Analysis of Railyard Congestion and Departure Delay Relationship: a Case Study from Swedish Railways

Authors:

Niloofar Minbashi[1] (minbashi@kth.se)

Markus Bohlin[1] (mbohl@kth.se)

Behzad Kordnejad[1] (behzad.kordnejad@abe.kth.se)

[1]Division of Transport Planning, School of Architecture and Built Environment, KTH Royal Institute of Technology, Stockholm, Sweden.

Corresponding Author:
Niloofar Minbashi, minbashi@kth.se, +46 8 790 70 54. KTH Royal Institute of Technology, Division of Transport planning, Brinellvägen 23, SE-100 44 Stockholm, Sweden.

ABSTRACT

In this paper we propose a macroscopic model framework for departure delay prediction from railyards. The railyard is a large area comprising three sub-yards (arrival, classification, departure). In fact, timely operation at railyard is dependent on coordinated operations in these sub-yards. More importantly, punctual functioning of railyards is crucial for increasing competitiveness of rail freight services throughout the network. Despite previous models, we considered railyard congestion at the arrival yard, time availability of each wagon at the classification yard, and time availability of locomotive at the departure yard. The core part of this paper analyzes the effect of congestion at arrival yard on departure delays. Punctuality data from two Swedish railyards for a seven-year period is used. The congestion is defined as the number of arriving trains three hours before each departure. The results showed that the highest number of delayed departures occur at congestion levels of 4-10, while correlation coefficient is around zero. Analysing the whole dataset reveals that these congestion levels are common for all departures not just the delayed ones. Therefore, we conclude that as three sub-yards are interrelated, a comprehensive definition of congestion at railyard level is required. An elaborate definition of congestion can make it a proper predictor for further delay prediction models.

Key words: Departure Delay Prediction, Congestion, Railyards, FR8HUB, Shift2Rail

INTRODUCTION

Rail freight has an essential role in creating and maintaining logistic networks in Europe. European Commission has set a target to increase from 18% to 30% of modal share for rail freight in 2030 [1]. In Sweden, export-oriented industries, such as, iron ore and steel are highly dependent on rail freight. Therefore, remaining competitive with other modes of transport while increasing the modal share for rail requires improvement in efficiency and reliability of rail freight services.

One of the areas that can improve reliability of rail freight networks is railyards. Railyards (referred as yards from now on in this paper) are large areas dedicated to arranging freight wagons. In fact, wagons spend most of their time in yards, for example, reports from two main North American railroads show that nearly 60% of wagon transit time was spent in yards [2], which might be the typical figure for all Class I railroads in North America [3]. This shows that enhanced operation at yards can significantly improve rail freight reliability.

A typical yard is comprised of three sub-yards called; arrival, classification, and departure. The main function of yards is train make-up. The arrival yard receives trains from various origins. Then, each arriving train is disassembled and wagons for different destinations are assigned for different tracks at the classification yard. In the classification yard, wagons are classified based on their destination forming departing trains. When a departing train is ready, a locomotive pulls it to the departure yard. At the departure yard, trains are ready to be dispatched to the railway network whenever the Infrastructure Manager permits them to leave.

The operation in the classification yard is the most labor-intensive one among the sub-yards and it takes up most of the time in train make-up process. As a result, studies have focused enormously on classification yard as the bottleneck of the yard. The main thinking in these studies is that improving the bottleneck will enhance yard performance. While the yard is a system and the sub-yards are interrelated and in constant interaction. In fact, optimizing the performance of the yard should be through systematic and comprehensive view including all three sub-yards. The interaction between sub-yards should be considered in train make-up processes because each sub-yard imposes various delays on departing trains. Moreover, the interaction between yard and network is influential on yard punctuality. In fact, understanding these interactions how they impact on timeliness of freight trains can help us to estimate freight train delays and strive to minimize them. This can improve punctuality of rail freight services in long run and decrease their negative effects on passenger trains. In our proposed approach we try to take into account these interactions; firstly, we consider delays that might be imposed on departing trains from all three sub-yards. Secondly, we assess the interaction between the yard and the network by taking into account the actual arrival times of the trains and wagons.

In this paper, we propose a model framework for departure delay prediction from the yard considering all three sub-yards. The main purpose of this paper is to present analysis on the interaction between the arrival and departure yard. We analyze the relationship between the departure delays and congestion at the arrival yard as the first step towards building our model.

The remainder of this paper is organized as follows. First section discusses the literature with main findings and gaps in the previous studies. Second section introduces our proposed approach as a macroscopic model framework for delay prediction. Third section is about our case study as well as the data we used. In Fourth section the results are discussed. Finally a brief conclusion is presented.

LITERATURE REVIEW

Yards have an important role in reliable performance of rail freight networks. Previous studies have mainly focused on optimizing operations at classification yard, such as, assignment of wagons to classification tracks, and sorting/rearranging wagons. In fact, previous studies showed that classification yard, as the critical part of a yard, can have direct impact on service reliability of the yard [4, 5, 6]. In line with this, the sorting and rearranging the wagons problem in the classification yard has received the highest attention [7]. There are two forms of sorting single-stage sorting and multistage sorting.

Single-stage sorting refers to the situation where the number of tracks in the classification yard is enough to make departing trains. This problem has been studied through various objective functions, such as, minimization of “throughput costs”, minimization of dwell time [3] [8], minimization of the number of used tracks [9].

Whereas, multistage sorting refers to the situation where the number of classification tracks is less than the number of departing trains, so the cars need to be pulled back from the classification yard to the arrival yard, hence, some tracks in the classification yard are allocated to pullback process and called mixing tracks. The main objective function for this problem has been minimization of the number of pull backs or expressed in another term “extra roll-ins” by an optimal track allocation at classification yard [6, 10,11, 12].

One of the main drawbacks of sorting models, particularly multistage sorting models, is that most of them don’t consider timetabled departure times of wagons, even if we try to optimize the yard by minimizing the dwell time of wagons, each wagon is still restricted to leave the yard according to a timetable. Therefore, future models should consider the scheduled departure time of the wagons in delay estimation models. In fact, the yard is part of the railway network and is in constant interaction with the railway network, as a result, departure delays from yards will affect other trains on the network.

Another drawback in the yard studies in general is that the arrival yard, the classification yard and the departure yard have been regarded as single and separate entities. The interaction between the three sub-yards hasn’t been fully considered. Except one study [5] that considers the yard as a whole system and tries to improve yard performance based on lean manufacturing concepts and minimizing dwell time of wagons to increase the yard efficiency. Otherwise, most of the studies have focused on the classification yard as the bottleneck in the yard, while this approach has not helped to optimize the whole yard performance substantially. Future studies should focus on studying all three sub-yards simultaneously and analyze how they can affect and interact with each other.

PROPOSED APPROACH

In this paper we propose a macroscopic model framework for departure delay prediction of the yards[13]. The actual departure time depends on punctual operations of all of the three sub-yards. Different operations can cause departure delay not to the same extent. The ideal way would be to encompass all the factors that engender delays. However, it’s not possible to gather data for all the factors or it would be hard to implement such a model computationally. Therefore, we chose three main factors from each sub-yard that impact departure delay significantly. This approach enables us to simplify the problem and build a macroscopic model.

The model has three components each representing one sub-yard; the component for the arrival yard is the congestion in terms of the number of arriving trains that are present in the yard three hours before the departure. The component for the classification yard is the arrival time of each wagon, and for the departure yard the delay in the arrival time of the locomotive has been considered.

The mathematical formulation for departure delay prediction is based on the following mathematical expressions. The function in (1) predicts the actual departure time. The delay is calculated as the difference between the scheduled and the actual departure time in (2), the notation can be found on Table 1.

$$t_b^{\Delta} = \alpha\left(t_a^{\Delta} - T_1\right) + \sum_i \beta_i\left(t_i^{\Delta} - T_2\right) + \gamma \#_i\left(t_b^{tt} - T_3 \leq t_i^{arr} \leq t_b^{tt}\right) \tag{1}$$

$$\delta_b = t_b^{\Delta} - t_b^{tt} \tag{2}$$

TABLE 1 Notation

Parameter	Description
t_b^{Δ}	the actual departure time of train b
t_a^{Δ}	the arrival time of the locomotive
t_i^{Δ}	the arrival time of wagon i
t_b^{tt}	the timetabled departure of train b
g_{ti}	the number of trains in the arrival yard when train i arrives the yard
γ	Waiting time in the arrival yard per train (time/train)
α, β_i	Correlation coefficients
b	Total number of departing trains in a specific time-span
δ_b	The difference between the actual departure time of train b and the timetabled departure of train b
T_1, T_2, T_3	Buffer Times

CASE STUDY

We intend to use empirical data from two main yards (Hallsberg and Malmö) in Sweden (See Figure 1). These two yards are suitable representative yards in Europe. The Hallsberg yard is a large yard with a conventional yard layout, and Malmö yard is a congested yard that connects Swedish Railways to the rest of European rail network.

In Sweden, the infrastructure manager is responsible for the arrival and departure yards, and the yard operator manages the classification yard. Hence, the databases related to each sub-yard are different and have different owners (infrastructure manager and yard operator). This makes the progress of the modelling rather difficult and we need to move forward step by step in the modelling process.

In this paper, we aim to evaluate the first component of the model which is the impact of congestion on the departure delay. The data to evaluate this is Swedish Railways' database for punctuality and disruptions. This database contains train information, such as, train number, train type, train path, scheduled/actual departure and arrival.

In order to analyse departure deviations from Hallsberg and Malmö, we used train information that belongs to 2010-12-12 to 2017-12-09 timespan. This time period encompass around 255 000 departures from Malmö and Hallsberg in total.

We selected two main variables for the analysis; 1) departure delay: the difference between actual departure time and scheduled departure time from the yard, 2) yard congestion: the number of arriving trains three hours before each departure. Three hours was chosen because it takes around three hours for a departing train to be prepared for departure. In fact, when the arrival yard is congested, the yard operator is involved with handling the cars between the arrival and classification. As a result, departure delay might occur.

According to the Swedish Transport Administration, trains are on time if they arrive or depart up to 5 minutes later than their scheduled time. Any deviation beyond this is late and ahead of this is early. Additionally, freight trains with delays more than 12 hours are due to exceptional causes and are removed from our database. The trains for which the actual arrival/departure time wasn't recorded were also removed from our analysis as well as the cancelled trains.

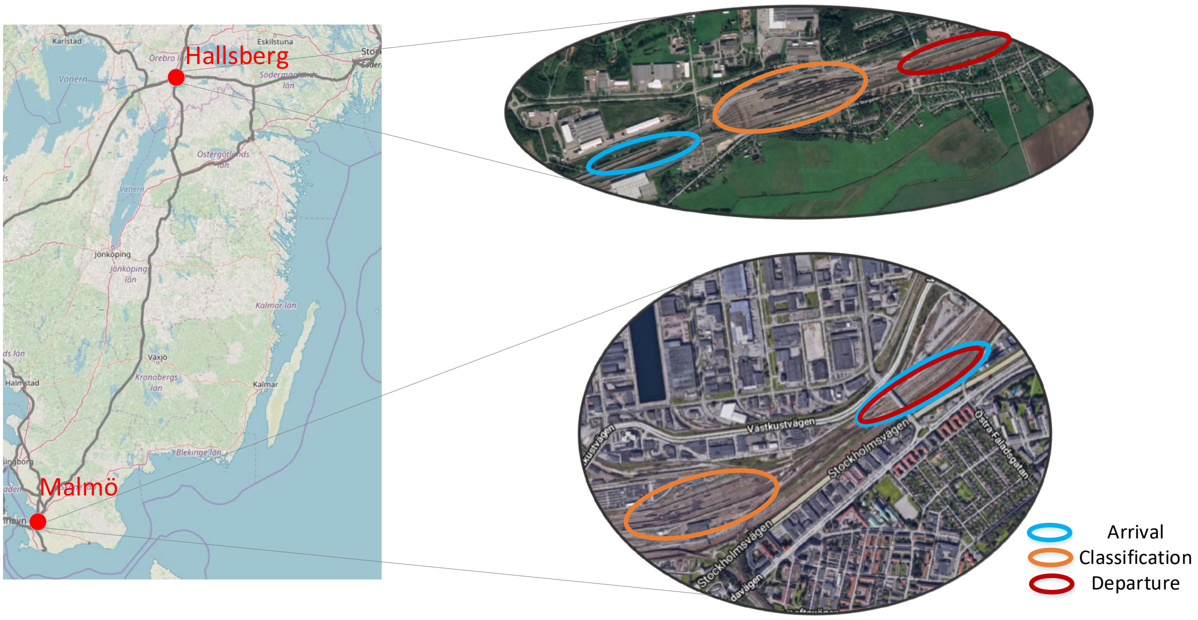

FIGURE 1 the Location of the case study ad yard layouts

RESULTS

In this section, we present the results of analysing the relationship between arrival yard congestion and departure delays in Figure 2 and Figure 3. The X axis represents the congestion as the number of arriving trains three hours before each departure. Three hours is the average time that it takes for preparing a departing train. We assume that congestion at yard can have an effect on departing trains. The Y axis represents the number of delayed departures in (a) and all departures in (b) and (c) over the year.

First, we analyse Malmö yard (Figure 3 (a)). The highest number of departure delays occur when congestion at the arrival yard is around 4-10 (the number is slightly different from year to year, but the trend is fairly the same for all years in our dataset, however, we show 2016 and 2017 as an example). As the operators claim, a viable explanation can be the yard layout; at Malmö yard arrival and departure yards are the same. As a result, when the yard operator is occupied with handling the arriving trains, it will have difficulty to dispatch departing trains on time. However, it's also important to test whether this level of congestion at yard is common for all departures, not just the delayed ones. Therefore, we depicted the number of all departures versus congestion levels at Figure 3 (b). We can see that 4 to10 is a common congestion level at Malmö yard. At this level of congestion, we have the highest number of departures in total. Figure 3 (c) shows that there is a relative distribution between delayed (27%), on time (12%) and early (60%) departures for congestion level from 4 to 10. This shows that we can not conclude firmly the relationship between departure delay and congestion at the arrival yard. Further analysis showed a Pearson correlation coefficient of -0.02 which confirms that there is no correlation between these two parameters.

We observed similar results for Hallsberg as well. An extract of the results of 2016 is shown on Figure 4. Likewise, the highest number of delays occurred when the congestion level was from 4 to 9 (Figure 4 (a)). In Figure 4 (b), we can see that distribution between delayed (21%), on time (10%) and early (69%) departures is fairly the same throughout these congestion levels. The Pearson correlation is -0.05 which validates lack of correlation.

To sum up, the selected parameter lacks a full representation of congestion at arrival yard. One drawback might be that congestion should be calculated at the yard level not just at the arrival yard level. If so, we need to elaborate congestion with these variables; the initial number of trains at the arrival yard, the number of trains being classified at the classification yard, and the number of trains ready at the departure yard. So far, a limitation to consider these variables is the difficulty to retrieve data; different databases need to be linked to give a proper estimation of them. Another variable that can be added in congestion estimation is track occupancy at sub-yard levels. Currently, we are working on elaborating congestion parameter by these variables.

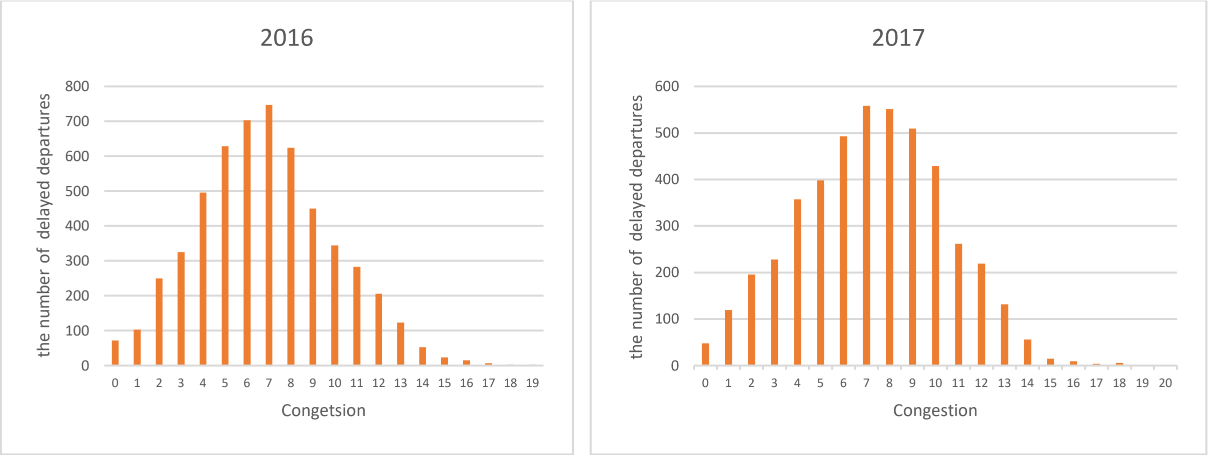

FIGURE 3 (a)

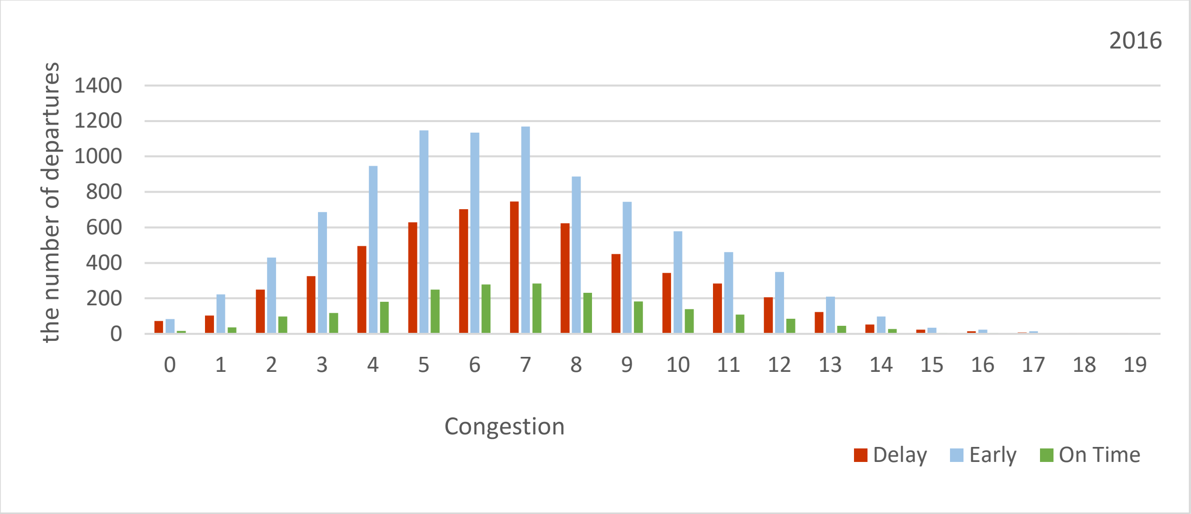

FIGURE 3 (b)

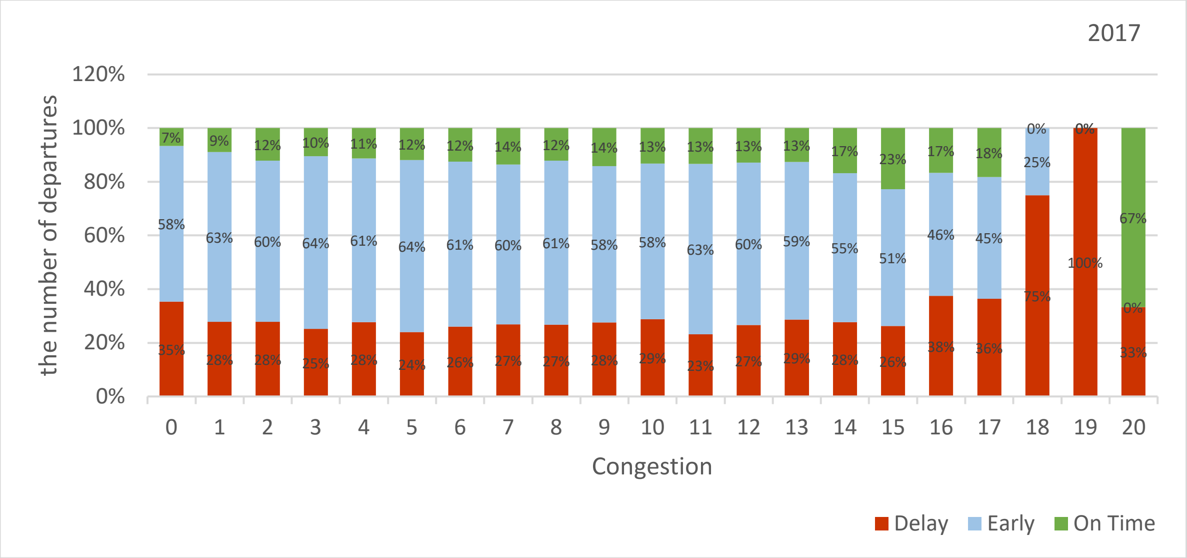

FIGURE 3 (c)

FIGURE 3 the relationship between yard congestion and departure delays for Malmö yard

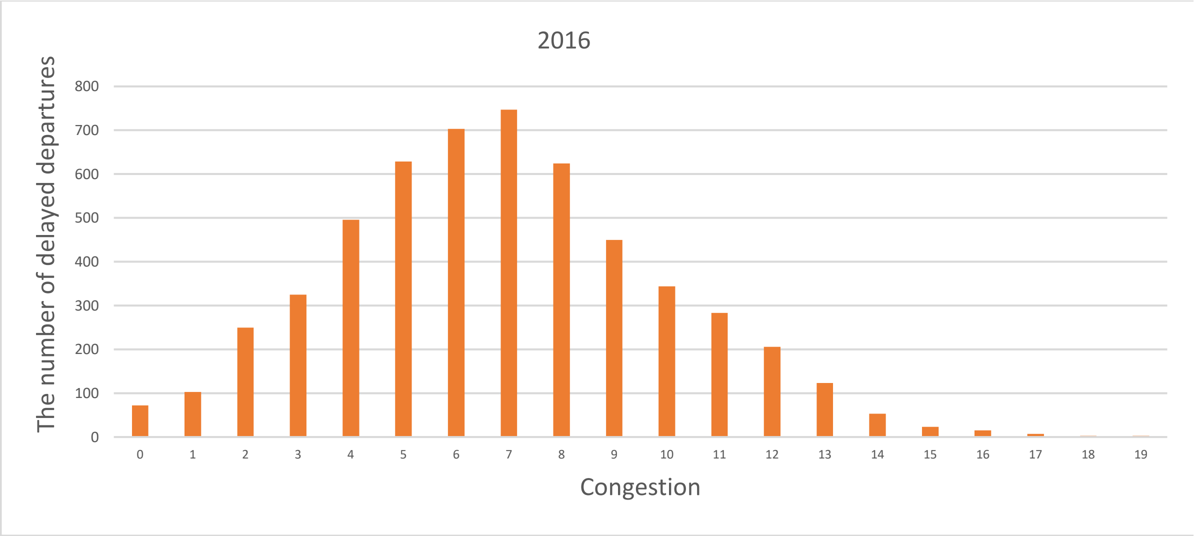

FIGURE 4 (a)

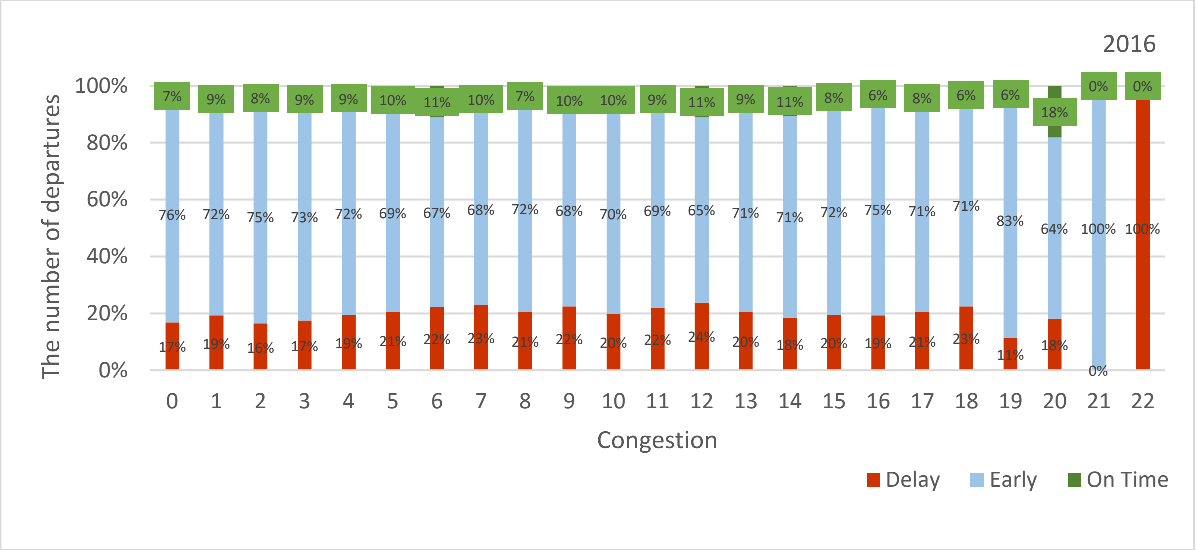

FIGURE 4 (b)

FIGURE 4 the relationship between yard congestion and departure delays Hallsberg yard

CONCLUSION

In this paper we introduced a model framework for departure delay prediction at yards. In spite of previous models, we considered yard congestion at the arrival yard, time availability of each wagon at the classification yard, and time availability of locomotive at the departure yard. We evaluated the effect of congestion using a seven-year period data from two main yards in Sweden (Malmö and Hallsberg). We defined congestion as the number of arriving trains three hours before each departure. The results showed that the highest number of delayed departure occur at congestion levels of 6-10 (Malmö) and 4-9 (Hallsberg). However, correlation coefficient was around zero. To confirm this, we analysed the whole dataset which revealed that these congestion levels are common for all departures. We conclude that current definition of congestion needs to be elaborated. It could be extended by calculating congestion at yard level not just the arrival yard. Future work is required to investigate this which is also a part of our current research.

ACKNOWLEDGEMENT

This study is financed by FR8HUB project within the European H2020 - Shift2Rail IP5, and we would like to show our gratitude to Magnus Wahlborg from Trafikverket for sharing his expertise with us throughout the project. We would like to thank Carl-William Palmqvist (Lund University) for sharing a part of his data with us.

REFERENCES

[1] Rail Freight Forward, "30 by 2030 Rail Freight strategy to boost modal shift," 2018.

[2] P. Logan, "Role of yard or terminal in operating performance and capacity," in *Presentation at 85th Annual Meeting of the Transportation Research Board, Washington, DC*, 2006.

[3] J. R. Dirnberger and C. P. L. Barkan, "Lean railroading for improving railroad classification terminal performance: Bottleneck management methods," *Transp. Res. Rec.*, no. 1995, pp. 52–61, 2007.

[4] C. Martland, P. Little, O. Kwon, and R. Dontula, "Background on Railroad Reliability. AAR Report No. R-803," Washington, DC, 1992.

[5] J. Dirnberger, "Development and application of lean railroading to improve classification terminal performance," 2006.

[6] M. Bohlin, F. Dahms, H. Flier, and S. Gestrelius, "Optimal freight train classification using column generation," *drops.dagstuhl.de*, 2012.

[7] N. Boysen, M. Fliedner, F. Jaehn, and E. Pesch, "Shunting yard operations: Theoretical aspects and applications," *Eur. J. Oper. Res.*, vol. 220, no. 1, pp. 1–14, 2012.

[8] T. Bektaş, T. G. Crainic, and V. Morency, "Improving the performance of rail yards through dynamic reassignments of empty cars," *Transp. Res. Part C Emerg. Technol.*, vol. 17, no. 3, pp. 259–273, Jun. 2009.

[9] R. S. Hansmann and U. T. Zimmermann, "Optimal sorting of rolling stock at hump yards," in *Mathematics--key technology for the future*, Springer, 2008, pp. 189–203.

[10] M. Bohlin, H. Flier, … J. M.-11th workshop on, and U. 2011, "Track allocation in freight-train classification with mixed tracks," *drops.dagstuhl.de*, 2011.

[11] M. Bohlin, S. Gestrelius, F. Dahms, M. Mihalák, and H. Flier, "Optimization Methods for Multistage Freight Train Formation," *Transp. Sci.*, vol. 50, no. 3, pp. 823–840, Aug. 2016.

[12] M. Bohlin, H. Flier, J. Maue, M. M.-4th I. S. On, and U. 2011, "Hump yard track allocation with temporary car storage," *diva-portal.org*, 2011.

[13] Fr8hub. Project, "Deliverable 3.2 Demonstration of FR8HUB Network Management Concept.," 2019.