2017 4th International Conference on Transportation Information and Safety (ICTIS), August 8-10, 2017, Banff, Canada
144

Data-driven Models for Predicting Delay Recovery in High-Speed Rail

Chao Wen[1]
National United Engineering Laboratory of Integrated and Intelligent Transportation, China, 610031
Department of Civil and Environmental Engineering, University of Waterloo, Canada, N2L 3G1
wenchao@swjtu.cn

Javad Lessan[2]
Department of Civil and Environmental Engineering, University of Waterloo, Canada, N2L 3G1
javadlessan@gmail.com

Liping Fu[3]
School of Transportation and Logistics, Southwest Jiaotong University, China, 610031
Department of Civil and Environmental Engineering, University of Waterloo, Canada, N2L 3G1
Intelligent Transport System Research Center, Wuhan University of Technology, Wuhan, China
lfu@uwaterloo.ca

Ping Huang[4]
School of Transportation and Logistics, Southwest Jiaotong University, China, 610031
huangping@my.swjtu.cn

Chaozhe Jiang[5*]
National United Engineering Laboratory of Integrated and Intelligent Transportation, Southwest Jiaotong University, China, 610031
Department of Civil and Environment Engineering, University of Waterloo, Canada, N2L 3G1
jiangchaozhe@swjtu.cn

***Abstract*---One of the main challenges arising in a high-speed railway (HSR) is predicting how fast a train, once delayed, can recover its operation. Accurate prediction of delay recovery in the downstream stations of a HSR line can help train dispatchers make adjustments to the timetables and inform the passengers of the expected delay to improve service reliability and increase passenger satisfaction. In this paper, we present the results of an effort to develop data-driven delay recovery prediction models using train operation records from the Centralized Traffic Control system (CTC) of Wuhan-Guangzhou (W-G) HSR in Guangzhou Railway Bureau. We first identified the main variables that contribute to delay, including total dwell (*TD*) time, running buffer (*RB*) time, magnitude of primary delay (*PD*), and individual sections' influence. Two alternative models, namely, multiple linear regression (MLR) and random forest regression (RFR), are calibrated and evaluated. The validation results on test datasets indicate that both models have good performance, with the RFR model outperforming the MLR in terms of prediction accuracy. Specifically, the evaluation results show that when the prediction tolerance is less than 3 minutes, the RFR model can achieve up to 90.9% of prediction accuracy, while this value is 84.4% for MLR model.**

Keywords---High-speed railway, Primary delay recovery, Data-driven approach, Multiple linear regression model, Random forest regression model

I. INTRODUCTION

Despite the availability of advanced communication and control technologies, HSR lines could still be subject to significant delay in their daily operations due to various unexpected disruptions such as weather events, power issues, and facilities failure (Khadilkar [2016]). For example, according to the statistics of the China Railway Corporation, the average departure punctuality for China's 20,000 km HSR network is as high as 98.8%. However, due to various disturbances during their operations, the average punctuality at the final destination stations is less than 90% (China-Railway-Corporation [2016]). Another example is Japan's HSR systems, which are internationally reputed for their service punctuality, they have still suffered an average delay of 0.9 minutes per train at their destination stations (Central-Japan-Railway-Company [2015]).

Punctuality is considered as one of the most important performance measures in any transit mode including HSR (Harris et al. [2013], and Lindfeldt [2010]). A train service, once delayed at a station or in a section, could continue arriving and departing late at the downstream stations, which was commonly referred as to delay propagation. When a train arrives late to a station, because of some exogenous delay, it may transfer some delay to other trains (Mattsson [2007]). A delayed train can cause delays to several other trains over a large operating area and time period. Even worse, the delay of just one train may cause a whole cascade of delays to other trains over the entire railway network (Vromans et al. [2006]) and further delays and conflicts at train interactions and transferring points (Delorme et al. [2009]). Major disturbances can propagate to other trains in the network, thus requiring short-term adjustments in the timetable in order to limit delay propagation (Corman et al. [2012]). How fast a train service can recover from its delay in the subsequent stations, called delay recovery, and is an important performance measure that shows the reliability and robustness of the service being provided. The ability to predict delay recovery plays an important role in estimating running and dwell times that are essential parts of managing any railway transport services (Kecman and Goverde [2015]).

The remainder of this paper is organized as follows. Section 2 briefly reviews delay management literature. Section 3 provides a formal description of the delay recovery problem as well as the operational data collected from the W-G HSR line. In section 4 we discuss the details of the presented models, the calibration and validation processes, and the comparison results of the proposed models. Finally, conclusions and future study directions are discussed in Section 5.

II. LITERATURE REVIEW

Delay management in preventing and recovering from delays has raised considerable interests in the literature (Corman and D'ariano [2012], and Kliewer and Suhl [2011]). Estimating and updating running times is a common interest about delay propagation, and it plays an important role during delay recovery. Workflow nets and triangular fuzzy numbers are applied by Wen et al. [2015] to study the issues associated with predicting train operation conflicts. However, the most common limitation of these work is that there are limited in incorporating real data on train operations to illustrate and validate the presented models practically.

There are two types of delay recovery models, one is based on simulation and theoretical hypotheses, and the other is based on historical data. In the former type, the issue of delay recovery was usually handled in the timetable rescheduling. A review on methods for real-time scheduling and recovery problems was carried out by Visentini et al. [2014]. They reviewed three classes of real-time schedule recovery: vehicle rescheduling for road-based services, train based rescheduling, and airline schedule recovery problems. Another overview of recovery models and algorithms for real-time railway disturbance and disruption management was presented in Cacchiani et al. [2014] that mainly summarized the methods on real-time timetable rescheduling of the rolling stock and crew duties. The size of primary delays (and possibly primary delays) and the ability of a timetable to absorb primary delays are characterized as two important ingredients for timetable robustness and performances measurement (Fernandez et al. [2004]). For the latter type, Hansen et al. [2010] presented a delay propagation model in which train path conflicts and dispatching decisions are taken into account, and parameters are estimated by offline statistical analysis of historical train detection data (Hansen et al. [2010]). However, this work does not address the delay probabilities at each station and in each section. Jouni uses a data-mining approach for analyzing rail transport delay chains, with data from passenger train traffic on the Finnish rail network. However, the data of train operations was limited only to one month record (Wallander and Ma¨kitalo [2012]). The empirical data from Dutch (Hansen et al. [2010]), Spanish (Cadarso et al. [2013]), Finnish (Wallander and Ma¨kitalo [2012]), Britain (Jaroszweski et al. [2015]), and Taiwan (Jong et al. [2010]) were used for modeling. Among this data, only the data of Taiwan came from the HSR. Simulation models were also applied to investigate the problem of timetable rescheduling and delay recovery (Keiji et al. [2015]). The concept of resilience has been introduced as a criterion to measure the system ability in absorbing perturbations, and its ability to recover rapidly from perturbations (Adjetey-Bahun et al. [2016]). Khadilkar [2016] studied the delay probability distributions and found the mean recovery rate of 0.13 minute/km based on historical data from the Indian Railways network, which is used in modeling delay recovery. However, these approaches cannot handle large-scale real-world railway networks that require fast and efficient solutions to cope with unexpected disturbances during trains' operation. Moreover, due to lack of detailed operational data, the presented models in the existing literature are mostly limited to incorporating real train operation records.

III. PROBLEM DEFINITION AND DATA DESCRIPTION

A. Problem Definition

Train operations on a HSR line is a dynamic process, affected by many factors including train schedule, operating constraints and various perturbations. When a train is held back on a track or at a station due to some external disturbances, the resulting delay is called *primary delay* or source delay. A late train could affect the operations of other trains at the downstream stations and sections. The delay caused to other trains is commonly referred as to *knock-on delay* or *secondary delay*. This study concerns with the primary delays. When a delay occurs to a HSR train, various mitigation actions could be taken, mostly initiated by dispatchers, such as adjusting running speeds and dwell times to absorb the incurred delay. The ability to predict delay recovery, especially at the terminal stations, plays an important role in adjusting running and dwell times to manage delay and reduce further delay in the downstream stations. The underlying process of delay propagation and recovery estimation is depicted in Figure (1), and formulated in Equations (1) and (2). Stations, indexed by k, are numbered sequentially from 1 to n. Consider a particular train servicing the line, starting from the first station (1) and running sequentially toward the terminal station (n). Assume the train had experienced a primary delay of PD_0 at a station (note that this delay could equivalently be initiated at the section preceding the station), which is the difference between the actual departure and the schedule departure time at the station. This delay will propagate downstream but could also be reduced through operational adjustments such as increasing train speed and reducing dwell time. The amount of delay recovered by a specific downstream station (k), denoted by RT_k, is defined as the difference between the amount of the primary delay the train suffered (PD_0), and its arrival delay at the station k, as shown in Equation (1).

$$RT_k = PD_0 - (AT_k - SAT_k) \tag{1}$$

where SAT_k, and AT_k are the scheduled arrival and actual arrival times of the train at station k. In this research, we are particularly concerned with predicting the delay recovery time at the destination station, i.e., RT_n.

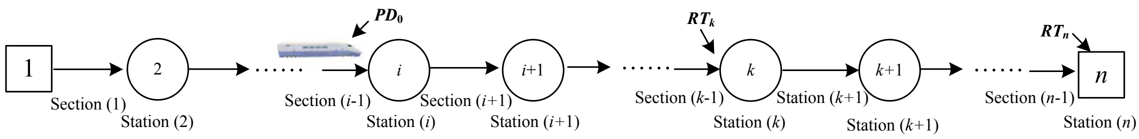


Fig.1. Sketch of train operation and delay propagation in a HSR line

The buffer time that scheduled in the sections and at the stations can be used for delay recovering after the train is delayed by (PD_0) in section (i-1), and there is Equation (2) and (3).

$$TD = \sum_{i} dt_i \tag{2}$$

$$RB = \sum_{i+1} rb_{i+1} \tag{3}$$

where TD and RB are the total dwell time and the total running buffer time respectively in the rest journey after the train is delayed. While dt_i is the dwell time at station i of the train, and there is dt_i=0 if the station is a non-stop station of the train. And rb_{i+1} is the running buffer time in section i +1, and there is rb_{i+1} if there is no buffer time scheduled in section i +1.

B. Wuhan-Guangzhou HSR

A case study is conducted in this research using daily train operation records from the W-G HSR. As shown in Figure (2), W-G HSR connects Wuhan (Hubei province) to Guangzhou (Guangdong province) with a total of 1096 km of double-tracks and 18 stations. In this research, only the operation data for trains connecting GuangzhouNorth and ChibiNorth were available (including 14 stations). The train operation data were obtained for the period from February 24, 2015 to November 30, 2015, including a total of 29,662 records. All HSR train operations in China are fully controlled by the CTC system that records all operational events and data on a second by second intervals. The trains are continuously monitored by a remote control system via the signal block system. When a train passes a station signal, this event is recorded by a remote control and simultaneously transmitted to the CTC system. Table (1) shows a sample of train operation records. We use these data to determine the arrival delays, departure delays, and dwell times at each station.

Fig.2. Map of Wuhan-Guangzhou HSR

TABLE.1. Train running records in a database

Train NO.	Date	Station	Scheduled Arrival	Scheduled Departure	Actual Arrival	Actual Departure
G634	2015-2-24	GZN	17:26:00	17:26:00	17:28:00	17:28:00
G6152	2015-2-24	QY	17:16:00	17:18:00	17:18:00	17:20:00
G9694	2015-2-24	YDW	19:00:00	19:02:00	19:00:00	19:03:00
G548	2015-2-24	SG	17:26:00	17:29:00	17:25:00	17:29:00

IV. MODELING FRAMEWORK OF DELAY RECOVERY

A. Delay Recovery Analysis

The primary delay (PD) is the main factor causing trains' late arrivals at, or departures from stations. In order to keep a timetable flexible and to recover trains from PD, a certain amount of buffer time is usually included in the timetable. The supplementary times are allocated in the form of additional minutes in dwell time at individual stations and running time over individual sections. Buffer time is the difference between the scheduled running time (dwell time) in the timetable and the minimum required running time (dwell time). To understand the potentials of delay recovery, we analyzed the arrival and departure events of the delayed trains during their dwelling at the stations or running over the sections case-by-case.

Especially, we see that the delayed trains have significant recovery potential, more than 4 minutes when they pass through the ZZW-CSS section. This shows that the ZZW-CSS section plays a significant role in recovering delayed trains. We use this finding to distinguish trains which are passing through this section (ZZW-CSS). Moreover, our findings support that the dwell times at the stations and buffer times in the sections can help reduce the propagation of delays in the network. As a result, we use *TD* and *RB* to include their effects in the recovering from delay.

So far we have identified the main factors that may contribute to the delay recovery of trains (*RT*), namely, *PD*, *TD*, *RB*. *PD* represents the severity (duration) of the delayed event. *TD* is the total scheduled dwell time for all stations between the points at which the train has experienced a primary delay and the terminal station. *RB* is the total (scheduled) buffer times in all of the downstream sections up to the terminal station. Moreover, as it was revealed, among all the stations and sections, the ZZW-CSS section has a significant effect of reducing the propagation of the passing trains' delay. To incorporate this factor, we introduce a binary variable *ZC* in our models to distinguish those delayed trains that pass through this section. Finally, the recovery time *RT* is the dependent variable in our models, representing the total delay recovered during the rest of the trip, once a train is delayed.

B. Model Calibration

Before the operational data described previously were used for model calibration, they were further processed. According to most European railway companies, trains arriving less than 5 minutes late are not considered to be delayed (Yuan [2006]). We followed the same convention by excluding these observations with a delay less than 5 minutes or above 60 minutes. Furthermore, the data that represent secondary delay records were also excluded. The filtered data set includes a total of 917 observations or events with primary delays. The summary statistics of the model variables are shown in Table (2), and the histogram of the samples given each variable is shown in Figure (3). Particularly, we see that the correlation between each pair of factors are relatively weak, allowing us to use them together as independent variables without the concern of multi-collinearity.

We divided our data into training and test datasets. To achieve a reliable model regarding stability measures, we attempted different sizes of random samples from our training data to compare the stability of the models. As a criterion we used the mean of squared residuals of 100 experiments under each proportion. Figure (4) shows the mean of squared residuals boxplots under different proportion of training data. The figure indicates that a higher training sample size results in a better stability level, with a more concentrated residuals distribution. In addition, when more than 70% of the samples are selected for training data, the models can acquire good performances. As a result, we took randomly 70% of the sample for modeling, and used the rest of 30% of the observations for validation.

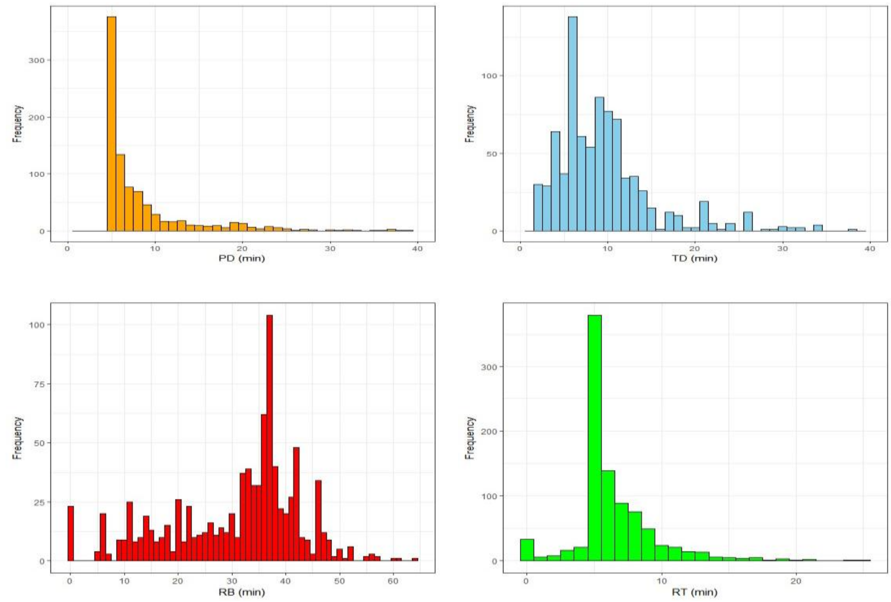


Fig.3. Distributions of the observations given each independent variable

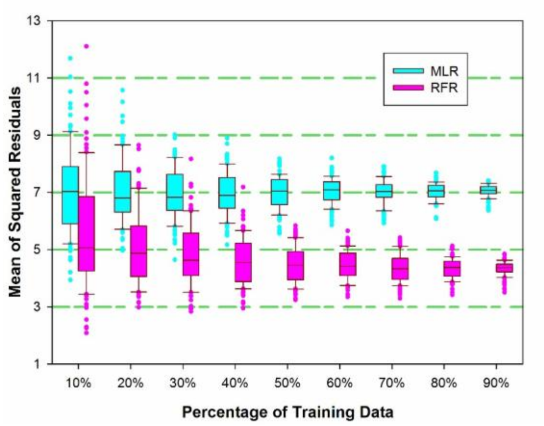


Fig.4. Mean of squared residuals for different size of training data (colored online)

TABLE.2. Summary of the descriptive statistics for the identified factors

Variable	N	Min*	Max*	Mean	SD*	Correlation *PD*	*TD*	*RB*	*RT*
PD	917	5.00	64.00	8.98	7.24	1	0.14	0.11	0.59
TD	917	0.00	58.00	8.90	6.53	0.14	1	0.45	0.34
RB	917	0.00	64.00	30.66	12.33	0.11	0.45	1	0.28
RT	917	0.00	25.00	6.38	3.10	0.59	0.34	0.28	1

[*] Values are in minutes.

C. Multiple Linear Regression Model

Generally, regression models are used to construct a relationship between two or more explanatory variables (independent) and a response (dependent) variable by fitting a linear equation to the data. We built a MLR model of PDR with R-program, the regression results are provided in Table (3). As can be seen, all of the regressors are highly significant, suggesting that they should all be included as part of the linear model. At the bottom of the table, the row "Model Summary" provides a summary of the test statistics of the regression model and variable coefficients including R^2, F-test, and t-test. We can say that with $R^2 = 0.85$, the MLR model shows a good fitness and relationship between the predictors and the response variable. Moreover, the p-value of F-test supports that the relationship is statistically significant, i.e., the dependent variable (RT) can be explained (at least) by one of the independent variables or a combination of these variables.

Overall, the regression results show that the MLR model for PDR makes sense. Therefore, the regression formula for the primary delay recovery reads as follows:

$$RT_n = 0.14 \cdot TD + 0.06 \cdot RB - 0.15 \cdot PD + 2.05 \cdot ZC \qquad (4)$$

The absolute value of the coefficients show that *PD* has the highest effect, and *TD* has more effect than *RB* on recovering from delays. Though *ZC* was considered as a variable in this model, it just plays the role of a constant value as it is a binary variable, so it has the least importance. The negative coefficient of *PD* implies that some part of the source delay (around 85%) already recovered at the arrival of a (primary) delayed train to its destination station. As a result, with the current train schedule, trains can considerably recover from primary delays that are less than 60 minutes. Indeed, the average arrival delay of late trains in our dataset (917 of observations), which is relatively low, i.e., $\sum_{k=1}^{n}(AT_k - SAT_k) \simeq 2.5$ minutes, confirms this.

TABLE.3. Regression Results

Factor	*Est.*	*Std. err.*	*t value*	*Pr(>|t|)*	*Sig.*
TD	0.14	0.02	8.78	0.00	***
RB	0.06	0.01	5.04	0.00	***
ID	-0.15	0.01	-10.64	0.00	***
YN	2.05	0.39	5.25	0.00	***
Model Summary	(a) Residual standard error: 2.79 on 638 degrees of freedom (b) Multiple R-squared: 0.85, Adjusted R-squared: 0.85 (c) F-statistic: 916.90 on 4 and 638 DF, p-value: < 0.00				

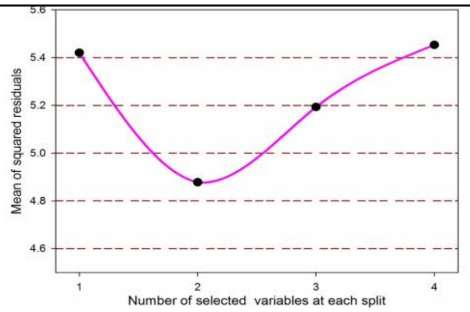
D. Random Forest Regression Model

Random Forest is a machine learning method for classification and regression. This modeling approach uses an ensemble of decision trees for mapping a relationship between a vector of predictor and dependent variables. The output of the model is the modes of the classes (classification) or mean prediction (regression) of the individual trees (Breiman [2001] and Breiman [1996]). In our RFR modeling there are four predictor variables, as a result for constructing each node of a decision tree, we considered four different scenarios each with a different number of variables. Figure (5) depicts the number of variables at each split, and the nodes (tree size) of the trees. Figure 5(a) shows the mean of the squared residuals (errors) under a different number of selected variables. As it can be seen, when two variables are selected, the smallest mean of squared residuals can be obtained. Figure 5(b) shows the complexity of a tree, and that the forest is quite complex when 74 to 132 nodes are included in each tree.

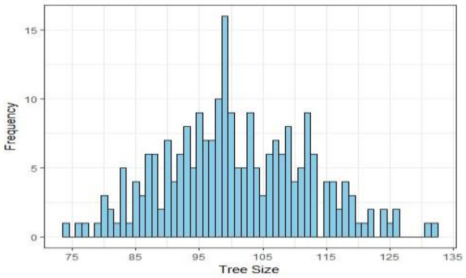
In order to find a reasonable forest scale (number of trees) and get a balance between the accuracy and computation speed, we examined prediction errors under different tree sizes. More explicitly, using RFR models with 1 to 500 decision trees, we examined prediction errors from the respective forest size, see Figure (6).

TABLE.4. Coefficients of the identified factors in RFR model

Variable	*PD*	*TD*	*RB*	*ZC*
Coefficient	49.33	14.51	11.85	6.67



(a) Mean squared residuals



(b) Tree's complexity

Fig.5. Errors, and tree's complexity under different number of variables (nodes)

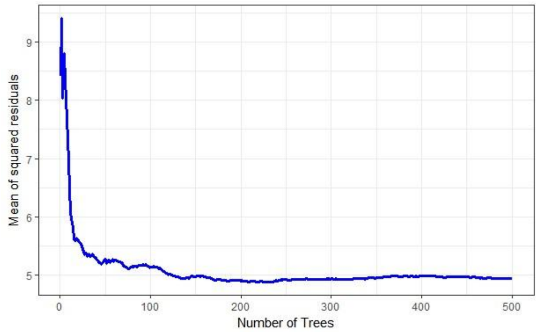


Fig.6.Prediction errors of models given different number of decision trees

The result shows that errors do not change significantly when the number of trees is higher than 200, suggesting that a forest with about 200 trees should be efficient for modeling PDR. Consequently, we built a forest with 231 decision trees where each tree has two variables on each node and the corresponding root mean squared residuals (RMSE) is 4.89 (see Figure 5(a)). There is no certain formulation for a RFR model, however, we can derive the main coefficients of the variables of the model, as provided in Table (4). The coefficients show that *PD* plays the most important role, then *TD*, and *RB* comes next, and the variable *CZ* has the least importance. These results are in accordance with what we concluded from the coefficients of MLR model.

E. Validation and Performance Comparison

We compared the prediction results obtained from the presented models with the respective observations in our test dataset. The results are shown in Figure (7), where the predicted values for PDR are called "MLR Predicted" and "RFR Predicted" respectively, while the corresponding values observed (*RT*) are named as "RT Actual". The vertical axis is the recovery time in minute, while the horizontal axis is the observation number in the dataset. The comparison of means under a t-test resulted a p-value less than of 0.00 for $\alpha = 0.05$, which supports that statistically the difference between predictions obtained from RFR model and MLR model are significant. The results of accuracy evaluation of the MLR and RFR models are plotted out in Figure (8). The assessments are based on comparing the residuals of the recovery times (*RT*) predicted by the models and those in the test dataset. The bars in the Figure (8) show the frequency of absolute residuals between the predicted values and the observations, while the curves show the accumulated residuals, given the defined intervals. We can see that in 122 instances, about 44.4% of the total test instances, the prediction error (absolute residual) by MLR model is less than 1 minute. While in the RFR model, we have 221 instance, about 79.3% of the total test instances, with an error less than 1 minute. We can see that all the absolute residuals of RFR model is less than 8 minutes, while the absolute residuals of MLR model reach up to 9 minutes. Especially, when the permitted absolute residuals are less than 3 minutes, the accumulated residuals is 84.4% for MLR model, and 90.9% for RFR model, indicating that both models present good performances in predicting PDR, yet RFR model has a higher prediction accuracy.

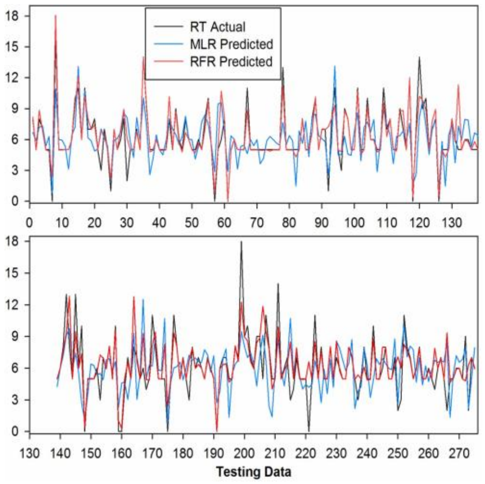


Fig.7. Prediction results of the models and the observations in testing data set (colored online)

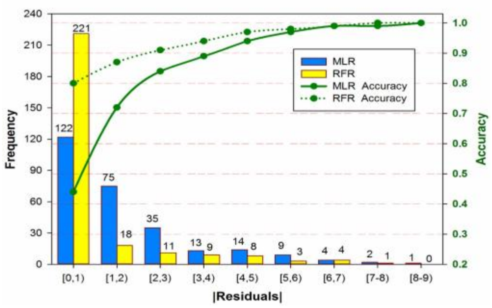


Fig.8. Accuracy comparison of the MLR and RFR models (colored online)

F. Effectiveness Comparison

To examine the effectiveness of the presented models, we compared the distribution of prediction residuals for both the training and testing datasets, respectively. The comparison results are depicted in Figure (9). It shows that a great proportion of the residuals are around zero in RFR model, meaning that the accuracy of the RFR model is higher. These results show that RFR model can predict the PDR with a higher accuracy and robustness level.

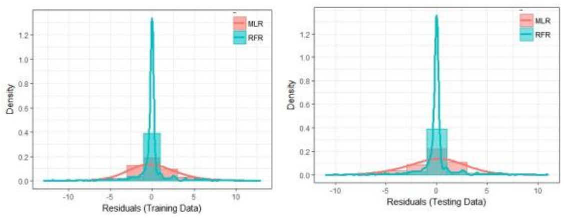


Fig. 9. Distributions of the predictions' residuals for both training and test datasets (colored online)

The comprehensive evaluation showed that the RFR model is more reliable than MLR model in predicting delay recovery.

V. CONCLUSIONS AND FUTURE RESEARCH

In this paper, two data-driven methods were proposed to tackle the problem of predicting delay recovery of HSR trains that suffer from a primary delay, namely, multiple linear regression and random forest regression. The validation tests indicate that both models can achieve good performance levels while the random forest model is superior to the multiple linear regression model in delay recovery prediction. Though the models are established using data from Wuhan-Guangzhou HSR, they can be extended to other HSR networks with similar characteristics. Furthermore, our models can be beneficial to both practitioners and researchers in the field of delay management.

The research presented in this paper represents the first step toward the development of a comprehensive service management system for a HSR network. Our ongoing work focuses on developing a general framework and model construct that can be applied to any HSR line. This will involve investigating different HSR lines using real operation data with a greater spatial and temporal coverage. We will also be examining the individual distribution of arrival and departure delays at the station- and section-levels. Additionally, we will consider the delay recovery distributions of each station and section. Lastly, we will combine the primary and knock-on delays into an integrated delay recovery model for optimizing service operations and timetabling of HSR.

ACKNOWLEDGEMENT

This work was supported by the National Nature Science Foundation of China [grant number 61503311] and National Key R&D Plan [grant number 2017YFB1200701]. We acknowledge the support of the Railways Technology Development Plan of China Railway Corporation [grant number 2016X008-J]. Parts of the work were supported by State Key Lab of Railway Control and Safety Open Topics Fund [grant RCS2017K008]. We are grateful for the contributions made by our project partners.

REFERENCES

[1] Harshad Khadilkar. Data-enabled stochastic modeling for evaluating schedule robustness of railway networks. *Transportation Science*, 2016.

[2] China-Railway-Corporation.Punctuality reaches 98.8%, 2016. URL http://news. gaotie.cn /yunying/2016-07-23/336706.html.

[3] Central-Japan-Railway-Company. Central Japan railway company annual report 2014, Aug 2015. URL http://english.jr-central.co.jp/company/ir/annu alreport/_pdf/annualreport2014.pdf.

[4] Nigel G Harris, Christian S Mjosund, and Hans Haugland. Improving railway performance in Norway. *Journal of Rail Transport Planning & Management*, 3(4):172–180, 2013.

[5] Lindfeldt. Evaluation of punctuality on a heavily utilised railway line with mixed traffic. Timetable Planning and Information Quality, page 115, 2010.

[6] Pavle Kecman and Rob MP Goverde. Predictive modelling of running and dwell times in railway traffic. *Public Transport*, 7(3):295–319, 2015.

[7] Lars-Go¨ran Mattsson. Railway capacity and train delay relationships. In *Critical Infrastructure*, pages 129–150. Springer, 2007.

[8] Michiel JCM Vromans, Rommert Dekker, and Leo G Kroon. Reliability and heterogeneity of railway services. *European Journal of Operational Research*, 172(2):647–665, 2006.

[9] Xavier Delorme, Xavier Gandibleux, and Joaqu´ın Rodriguez. Stability evaluation of a railway timetable at station level. *European Journal of Operational Research*, 195(3): 780–790, 2009.

[10] Francesco Corman, Andrea D'Ariano, Dario Pacciarelli, and Marco Pranzo. Bi-objective conflict detection and resolution in railway traffic management. *Transportation Research Part C: Emerging Technologies*, 20(1):79–94, 2012.

[11] Francesco Corman and Andrea D'ariano. Assessment of advanced dispatching measures for recovering disrupted railway traffic situations. *Transportation Research Record: Journal of the Transportation Research Board*, (2289):1–9, 2012.

[12] Natalia Kliewer and Leena Suhl. A note on the online nature of the railway delay management problem. *Networks*, 57(1):28–37, 2011.

[13] Chao Wen, Jianguang Li, Qiyuan Peng, Bingyu Li, and Juanjuan Ren. Predicting highspeed train operation conflicts using workflow nets and triangular fuzzy numbers. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 229(3):268–279, 2015.

[14] Monize Saˆmara Visentini, Denis Borenstein, Jing-Quan Li, and Pitu B Mirchandani. Review of real-time vehicle schedule recovery methods in transportation services. *Journal of Scheduling*, 17(6):541–567, 2014.

[15] Valentina Cacchiani, Dennis Huisman, Martin Kidd, Leo Kroon, Paolo Toth, Lucas Veelenturf, and Joris Wagenaar. An overview of recovery models and algorithms for real-time railway rescheduling. *Transportation Research Part B: Methodological*, 63: 15–37, 2014.

[16] A Fern´andez, AP Cucala, and JI Sanz. An integrated information model for traffic planning, operation and management of railway lines. *WIT Transactions on The Built Environment*, 74, 2004.

[17] Ingo A Hansen, Rob MP Goverde, and Dirk J van der Meer. Online train delay recognition and running time prediction. In *Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on*, pages 1783–1788. IEEE, 2010.

[18] Jouni Wallander and Miika M¨akitalo. Data mining in rail transport delay chain analysis. *International Journal of Shipping and Transport Logistics*, 4(3):269–285, 2012.

[19] Luis Cadarso, Angel Mar´ın, and G´abor Mar´oti. Recovery of disruptions in rapid transit´ networks. *Transportation Research Part E: Logistics and Transportation Review*, 53:15–33, 2013.

[20] David Jaroszweski, Elizabeth Hooper, Chris Baker, Lee Chapman, and Andrew Quinn.The impacts of the 28 June 2012 storms on uk road and rail transport. *Meteorological Applications*, 22(3):470–476, 2015.

[21] JC Jong, TH Lin, CK Lee, and HL Hu. The analysis of train reliability for the taiwan high speed rail. *WIT Transactions on The Built Environment*, 114:169–180, 2010.

[22] Kariyazaki Keiji, Hibino Naohiko, and Morichi Shigeru. Simulation analysis of train operation to recover knock-on delay under high-frequency intervals. *Case Studies on Transport Policy*, 3(1):92–98, 2015.

[23] Kpotissan Adjetey-Bahun, Babiga Birregah, Eric Chaˆtelet, and Jean-Luc Planchet. A model to quantify the resilience of mass railway transportation systems. *Reliability Engineering & System Safety*, 153:1–14, 2016.

[24] Francesco Corman, Andrea D'Ariano, Dario Pacciarelli, and Marco Pranzo. A tabu search algorithm for rerouting trains during rail operations. *Transportation Research Part B: Methodological*, 44(1):175–192, 2010.

[25] Leo Breiman. Random forests. Machine learning, 45(1):5–32, 2001.

[26] Leo Breiman. Bagging predictors. *Machine learning*, 24(2):123–140, 1996.