizon, the aim is to determine the schedule (time to be rolled-in) and the routing of cars. The objective function is the minimisation of the outbound trains delays. The authors split the problem into three subproblems: the hump sequence problem (HSP), the block to track assignment problem (BTAP) and the pullout allocation problem (PAP), where a pullout operation consists of pulling a series of cars from the classification to the departure tracks again using a specific locomotive. The HSP consists in deciding the sequence according to which the cars are to be pushed to the hump (one can refer to the roll-in sequence problem of Figure 9). The quality of a sequence depends on the departure day of the final car to be processed. The BTAP seeks the best assignment of blocks to classification tracks, where a block represents the set of cars with the same destination. Finally the PAP decides which pullouts to perform from the classification tracks to the departure track. This problem concerns the outbound train operations as discussed in Section 2.3 and the move the cars to the departure track activity. Two of these problems, the HSP and the PAP, are solved by IP while the BTAP is solved by a heuristic. The overall algorithm starts by solving the IP of the HSP. This can be done once since the roll-in sequence remains fixed for the whole process. Then the authors use a greedy algorithm to solve the BTAP and obtain the schedule of the cars from the hump to the classification tracks. The IP of the PAP is solved once all the cuts went down the hump or the classification tracks are full. Ref. [26] tackle the same problem using a simulation approach based on a rule-based framework. Another approach is advocated by ref. [27], who first solve a rolling horizon mixed integer linear program (MILP) representing a high-level description of the problem, where some microscopic objects like tracks or cars are aggregated into larger objects. Using the solution obtained from such an aggregated model, a microscopic detailed solution is obtained using a heuristic algorithm, which is competitive with the two previous works.

Finally, ref. [28] propose an IP model for the railroad yard operations plan problem. The authors consider a rather high level of detail to describe complex yard operations. For instance, they take into account the different types of locomotives (roll-in and pullback). The mathematical model is based on a flow model and it is solved with the help of lot-sizing problems valid inequalities.

The algorithms discussed in this section focus on a unique stage but they can be used as a sub-routine to compute a multistage classification schedule as underlined by ref. [8].

4.2 | Methods for the multistage classification with mixing tracks

As for the single-stage classification, the problems considered with multistage classification with mixing tracks belong to the cut and cars operations described in Section 2.3. Here again, the typical objective of the problem is minimising outbound train delays, with the additional complexity added by the small number of tracks available for building them.

Ref. [29] propose a non-linear IP model for the cars classification problem. These works have been labelled as single-stage methods in some previous review papers. However, we think they more suitably fit in the multistage methods since they consider at least one pullback from the classification tracks to the hump. Indeed, the operations studied include three pullbacks per day (one every 8 h). The main inputs of the model are the arrival times of inbound trains, the expected departure times of outbound trains and a feasible sequencing plan for cars. The output is the solution of the IP which returns a roll-in sequence. The model minimises an exponential function based on the delays of outbound trains. As in ref. [15], discussed in Section 4.1, a distinctive feature of this model is the possibility for a car to be delayed and inserted in a later outbound train with the same destination. In ref. [29], the author runs an IP solver for a limited amount of time. The process stops the branch-and-bound algorithm and returns the best found solution (which could be an unproven optimal solution). This technique is called a truncated branch-and-bound. This work is extended later in refs. [30, 31]. Ref. [30] focuses on high priority cars which need to make it on time to their assigned outbound train. These cars are sometimes stuck between low priority cars, which requires complex and costly operations to ensure that the high priority car can make its connection (this is called cherry picking). The method advocated is to split inbound trains into blocks and to map these blocks to a specific outbound train instead of a mere destination. Ref. [31] elaborates on this aspect by allowing to reassign outbound trains to cars before humping, possibly even an outbound train leaving earlier.

In the variant of the problem tackled by ref. [3], where the schedules of inbound and outbound trains are known, the classification tracks are reserved for building specific outbound trains. The classification tracks have a limited and non-homogeneous length. A mixing track is used to store cars and to wait for the reserved classification tracks to be available. The authors propose two IPs to tackle the problem with a single mixing track. The first one has an exponential number of variables and is solved with a branch-and-price method. The second one is a compact model, i.e. involving a polynomial number of variables and constraints in the problem size parameters. The latter obtains the best results. We report below the compact model since we consider it allows to represent quite well the state of the art on this problem without the burden of an extremely complex formalisation.

It models the problem in terms of pairs of trains scheduled in immediate succession on the same classification track. Each series of consecutive cars from the same inbound train which share the same outbound train is handled as a single car.

Let $\mathcal{A}$ and $\mathcal{B}$ be the set of classification tracks (the mixing track is not included) and the set of outbound trains, respectively. $\mathcal{B}_a$ gathers all the outbound trains allocated to track $a \in \mathcal{A}$. $p \in \mathcal{P}$ represents a stage among the set of stages $\mathcal{P}$. A stage includes all the movements performed between two pullbacks. $l_p(b, b')$ denotes the total length of these cars. $c(b, b')$ is the number of extra roll-ins. It corresponds to the number of cars going to the mixing tracks before being pulled back and then pushed again to the hump. u and v are virtual trains defining respectively the predecessor of the first train and the successor of the last train on any track. The only variables of the model are the threefold indexed binary variables noted $x_{bb'a}$, $b, b' \in \mathcal{B}_a$, $a \in \mathcal{A}$ and defined as $x_{bb'a} = 1$ if train b' is scheduled immediately after train b (represented as $b \prec b'$) on track a and 0 otherwise.

The IP of ref. [3] is:

$$\min \sum_{a\in\mathcal{A}} \sum_{b,b'\in\mathcal{B}_a : b\prec b'} c(b, b') x_{bb'a} \tag{1}$$

subject to:

$$\sum_{a\in\mathcal{A} : b'\in\mathcal{B}_a} \sum_{b\in\mathcal{B}_a : b\prec b'} x_{bb'a} \geq 1, \qquad b' \in \mathcal{B}, \tag{2}$$

$$\sum_{b'\in\mathcal{B}} x_{ub'a} \leq 1, \qquad a \in \mathcal{A}, \tag{3}$$

$$\sum_{a\in\mathcal{A}} \sum_{b,b'\in\mathcal{B}_a : b\prec b'} l_p(b, b') x_{bb'a} \leq k^{\text{mix}}, \qquad p \in \mathcal{P}, \tag{4}$$

$$\sum_{b\in\mathcal{B}_a : b\prec b'} x_{bb'a} = \sum_{b\in\mathcal{B}_a : b'\prec b} x_{b'ba}, \quad a \in \mathcal{A}, b' \in \mathcal{B}_a \backslash \{u, v\}, \tag{5}$$

$$x_{bb'a} \in \{0, 1\}, \qquad a \in \mathcal{A}, b, b' \in \mathcal{B}_a. \tag{6}$$

The objective function (1) minimises the number of extra roll-ins. Indeed, if $x_{bb'a} = 1$, b and b' are in consecutive order on track a and $c(b, b')$ gives exactly this number. Constraints (2) force the outbound trains to be in a sequence. If the train is the first one, it is virtually the second one right after u. Similarly, if it is the last one, it is virtually the one before v. At most one train can be the first one on each track, thanks to Constraints (3). Constraints (4) allow a series of cars on the mixing tracks at each stage only if its length is shorter than the one of the track k^{mix}. The last Constraints (5) are based on typical flow conservation constraints and force outbound trains to have the same number of predecessor and successor. Finally, the integrality Constraints (6) ensure binary values for the $x_{bb'a}$ variables. This model is shown to succeed in solving instances representing the largest hump yard in Scandinavia in less than 20 min, as will be discussed in Section 7.

4.3 | Methods for the multistage classification with car ordering

When a specific car ordering has to be achieved in outbound trains, the problem concerns the classify cars composite activity discussed in Section 2.3. To obtain this order, several stages may be necessary. Simple sorting algorithms are typically used, such as sorting-by-block, sorting-by-train, triangular sorting, or geometric sorting [6, 8 32–34].

Figure 10 represents the main steps of the sorting-by-block algorithm. The idea is to use one classification track per car of a given outbound train. Although the algorithm does not specify what to do with the cars aimed at other trains, they are usually stored on a mixing track. However, if the number of classification tracks is higher than or equal to the total number of cars, the classification can be single-stage as in Figure 10. Once the cars are stored in the classification tracks, they are pulled one by one to the chosen departure track in the correct order. This algorithm can be efficient when outbound trains are formed by a low number of cars and yards have many classification tracks. Indeed, the orders of cars are quite easy to obtain with this algorithm, although the number of tracks used is high.

The sorting-by-train algorithm is represented in Figure 11. The main idea is to dedicate each classification track to an outbound train. While the cuts are rolled-in, the cars are sent to the classification tracks according to their target outbound train. In the example, the operators have to build two trains, sorted as (4-5-6) and (1-2-3). Then, the sorting-by-train algorithm uses two classification tracks, one for each outbound train. After the first stage of roll-in, they obtain trains (4-5-6) and (1-3-2). In another stage, the series of car (3-2) is pulled back to send (3) to another classification track. Car (3) is then pulled back to be finally sent to the same track as (2-1), ending up on its left. Since the two trains have the right order, they are sent to the departure track.

The triangular sorting algorithm indexes the cars according to a sophisticated calculation based on the length of the trains. The obtained indexes lead to an assignment to the classification tracks. This algorithm uses fewer tracks than the two previous ones but is very costly in terms of pullbacks. An illustrated example is given in ref. [6]. The geometric sorting algorithm is an evolution of the triangular sorting algorithm and considers some sorting-by-train algorithm principles [35].

In general, the multistage classification with car ordering can be seen as a sorting problem with n stacks, introduced in ref. [36] and solved with the sorting stack procedure, with n being the number of tracks corresponding in the analogy to $n-1$ "classification stacks" and 1 "hump stack". Since this work, no paper was proposed with new algorithms for such a fundamental problem.

To present, analyse, and develop classification methods, ref. [13] propose a powerful encoding of classification schedules. This encoding can be used to analyse the efficiency of commonly used multistage methods, as shown in the paper for the simple methods introduced above. Through the encoding, the authors prove the optimality of a variant of the geometric sorting in terms of sorting steps, considering presorted input.

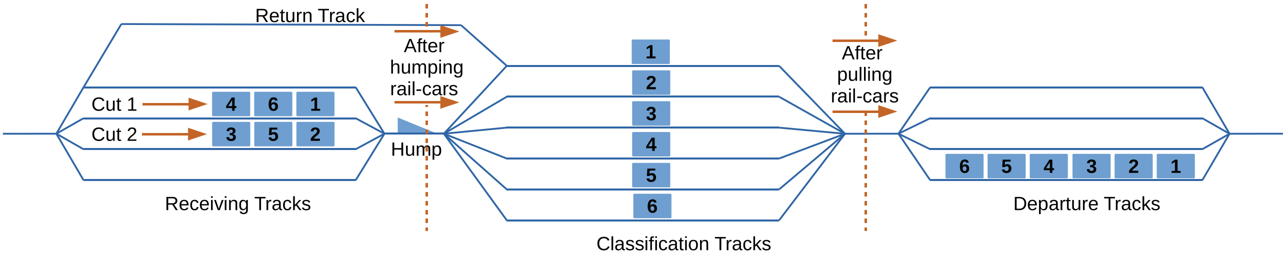

FIGURE 10 Three main steps of the sorting-by-block algorithm

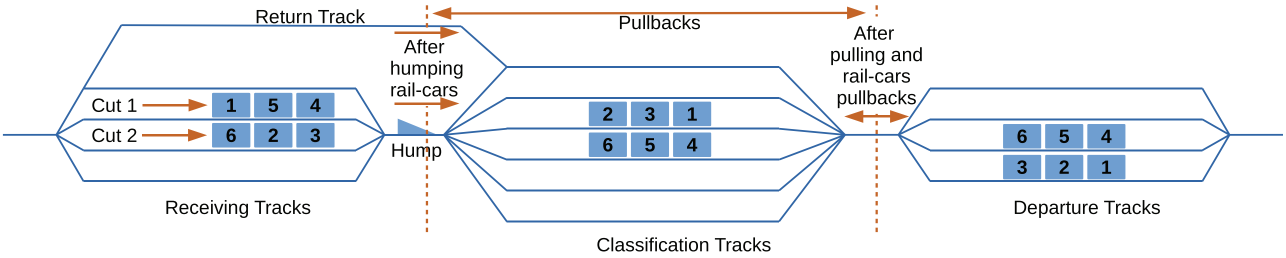

FIGURE 11 Three main steps of the sorting-by-train algorithm

Another IP model is presented by ref. [37], who minimise both the number of sorting steps and the number of cars rolled-in during the classification process. In their bi-objective algorithm, the authors solve a sequence of IP. Moreover, they consider the case of a yard with multiple humps. In this case, two classification activities, one per hump, can be run in parallel. In the modelling proposed, the activities deal with independent cars partitions: there is one shunting engine operating on each hump, and each available classification track is accessed from exactly one hump; furthermore, every outbound train is composed using only cars from exactly one partition. The assignment of trains to partitions is part of the optimisation process, and it is hence included in the IP model. An additional contribution of ref. [37] is the application of a microscopic yard simulator for assessing the performance of the algorithm. The simulator is named Villon and is developed by Simcon, in Slovak Republic [38]. Such an application allows the evaluation of the performance independently from simplifying modelling hypothesis which are necessarily present in optimisation algorithms. The same case study is used in ref. [39], which considers the order in which the inbound trains arrive in the yard. They propose an iterative procedure based on distributed constraint optimisation which obtains heuristic solutions.

Considering a quite specific utilisation of classification tracks, ref. [40] propose and IP model and a tabu search algorithm for the multistage problem with car ordering. Classification tracks are grouped into empty, clean and mixed tracks. On clean tracks, outbound trains are built. Mixed tracks are divided into temporary and dirty tracks: the former host cars ordered as in their aimed outbound trains; the latter are occupied by cars in scattered order.

Ref. [41] handle a multistage classification problem where both the number of sorting steps and the total number of movements are minimised, instead of handling those criteria lexicographically as usually done. Several trains are formed simultaneously and they consider the problem both at the tactical and strategic levels, i.e. the number and length of tracks are considered as variables. However, it is assumed that the inbound flow of cars is given and the number of cars does not vary over time. A MILP is devised to deal with the problem but given its intrinsic complexity, a reduced variable neighborhood search algorithm is also proposed. For instances with more than 200 cars, the metaheuristic algorithm is shown to return better results than a commercial solver running the MILP.

Ref. [42] propose algorithms for robust classification schedules while considering multiple inbound and outbound trains. They study a specific type of robustness called "recoverable robustness" (ref. [43]). Compared to strict robustness, recoverable robustness excludes cases that hardly ever occur in practice. The disruptions of their model only involve changes appearing in the sequence of inbound trains cars but not the potential unavailability of classification tracks. In addition to providing a feasible solution for the expected sequence of cars rolling over the hump, their algorithms build sufficiently flexible solutions so that new stages can be feasibly inserted in the schedule if the sequence is disturbed (e.g. a delayed inbound train). The algorithms have recovery parameters in order to trade off between a fast schedule and a robust one. Although their model considers many cases, the authors define classes of scenarios to experiment on those which are most probable, avoiding extreme and too rare cases. Ref. [44] discussed the notion of "recoverable robust solution" from ref. [45] in the context of classification problems. They consider three different cases: in the first, only the length of the classification tracks (with the number of cars which fit into one track) is bounded; in the second, only the number of classification tracks is bounded; in the third case, neither the length of classification tracks nor their number are bounded. Only one type of disruption is taken into account at the same time. The disruptions are the following: one car with unexpected position, one new car, one missing car (the simplest case, since nothing has to be changed in the initial schedule) and one unavailable track.

5 | EXTENDED YARD MANAGEMENT PROBLEMS

Though the vast majority of existing works focuses on the car humping sequence and classification track allocation, a few works exist which attempt at including additional parts of the yard operations and we describe them below.

For the single-stage classification, the work of ref. [46] considers the classification and train assembling sequences as well as the inspection sequences, as they argue that this may be an important bottleneck if the number of available inspectors is low. The assignment of trains to tracks (receiving, departure and classification) is ignored as the authors do not consider it a bottleneck in general. The number of available humping and yard locomotives is also limited. The aim is to minimise the total dwell time of the cars in the yard. The authors propose a MILP for this problem and solve it on different scenarios with different numbers of locomotives. This is one of the very few works which consider the inspection process and its impact on the scheduling of the overall yard operations. Another such work is ref. [47], who consider some uncertainty in their optimisation problem based on typical operations in the yard. From one scheduled sequence of inbound trains and one expected sequence of outbound trains, they optimise the operations schedule of car separation and combination. Operations related to the switch and the yard engines are considered. To tackle the problem, the authors first develop a MILP based on fuzzy set theory to handle imprecise information (ref. [48]). Second, they propose a metaheuristic which is a hybrid method of genetic algorithm and local search techniques. A complete solution of the model amounts to taking the following decisions: affect a scheduled outbound train for each car, authorise some potential delay for some operations and schedule the time to assemble the outbound trains. The goal is to assemble the maximum number of cars in outbound trains, with cars spending the minimum possible time in the yard.

Another example of a single-stage problem, which has similarities to the pullout allocation problem (see Section 4.1), is ref. [49], who formulate as an IP the problem of optimising the sequence to follow when assembling outbound trains. The objective is to find a good balance between the minimisation of the dwell time of each car and the minimisation of the outbound trains delays at departure. The time horizon of interest is split in intervals of several hours treated sequentially. In the paper these intervals are called stages, although they have nothing to do with pullbacks. In addition to the humping sequence and the classification track allocation, the model also includes the allocation of the receiving and departure tracks in the yard.

Finally, among the class of multistage problems, ref. [7] propose a comprehensive IP model for track allocation and roll-in operations timing. The authors consider a specific yard layout with mixing track, and no departure tracks. According to the activity diagrams depicted in Section 2.3, this work focuses on Inbound train operations, cut and cars operations, including mixing track pullback, and outbound train operations. The absence of departure tracks and of the need for car sorting implies that only the couple cars, couple the train locomotive. and test and check for departure activities are taken into account. The paper covers the first three problems of Figure 9. For inbound trains, the model takes the arrival track capacity into account. For the classification track allocation, it is a variant of the model by ref. [3], modulo a few differences due to the layout of the yard. Classification tracks are allocated as groups of tracks (clusters) with the same length. Shunting operation start times are variables of the model while arrival and departure times of inbound and outbound trains are given as input. The main part of the work focuses on scheduling pullbacks, differently from ref. [3] who look for the best sequence on each track without any time consideration. Scheduling constraints ensure that roll-ins and pullbacks do not overlap on the hump (in the considered yard layout, both roll-ins and pullbacks use the hump). Every possible combination of activities is taken into account and modelled into constraints. For example, constraints formulate how a pullback can fit between a roll-in and a departure, between a departure and a roll-in, or even between two departures. Many precedence constraints are also developed, e.g. if the classification tracks are all busy, one of them has to be freed before new cars can be rolled-in to create a new outbound train on it. In the end, all of these constraints have to be activated depending on the case to be dealt with. The paper optimises a multistage yard model with two objective functions. The first one minimises the work effort which is in fact the minimisation of the number of pullbacks from the classification tracks to the hump. It actually minimises the number of shunting operations like the number of couplings, decouplings, checks, and shunting locomotive movements. The second objective is the minimisation of the track costs, that is, of the weighted sum of the number of receiving tracks and the number of classification tracks used. The resulting model has many big-M constraints which, together with the overall large number of constraints, make the resolution of the model very complex. The results presented show how difficult it is to find the optimal solution to relevant instances. However, the model is shown to achieve interesting performance in several situations.

Finally, the latest contribution regarding integrated models and algorithms for single-stage shunting is the one of ref. [14]. Their problem considers the assignment of receiving, classification, and departure tracks, the outbound train selection for each car as well as the scheduling of all yard resources, including tracks, shunting, and pulling locomotives and yard inspectors. The problem is considered static and does not consider the routing of trains and locomotives on the yard. The authors propose a general MILP formulation. Since it is of limited use on real-world instances, they also devise an adapted large neighbourhood search heuristic with simulated annealing as an outer framework, which performs well against the MILP on instances which can be solved exactly. The sequence of actions to realise for each train is modelled using a directed graph.

In Table 1, we summarise the analysis of the state of the art underlying the problems dealt with by each paper and the techniques used to do so.

6 | OTHER YARD MANAGEMENT PROBLEMS

The literature also proposes algorithms for other, less commonly considered, yard management problems. In this section, we briefly review a few examples of such problems and the solution methods adopted. Among them, ref. [50] propose an IP model aiming at dynamic empty car assignment to outbound trains. The hypothesis is that all cars, full and empty, are planned to leave a yard on a specific outbound train. While for full cars this is considered a constraint, the assignment of empty ones is taken as modifiable, provided that the planned number of empty cars of each type (box cars, flat cars, gondolas etc.) departs on each outbound train. Indeed, in case of delayed arrival of inbound trains, the planned assignment may increase the idle-time empty cars spend in yards. The authors aim to minimise the empty car time in yard. with an algorithm based on a sliding time window principle, in which only empty cars in not yet rolled-in cuts can be re-assigned.

The variability in the trains arrival time with respect to the original timetable is taken into account in ref. [51] through a dynamic reoptimisation approach which reassigns trains to receiving tracks in a yard to minimise the total delay of trains. Shunting and routing operations, however, are not taken into account. The authors solve the problem with a MILP model and with two heuristic approaches, namely a genetic algorithm and a first-scheduled first-served heuristic for comparison purposes. Different approaches using event-based simulation (and handling the different yard resources, such as inspection teams) for a shunting yard management with uncertainty are proposed, e.g. in refs. [52] and [53].

An approach for dealing with routing in shunting yards is proposed in ref. [54]. It relies on a knowledge representation technique called answer set programming and exploits the knowledge of experts in yard management to find the best routes for trains inside a yard. The method is more efficient for very complex problems with complex sets of rules for deciding the routes.

An increasing amount of literature is dedicated to a different type of yard, where trains transport containers which are loaded and unloaded without the need for uncoupling the cars. Such yards are usually called trans-shipment yards and can feature rail-rail containers exchanges as well as intermodal exchanges, e.g. ship-rail or rail-truck trans-shipment. The management of such yards is a very complex task since the containers are transported by cranes which must be scheduled efficiently, and possibly by shuttle cars moving along sorters for long-distant container moves along the spread of the yard. Moreover, the containers sometimes need to be stored in a storage area when their destination vehicle is not available yet. A frequent objective in such problems is the total processing time or the makespan of the cranes, as they are linked to the processing time of the trains themselves or the cost of operating the yard. Some of the complex decisions to make on such yards are, e.g.:

1. Partitioning the trains into bundles on the parallel tracks, respecting arrival times and due dates as well as tracks capacity.
2. Assigning the trains to a parking position on the tracks.
3. Assigning a position to the containers on the trains or ships.
4. Assigning the containers to be moved to a single crane or a crane pair.
5. Sequencing the cranes operations.
6. Assigning the shuttle cars to split moves which use a pair of cranes.

An earlier example of optimisation approach for transshipment yards is the work of ref. [55], which deals with the problem of container loading in rapid rail-rail transshipment yards. Containers are typically first accumulated in specific sites before being moved to outbound trains. The problem considered in the paper consists in determining the initial loading site of containers and their reloading place on outbound trains to minimise their transfers within the yard and therefore the use of handling equipment. Consequences of this optimisation is the reduction of the train processing time in the yard and thus the total time for correspondence. The authors propose heuristics and an IP model for different variants of the problem.

Many works in the literature tackle the optimisation of gantry cranes on transshipment yards. While the different tasks of the yard operations used to be tackled individually, ref. [56] propose a MILP which handles the first three aspects of the yard listed above in an integrated manner. Ref. [57] tackle the last two problems, i.e. crane and shuttle car sequencing, as their schedules are intricately linked together. After showing the NP-hardness of the problem, they propose a heuristic decomposition where the subproblems obtained are solved by specific DP algorithms. Ref. [58] instead seek to minimise the number of containers that need to be stored in the storage area. They reformulate the problem as an acyclic partitioning problem on a directed graph and devise a MILP with valid inequalities, which are based on the derivation of several solution properties.

An example of a recent work on intermodal transshipment yards is the one of ref. [59], who focus on the optimal loading of containers, maximising the cars utilisation, on double stack trains. The problem is solved using a MILP which takes into account all the loading constraints of the trains. We refer the interested reader to ref. [2] for a specific survey on transshipment yards.

Finally, a rather recent trend considers the use of machine learning (ML) methods to find optimised solutions for several aspects of shunting yard operations and can help improve the input data for the problems discussed in previous sections. Ref. [60] introduces a reinforcement learning method to assemble several outbound trains simultaneously in a shunting yard. The yard layout is very simple and includes only sub-tracks where cars are temporarily parked and a main track to form outgoing trains, with cars in the right order. The aim is to minimise the total processing time considering only one shunting locomotive. Other ML methods are used, e.g. to improve estimates on trains delay, such as ref. [61] who use support vector regression to predict updated ETAs of incoming trains in real-time; or ref. [62] where the departure delay of outbound trains is estimated using tree-based methods and the synthetic minority oversampling technique. ML techniques can also be of use to improve the driving profiles of freight trains by estimating their change in mass over the shunting process [63] or to help regulate of the speed of cars that slide freely from the hump [64].

7 | CASE STUDIES

Although most of the papers cited in the above sections assess the proposed algorithms on academic instances, some works tackle case studies representing operations in real yards. To give the feeling of what the literature may be capable of dealing with in practice, we gather in this section the information available on these case studies.

München–Nord yard—Germany: Ref.[18] tackles instances based on operations in the München–Nord yard. They consider a 1-day time horizon, with up to 600 cars arriving. München–Nord is a medium-sized German shunting yard. It is a hub for Germany, Austria, Czech Republic, and Italy. This yard is highly automated and has one hump. The number of receiving tracks, classification tracks and departure tracks are respectively 14, 40, and 13. 356 switches are installed and the total length of tracks is 120 km. In 2017, 1379 inbound trains/month have been received and 1441 outbound trains/month have been built. There are around 500 transit trains each month, i.e. trains which cross the yard but do not need shunting. Finally, the hump capacity is about 250 cars per h. This capacity is large enough since only 110 cars per h are actually going over the hump (ref. [65]).

Hallsberg yard–Sweden: The Hallsberg yard, considered in ref. [3] and located in Sweden, is the largest yard in Scandinavia. Being right next to a passenger yard, it is a strategic point of the Swedish rail freight, especially in the north–south freight corridor. Ref. [3] generate their instances from data related to this yard. Two tracks are passing over the hump but only one can be used at the same time due to safety constraints. The yard has eight receiving tracks with lengths ranging from 595 to 693 m and 32 classification tracks with lengths ranging from 374 to 760 m. Among the classification tracks, two are typically used as mixing tracks and the 30 others as formation tracks. Therefore, multistage operations are executed in the yard. It also has 12 departure tracks with lengths ranging from 562 to 886 m. There are 170 switches and 60 km of total track length. In 2016, around 1161 inbound trains per month have been uncoupled to build other 1197 outbound trains per month. The yard has a capacity of around 167 cars per h while, today, an average of 102 cars per h is handled.

Sävenäs yard–Sweden: Ref.[7] studies the Sävenäs yard in Gothenburg, Sweden. In the case study tackled, only two rollback operations are allowed per day, in average over 4 days. The 4-day instances available contain in average 81 inbound and 100 outbound trains, and 1592 cars. The lengths of the classification tracks vary from 360 to 829 m. These tracks are grouped in clusters according to their length (see the detail in ref. [7]). Two classification tracks of 541 m each are used as mixing tracks.

Lausanne Triage yard—Switzerland: The Lausanne Triage yard is the object of the experimental analysis of ref. [37]. It is the third largest yard in Switzerland after Bâle–Muttenz and Zürich–Limmattal. The yard has 11 receiving and 38 classification tracks, linked by two humps and a return track for multistage operations. They measure in total 62 km. There are no departure tracks and the outbound trains leave directly from the classification tracks. In the 1-day instances considered, 1346 cars must be handled to create 22 outbound trains.

MacMillan yard–USA: Ref. [50] consider real data coming from a week operation at the MacMillan yard. It is the second largest yard in Canada, and it is operated by Canadian National Railway (CN). It measures approximately 3 km in length and 1 km in width. The yard shunts over a million cars per year, half of them being empty. The data tackled include more than 4000 empty cars and 272 cuts.

Taschereau Yard—Canada: Refs. [15] and [16] work on the Taschereau Yard in Canada, operated by CN. In the real day considered, eight inbound trains and six outbound trains are managed. No further details are given on the yard layout but for the fact that a large number of classification tracks is available. While in 1983 this instance may have been challenging, it may constitute more of an illustrative example than an actual case study for today's computing possibilities.

Zhengzhoubei yard—China: Ref. [49] deal with three Chinese yards. First, the Zhengzhoubei yard, with 26 receiving and 34 departure tracks and two humps. Seven locomotives are available for managing about 160 inbound and 160 outbound trains including up to 8000 cars going to 36 destinations. Second, the Fengtaixi yard, which also has two humps connecting 23 receiving and 24 departure tracks. Five locomotives are available for 115 inbound and 115 outbound trains including up to 5300 cars going to 28 destinations. Finally, the Chengdudong yard manages 65 inbound and 65 outbound trains with two locomotives. Only one hump is present for shunting up to 2200 cars going to 12 different destinations. The yard includes seven receiving and five departure tracks.

Note on the RAS competition yard on a realistic case: A realistic case study was proposed at the problem solving competition organised by the RAS of the Institute for Operations Research and Management Science [25], which was used in refs. [26], [24], and [27]. The instances are composed of 702 inbound trains with a total of 52,246 cars, over a period of 42 days. Their lengths can range up to 9800 feet with almost 175 cars. There are 16 distinct outbound trains per day with maximum length between 6800 and 9000 feet. Three instances are proposed with a number of classification tracks ranging from 42 to 58 and a total track length ranging respectively from 96,354 to 122,489 feet.

8 | CONCLUSION

In this paper, we proposed a literature review on optimisation algorithms for shunting yard management. One of the main contributions of our survey is to formalise the activities on a shunting yard using a UML representation, in order to clarify the different yard operations and processes and their interactions. We helped clarify the links between different approaches in the literature by including, in an appendix, a glossary of different synonyms and acronyms commonly used in different research communities. We also provided a description of the different case studies used in the existing literature to give an idea of the real case problems that have been dealt with today. Finally, we reported an up-to-date survey of the research contributions on optimisation problems for the management of shunting yards, including the most recent contributions. The main characteristics of these works are summed up in Table 1, where they are classified according to the type of shunting process considered.

In order to keep the survey contained, we chose to focus specifically on shunting yards, as opposed, e.g. to transshipment yards where containers are loaded and unloaded from the trains. We also focused on a type of yard operating car load services, as opposed to full train load services. We have followed the practice in the optimisation literature about shunting yards and ignored a certain number of aspects which are indirectly linked to the yard management process. Such aspects include car registration and car or locomotive maintenance schedule, or even the management of the breaking process of cars and trains, which can all influence the processing time the different tasks on the yard.

Our analysis shows that most contributions deal with a limited part of the yard operations, namely the car shunting. As can be seen on Table 1, most of the existing works focus on the scheduling of inbound trains over the hump and the allocation of the classification tracks, with a more or less detailed scheduling of the yard locomotives. Only a few works try to consider a more global yard management problem, e.g. the allocation of the receiving or departure tracks or the schedule of inspection teams. None of the works in Table 1 considers the routing of either trains or locomotives in the yard at a microscopic level, which may underestimate conflicts and delays of the locomotives for the humping or pulling process of the cars. Considering several of the yard aspects often results in more complex formulations, which would require further progress on the models themselves or on alternative solution techniques, such as exact or heuristic decomposition methods or efficient heuristic algorithms.

Moreover, few works test the results of the proposed algorithm with simulation. This should probably become common practice to allow the research community to assess the contribution of optimisation algorithms independently from the simplifying hypothesis considered in the models. The uncertainty inherent to the management of a freight yard is also rarely taken into account. The frequent delays faced by trains on a railway infrastructure calls for the use of either stochastic or dynamic reoptimisation frameworks which work in closed loop with frequently refreshed data about incoming trains and yard resource state.

In practice, yard resource assignments and activity schedules are still mainly decided manually, with no optimisation method applied. IT systems are developed in every yard but they are mainly used for supervising operations and are not ready to integrate optimisation-based decision support tools. The literature review we presented shows that research is quite ready to support the development of such tools, but significant effort of knowledge transfer is still necessary.

ACKNOWLEDGMENTS

This work was partially supported by the OptiYard project. The OptiYard project has received funding from the Shift2Rail Joint Undertaking under the European Union's Horizon 2020 research and innovation programme under grant agreement No. 777594.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as it surveys the data and findings of other scientific contributions.

ORCID

Samuel Deleplanque https://orcid.org/0000-0003-4119-6006
Pierre Hosteins https://orcid.org/0000-0003-4186-9127
Paola Pellegrini https://orcid.org/0000-0002-6087-651X

REFERENCES

1. Caprara, A., Kroon, L., Toth, P.: Optimization problems in passenger railway systems. Wiley Encyclopedia of Operations Research and Management Science, pp. 3896–3905. Wiley, New York (2010)
2. Boysen, N., Fliedner, M., Jaehn, F., Pesch, E.: A survey on container processing in railway yards. Transp. Sci. 47(3), 312–329 (2013)
3. Bohlin, M., Gestrelius, S., Dahms, F., Mihalák, M., Flier, H.: Optimization methods for multistage freight train formation. Transp. Sci. 50(3), 823–840 (2016)
4. Bontekoning, Y., Priemus, H.: Breakthrough innovations in intermodal freight transport. Transp. Plann. Technol. 27(5), 335–345 (2004)
5. Gatto, M., Maue, J., Mihalák, M., Widmayer, P.: Shunting for dummies: An introductory algorithmic survey. Robust Online Large-Scale Optim. 5868, 310–337 (2009)
6. Boysen, N., Fliedner, M., Jaehn, F., Pesch, E.: Shunting yard operations: theoretical aspects and applications. Eur. J. Oper. Res. 220(1), 1–14 (2012)
7. Gestrelius, S., Aronsson, M., Joborn, M., Bohlin, M.: Towards a comprehensive model for track allocation and roll-time scheduling at marshalling yards. J. Rail Transp. Plann. Manage. 7(3), 157–170 (2017)
8. Bohlin, M., Hansmann, R., Zimmermann, U.T.: Optimization of railway freight shunting. In: Handbook of Optimization in the Railway Industry, pp. 181–212. Springer, Cham (2018)
9. Marinov, M., Viegas, J.: A simulation modelling methodology for evaluating flat-shunted yard operations. Simul. Modell. Pract. Theory 17(6), 1106–1129 (2009)
10. Petersen, E.: Railyard modeling: part i. prediction of put-through time. Transp. Sci. 11(1), 37–49 (1977)
11. Di Stefano, G., Koči, M.L.: A graph theoretical approach to the shunting problem. Electron. Notes Theor. Comput. Sci. 92, 16–33 (2004)
12. Hansmann, R.S., Zimmermann, U.T.: Optimal sorting of rolling stock at hump yards. In: Krebs, H.-J., Jäger, W. (eds.), Mathematics – Key Technology for the Future: Joint Projects Between Universities and Industry 2004–2007, pp. 189–203. Springer, Berlin, Heidelberg (2008)

13. Jacob, R., Márton, P., Maue, J., Nunkesser, M.: Multistage methods for freight train classification. Networks 57(1), 87–105 (2011)
14. Ruf, M., Cordeau, J.-F.: Adaptive large neighborhood search for integrated planning in railroad classification yards. Transp. Res. Part B 150, 26–51 (2021)
15. Yagar, S., Saccomanno, F., Shi, Q.: An efficient sequencing model for humping in a rail yard. Transp. Res. Part A: Gen. 17(4), 251–262 (1983)
16. Saeednia, M., Bruckmann, D., Weidmann, U.: Event-based model for optimizing shunting yard operations. Transp. Res. Rec. 2475(1), 90–94 (2015)
17. Jaehn, F., Rieder, J., Wiehl, A.: Minimizing delays in a shunting yard. OR Spectrum 37(2), 407–429 (2015)
18. Jaehn, F., Rieder, J., Wiehl, A.: Single-stage shunting minimizing weighted departure times. Omega 52, 133–141 (2015)
19. Glover, F., Laguna, M.: Tabu search. In: Handbook of Combinatorial Optimization, pp. 2093–2229. Springer, Berlin, Heidelberg (1998)
20. Dahlhaus, E., Horak, P., Miller, M., Ryan, J.F.: The train marshalling problem. Discrete Appl. Math. 103(1–3), 41–54 (2000)
21. Rinaldi, F., Rizzi, R.: Solving the train marshalling problem by inclusion-exclusion. Discrete Appl. Math. 217, 685–690 (2017)
22. Falsafain, H., Tammanaei, B.: A novel dynamic programming approach to the train marshalling problem. IEEE Trans. Intell. Transp. Syst. 21(2), 701–710 (2020)
23. Boysen, N., Emde, S., Fliedner, M.: The basic train makeup problem in shunting yards. OR Spectrum 38(1), 207–233 (2016)
24. Haahr, J.T., Lusby, R.M.: A matheuristic approach to integrate humping and pullout sequencing operations at railroad hump yards. Networks 67(2), 126–138 (2016)
25. INFORMS: RAS problem statement - railroad hump yard block-to-track assignment. https://www.informs.org/Community/RAS/Problem-Solving-Competition/20142014-RAS-Problem-Solving-Competition (2014). Accessed: 30 September 2014.
26. Khadilkar, H., Sinha, S.K.: Rule-based discrete event simulation for optimising railway hump yard operations. In: 2016 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), pp. 1151–1155. IEEE, Piscataway, NJ (2016)
27. Raut, S., Sinha, S., Khadilkar, H., Salsingikar, S.: A rolling horizon optimisation model for consolidated hump yard operational planning. J. Rail Transp. Plann. Manage. 9, 3–21 (2019)
28. Shi, T., Zhou, X.: A mixed integer programming model for optimizing multi-level operations process in railroad yards. Transp. Res. Part B: Methodol. 80, 19–39 (2015)
29. Kraft, E.: A hump sequencing algorithm for real time management of train connection reliability. J. Transp. Res. Forum 39(4), 95–115 (2000)
30. Kraft, E.: Priority-based classification for improving connection reliability in railroad yards: part i. integration with car scheduling. J. Transp. Res. Forum 56, 93–105 (2002a)
31. Kraft, E.: Priority-based classification for improving connection reliability in railroad yards: part ii. dynamic block to track assignment. J. Transp. Res. Forum 56, 107–119 (2002b)
32. Daganzo, C.F.: Dynamic blocking for railyards: part i. homogeneous traffic. Transp. Res. Part B: Methodol. 21(1), 1–27 (1987)
33. Daganzo, C.F.: Dynamic blocking for railyards: part ii. heterogeneous traffic. Transp. Res. Part B: Methodol. 21(1), 29–40 (1987)
34. Daganzo, C.F., Dowling, R.G., Hall, R.W.: Railroad classification yard throughput: The case of multistage triangular sorting. Transp. Res. Part A: Gen. 17(2), 95–106 (1983)
35. Daganzo, C.F.: Static blocking at railyards: sorting implications and track requirements. Transp. Sci. 20(3), 189–199 (1986)
36. Knuth, D.E.: The Art of Computer Programming, Volume 1: Fundamental Algorithms. Addison-Wesley, Boston, MA (1973)
37. Márton, P., Maue, J., Nunkesser, M.: An improved train classification procedure for the hump yard lausanne triage. In: 9th Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems (ATMOS'09), Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl (2009)
38. Adamko, N., Klima, V.: Optimisation of railway terminal design and operations using Villon generic simulation model. Transport 23(4), 335–340 (2008)
39. Sato, D., Borges, A., Márton, P., Scalabrin, E.: I-DCOP: train classification based on an iterative process using distributed constraint optimization. Proc. Comput. Sci. 51, 2297–2306 (2015).
40. Zhang, Y., Song, R., He, S., Li, H., Guo, X.: Optimization of classification track assignment considering block sequence at train marshaling yard. J. Adv. Transp. 2018(4), 1–11 (2018)
41. Belošević, I., Ivić, M.: Variable neighborhood search for multistage train classification at strategic planning level. Expert Syst. Appl. 33, 220–242 (2018)
42. Büsing, C., Maue, J.: Robust algorithms for sorting railway cars. In European Symposium on Algorithms, pp. 350–361. Springer (2010)
43. Liebchen, C., Lübbecke, M., Möhring, R., Stiller, S.: The concept of recoverable robustness, linear programming recovery, and railway applications. In: Robust and Online Large-Scale Optimization, pp. 1–27. Springer, Berlin, Heidelberg (2009)
44. Cicerone, S., D'Angelo, G., Di Stefano, G., Frigioni, D., Navarra, A.: Recoverable robustness for train shunting problems. Algorithmic Oper. Res. 4(2), 102–116 (2009)
45. Liebchen, C., Lübbecke, M., Möhring, R.H., Stiller, S.: Recoverable robustness. ARRIVAL-Project-Technical Report. http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=88721EFAAD2A7CEE971F4BA18AD0919D?doi=10.1.1.72.2675&rep=rep1&type=pdf (2007). Accessed Aug 2007
46. Li, H., Jin, M., He, S.: Sequencing and scheduling in railway classification yards. Transp. Res. Rec. 2475(1), 72–80 (2015)
47. He, S., Song, R., Chaudhry, S.S.: Fuzzy dispatching model and genetic algorithms for railyards operations. Eur. J. Oper. Res. 124(2), 307–331 (2000)
48. Zadeh, L.A.: Fuzzy sets. Inf. Control 8(3), 338–353 (1965)
49. He, S., Song, R., Chaudhry, S.S.: An integrated dispatching model for rail yards operations. Comput. Oper. Res. 30(7), 939–966 (2003)
50. Bektaş, T., Crainic, T.G., Morency, V.: Improving the performance of rail yards through dynamic reassignments of empty cars. Transp. Res. Part C: Emerg. Technol. 17(3), 259–273 (2009)
51. Schasfoort, B., Gkiotsalitis, K., Eikenbroek, O., van Berkum, E.: A dynamic model for real-time track assignment at railway yards. J. Rail Transp. Plann. Manage. (2020)
52. Saeednia, M.: A framework for managing resources at hump shunting yards. In: 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC 2020), pp 1–6. IEEE, Piscataway, NJ (2020)
53. Licciardello, R., Adamko, N., Deleplanque, S., Hosteins, P., Liu, R., Pellegrini, P., Peterson, A., Wahlborg, M., Za'tko, M.: Integrating yards, network and optimisation models towards real-time rail freight yard operations. Ingegneria Ferroviaria 75(6), 417–447 (2020)
54. Adlbrecht, J.-A., Hüttler, B., Ilo, N., Gronalt, M.: Train routing in shunting yards using answer set programming. Expert Syst. Appl. 42, 7292–7302 (2015)
55. Bostel, N., Dejax, P.: Models and algorithms for container allocation problems on trains in a rapid transshipment shunting yard. Transp. Sci. 32(4), 370–379 (1998)
56. Cichensky, M., Jaehn, F., Pawlak, G., Pesch, E., Singh, G., Blazewicz, J.: An integrated model for the transshipment yard scheduling problem. J. Scheduling 20, 57–65 (2017)
57. Fedtke, S., Boysen, N.: Gantry crane and shuttle car scheduling in modern rail–rail transshipment yards. OR Spectrum 39, 473–503 (2017)
58. Albareda-Sambola, M., Marìn, A., Rodrìguez-Chí, A.M.: Reformulated acyclic partitioning for rail-rail containers transshipment. Eur. J. Oper. Res. 277, 153–165 (2019)
59. Ng, M., Talley, W.K.: Rail intermodal management at marine container terminals: loading double stack trains. Transp. Res. Part C 112, 252–259 (2020)
60. Hirashima, Y.: A new reinforcement learning for multi-train marshaling with time evaluation. IFAC PapersOnLine 53(2), 11144–11149 (2020)

61. Barbour, W., Martinez Mori, J., Kuppa, S., Work, D.: Prediction of arrival times of freight traffic on us railroads using support vector regression. Transp. Res. Part C 93, 211–227 (2018)
62. Minbashi, N., Bohlin, M., Palmqvist, C.-W., Kordnejad, B.: The application of tree-based algorithms on classifying shunting yard departure status. J. Adv. Transp. 2021, 1–10 (2021)
63. Feng, W., Dong, W., Zhai, S., Zhang, G., Sun, X., Ji, Y.: A deep reinforcement learning method for freight train driving based on domain knowledge and mass estimation network. In: 2021 6th International Conference on Machine Learning Technologies (ICMLT 2021), pp. 41–46. ACM Press, New York (2021)
64. Shabelnikov, A., Olgeizer, I., Sukhanov, A.: Machine learning in calculating speeds in a railway sorting yard. IOP Conf. Ser.: Mater. Sci. Eng. 918, 012067 (2020)
65. ARCC: Real-Time Yard Management. European Commission, pp. 1–73. European Commission, Brussels (2017)

How to cite this article: Deleplanque, S., Hosteins, P., Pellegrini, P., Rodriguez, J.: Train management in freight shunting yards: Formalisation and literature review. IET Intell. Transp. Syst. 16, 1286–1305 (2022). https://doi.org/10.1049/itr2.12216

APPENDICES

GLOSSARY

Writing a literature review on the freight trains management in yards is not an easy task since the different keywords used to describe the layouts of yards or to define its operations can be different from one paper to another. This is especially true comparing the literature on railway systems and the literature on operations research. This appendix tries to gather the different synonyms (Syn) and gives the definitions (Def) in order to help the readers on this topic. The definitions without any references have been suggested by operators working at the Česká Třebová yard in Czech Republic during the OptiYard project.

Block:. Def: Blocks are managed through several yards with models generally based on a macroscopic level. A block is a set of cars (or groups) that take a common itinerary over potentially many classification yards. A block is not broken up at intermediate classification yards (ref. [13]).

Car:. Syn: Train car, railroad (US), railcar, wagonload, wagon, train wagon, railway wagon, railway carriage, coach (for passengers trains). Def: Defines the smallest element which can be decoupled in a freight yard to be recombined in a new train.

Classification Track:. Syn: Marshalling track. Def: This type of tracks is used to create new trains from the ones arriving from the receiving tracks. The cars are dispatched to the classification tracks according to the needs of future outbound trains. The set of classification tracks might be called classification bowl. In some yards, the classification bowls contain one or several mixing tracks which are only used to store temporarily cars (they could be from different inbound trains). In this case, the classification tracks which are not mixing tracks are usually called formation tracks. Comparing the works of refs. [3] and [7], we easily see that this specific usage of classification tracks is not always fixed in advance and can be changed dynamically. This specification has naturally an impact on the optimisation process.

Connection:. Def: In the context of a shunting yard, a connexion identifies one inbound train and one outbound train.

Cut:. Def: A cut is a set of cars rolling from the same receiving track to the hump. Generally, a shunting locomotive pushes a full cut for shunting over the hump. According to the layout of the yard and especially to the number and the length of the receiving tracks, a cut can be a single inbound train or, in rare cases, a series of inbound trains without locomotives (ref. [6]).

Decoupling and coupling processes:. Syn: Uncoupling/disassembling/detaching/separation processes and assembling/attaching. Def: In order to make the humping process, the operators decouple the pairs of cars on the receiving tracks to send them in different classification tracks through the hump. Once the cars are on the classification tracks, the operators will couple them in order to create the outbound trains to be dispatched to the departure tracks, or in order to be pulled back. In this case, without the coupling process, only the first car would be pulled. The operators performing the coupling operations are called the couplers.

Departure track:. Def: All the new trains combined in the classification tracks are stored in a departure track, on which they wait to enter the general railway network.

Estimated time arrival/departure (ETA/ETD):. Syn: Arrival/departure time. Def: The estimated time arrival of a train, is the time the operator expects an inbound train to enter a yard. The estimated time departure is the time for an outbound train to leave the yard and to enter the general network.

Group:. Def: Groups are considered into one given yard with models generally based on a microscopic level. More specifically, a group is a set of cars coming from one or several inbound trains and which will be gathered in the same outbound train. Groups are more frequent in European yards compared to US yards where blocks are more often considered (ref. [13]).

Hump:. Syn: Hill. Def: A hump is the place where the cuts are decoupled for ordering its cars. This is usually a tiny hill, natural or not, from where the cars can go down to the classification tracks through the action of the gravity force. A so-called shunting locomotive just has to push them until they go downhill.

Inbound and outbound trains:. Syn: Input trains and output trains. Incoming and outgoing trains; starting and departing trains. Def: The trains arriving in the yard and then stored in a receiving track are called inbound trains. Once the new trains are recombined as expected, they are stored in a departure track and called outbound trains.

Receiving track:. Syn: Arrival track. Def: This is the first type of tracks a train going in the yard will use. The trains are stored on these tracks before being decoupled. The set of all the receiving tracks is called receiving/arrival bowl/yard.

Roll-in operations:. Def: An uncoupled car (or a cut) is pushed to the hump until the cars roll over the hump into the classification tracks. The last operation is called a roll-in. The opposite operation is the pullback defined above. By analogy to the sorting stack algorithm, the roll-in operation is a pop operation and a pullback is a push operation.

Shunting yard:. Syn: Classification yard, marshalling yard. Def: Define the whole yard used to combine new output trains from input trains according to the demand.

Stage:. Syn: step. Def: A stage starts from a series of roll-in operations to the cars dispatched in the classification tracks. We consider the definition of [3] where a Stage is simply defined by "The time from the start of a pullback to the next pullback". In this paper, we split the problems into two categories: the single-stage and the multistage also called as ito-shunting and hump-shunting, respectively (ref. [8]).

Track pull:. Syn: Return track, rehump track. Def: This special track allows the cars to be pulled back from one classification track to a receiving track. The use of such a track is done for "re-roll-in" operations if some cars need to go another time down the hump. From a decisional point of view, the existence of such tracks will determine the type of classification problem that the operator has to manage: the single-stage classification does not use any track pull unlike the multistage classification. In the literature, the pullback operations–sometimes called *t*–*t*-moves for track to track moves—very often take all the cars stored in the classification yard. But, in reality, nothing forces the operator to do such a thing, he can pullback less cars from the front of the track (ref. [6]).