2015 IEEE 18th International Conference on Intelligent Transportation Systems

Data driven approaches for passenger train delay estimation

Ren Wang, *Student Member, IEEE,* and Daniel B. Work, *Member, IEEE*

***Abstract*— Train delay is a critical problem in railroad operations, which has led to the development of analytical and simulation based approaches to estimate it. With the recent advances in sensing and communication technologies, train positioning information is now available to support new data driven methods for train delay estimation. In this work, two data driven approaches are proposed to estimate train delays based on historical and real time information. A historical regression model is proposed to estimate future train delays at each station using only past performance of the train along the route. Next, several variations of an online regression model are proposed to estimate delay using delay information of the trains at earlier stations along the current trip, as well as delay information of other trains that share the same corridor. The proposed methods are tested with data collected on 282 Amtrak trains (the largest US passenger railroad service) from 2011 to 2013, which consists of more than 100,000 train trips. Compared to prediction based on the scheduled time table, the proposed historical regression model improves the RMSE estimate of delay by 12%, while the online proposed model improves the RMSE estimate of delay by 60%.**

I. Introduction

For a given train, the delay is defined as the difference between the true running time and the free running time. The variabilities associated with train operations (e.g., equipment maintenance, station dwell time, weather) may contribute to the travel time delay, which consequently impacts the efficiency of railroad operations [1]. In the US, Amtrak passenger trains have priority over freight trains, and yet the average on–time rate of Amtrak is less than 50% [2]. Moreover, the average delay for several trains can reach as high as 50 minutes (e.g., the Adirondack train at Fort Edward station in 2013 [2]). In the presence of this variability, it is important to estimate train delays so that strategies can be developed to improve the railroad operation efficiency.

The objective of this article is to develop new data driven methodologies to estimate passenger train delays and to assess their performance on a large dataset of more than 100,000 trips. In the past, many analytical models and simulation approaches have been proposed to estimate train delays. While these approaches have merit due to their elegance (analytical approaches) and realism (simulation based approaches), application of either approach constitutes a major model building or calibration task. For complex systems, analytical methods require some degree of abstraction to maintain tractability. Simulation based approaches can model the complexity of the realistic train operations, but require extensive effort to accurately calibrate the model.

With the recent advances in sensing and communication technology, train positioning data is now available to improve train delay estimation through data driven methods. For example, regression models can be constructed to estimate delay, where the parameters associated with the regression models are calibrated by learning from historical data. Compared to the analytical methods and simulation methods, data driven approaches can be easily generalized and deployed to estimate train delays for any train, as long as training data is available. Note this necessarily prevents the applications of these methods for scenario planning, which analytical or simulation approaches are more appropriate.

The main challenge associated with data driven approaches for passenger train delay estimation is data availability. First, accurate data may not be available, or it may be sparse or incomplete. For example, the Amtrak data considered in this work does not contain records between stations, and no information is publicly available about the freight traffic, which shares the same track. Moreover, the data is incomplete, and some delays are never recorded. In spite of these limitations, this work shows standard regression models can significantly improve passenger train delay estimation compared to the predictions based on the scheduled time table. While additional refinements are certainly warranted, data driven approaches appear promising for delay estimation.

The main contributions of this article are summarized as follows. This article proposes two data driven approaches for passenger train delay estimation. A historical regression model is designed to predict train delays before the current trip starts, and online regression models are proposed to provide a more accurate train delay estimate after the trip begins, using the delay recorded at the upstream station on the current trip and the delay recorded by other nearby trains. Data from 282 Amtrak trains (over 100,000 trips), are used to illustrate and test the proposed algorithms. The estimation results show the proposed historical regression model improves the *route mean square error* RMSE by 12% and the online regression model improves the RMSE by 60%, compared to prediction based on the scheduled time table.

The reminder of the article is organized as follows. In Section II, we review the existing work for train travel time delays. In Section III, data driven autoregressive approaches are proposed for train delay estimation. The proposed methods are implemented and tested with Amtrak data, and the estimation results of the proposed methods are shown in Section IV. In Section V, we conclude the proposed

R. Wang is with the Department of Civil and Environmental Engineering, University of Illinois at Urbana Champaign, IL. email: (renwang2@illinois.edu).

D. Work is with the Department of Civil and Environmental Engineering and Coordinated Science Laboratory, University of Illinois at Urbana Champaign, IL. email: (dbwork@illinois.edu).

978-1-4673-6596-3/15 $31.00 © 2015 IEEE
DOI 10.1109/ITSC.2015.94

methods significantly improve delay estimation compared to the scheduled time table, and note the need for further work on capturing cascading delays and other delay related factors.

II. Related work

In this section, we provide a brief review on train delay estimation, following the excellent and detailed review by Murali et al. [3].

A. Analytical delay estimation methods

Frank [4] proposed to estimate train travel time by studying the accumulation of trains when traffic exceeds the capacity. The model assumes that no overtakes are allowed, the departing times are uniformly distributed, and that the speed of each train is unique and constant. This work was later extended by Peterson [5] and Chen et al. [6], where factors such as overtakes, different speeds, priority systems, and uncertainties associated with train departure time were considered. Carey et al. [7] used stochastic approximation to analyze the effects of headways on knock–on delays of trains, and the impacts of dispatching strategies on train delays and passenger waiting time were analyzed by Özekici et al. [8]. Higgins et al. [9] proposed a model to quantify the expected positive delay for individual passenger trains and track links in an urban rail network. A stochastic modeling approach proposed by Yuan [10] estimate train delays and delay propagation in stations by using probability distributions to model train events and process times based on empirical data.

While the analytical methods provide explicit mathematical relationships to estimate delays, the delays caused by the complex interactions among trains, the variabilities among train operations, and operating parameters cannot be fully captured by the mathematical formulations. As a result, simulation based models that incorporate these aspects have also been developed for train delay estimation.

B. Simulation methods

Peterson et al. proposed a structured model [11] to simulate the movement of trains over the rail line, which supports arbitrary number of trains with different speeds and priories on single or multiple tracks rail networks with sidings, switches and cross–overs. Dessouky et al. [3], [12], [13] proposed simulation models that are able to simulate train movements over single and double track lines, junctions, and terminals and to model rail networks that consist of multiple trackage configurations and speed limits. Their proposed simulation model [13] has been tested and validated with the movements of passenger and freight trains in Los Angeles.

The simulation based models have also been used to calibrate the parameters of the analytical models for delay estimation. Hallowell [14] proposed an analytical model to study the impacts of randomness of departure times and dispatch policies on delays. The model is calibrated and validated through rail operating data generated by extensive Monte Carlo simulations using an optimal meet and passes planning model.

The main advantage of the simulation models is that they are capable of incorporating the sophisticated interactions of trains on complex infrastructure, and the resulting delays can be easily estimated once the model is calibrated. However, simulation models are still an approximation to the true rail operations, and thus are limited to considering the delay factors which are explicitly modeled. Moreover, simulation approaches usually require extensive simulations in order to generate delay distributions of different operating conditions [3].

C. Data driven methods

Data driven approach becomes an alternative approach for train delay estimation in the recent decade. Forman [15] applied linear regression on a first class freight railroad (BNSF), with the goal to determine the factors contributing to delay. The model is performed on eight districts with widely varying traffic patterns and track configurations. Train delays are estimated for each of the eight districts, and factors such as horsepower per ton, track geometry, train priorities, meets, passes, overtakes, train spacing variabilities, are considered. It is the first work which uses regression models on US freight rail data for delay prediction. Compared to freight trains, passenger trains are scheduled traffic and stop frequently in order to pick up and drop passengers. Moreover, the number of factors considered in the regression is significantly smaller, since data on passenger trail is significantly more limited compared to data available (internally) to the freight railroad.

Kecman and Governed [16] present a microscopic model to predict train travel time and delay for railroad networks. Historical track occupation data are used to train the parameters in the microscopic model to estimate train delays. The proposed model are tested on a corridor in Netherlands in a simulated real-time environment. Another work that uses data driven approach for train delay estimation is performed by Hansen et al. [17], where an online model is trained with historical track occupation data, and implemented on a segment of Dutch railway corridor. In the US, railroad track occupation data are not public available. The regression models proposed in this article use train departure time records at stations to estimate train delays.

Recently, a tool was created by Google and Amtrak to track the Amtrak trains and provide arrival time prediction [18]. However, here are no published results on the algorithms or their accuracy. Thus, this work presents a first quantitative and data driven study on methods for estimation of passenger train delays in the US.

III. Methodology

In this work, we develop two approaches to estimate train delays. The first method is a historical regression model developed by assuming delays from one trip to the next follow an vector autoregressive process. This model predicts train delays at each station before the current trip starts based on the delay recorded in the past trips. Next, two variations of an online regression model are developed, which aims

at providing accurate train delay estimation by using delay information of the train at earlier stations long the current trip, as well as delay information of other trains that share the same corridor.

A. Historical regression model

Passenger train delay can be assumed to follow a vector autoregressive process [19] , because passenger trains operate on a fixed frequency (e.g., daily) and schedule. As a result, prior delays on previous trips bring information to estimate the train delay at each station for the current trip. The vector autoregressive process predicts train delays at each station along the route simultaneously based on the prior delays on previous trips. The historical regression model constructed by a vector autoregressive process of order p is described as follows:

$$y_t^i = A_1 y_{t-1}^i + \cdots + A_p y_{t-p}^i + \nu + u_t, \tag{1}$$

where $y_t^i = \left(y_{1,t}^i, \cdots, y_{k,t}^i, \cdots, y_{K,t}^i\right)^T \in \mathcal{R}^K$ denotes the vector of train delays on trip t for train i. Here, $y_{k,t}^i$ is a scalar that denotes the train delay at station k on trip t for train i, with $k = 1, \cdots, K$, and K is the total number of stations on the train trip. The matrix $A_m \in \mathcal{R}^{K \times K}$, with $m = 1, \cdots, p$, denotes the relationships of delays among the current and past trips, and among stations. The variable $\nu = (\nu_1, \cdots, \nu_K)^T \in \mathcal{R}^K$ is an intercept term which allows constant delays. The variable $u_t = (u_1, \cdots, u_K)^T \in \mathcal{R}^K$ is the *white noise* which denotes the error between the predicted $\hat{y}_t$ and the true y_t, where $\hat{y}_t$ is given as:

$$\hat{y}_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + \nu. \tag{2}$$

Model (1) is also called a vector autoregressive process with lag p, since the vector y_t is computed using only the system state in the previous p trips. To apply the model, we first select p, and then train the parameters A_m and ν by using a least squares fit on the historical data. Then, the vector autoregressive process with the trained parameters can be used for prediction.

B. Online regression model

The historical regression model can predict train delays at each station before the current trip starts. After the trip begins, the accuracy of the train delay estimation at a station can be further improved if delays of the train at its upstream stations are known, and if the delays of another trains that may interact with the current train are known. In this section, an online regression model is proposed to incorporate such information for train delay estimation by using an autoregressive process [20]:

$$y_{k,t}^i = a_1 y_{k-1,t}^i + \cdots + a_p y_{k-p,t}^i + c_k + \Phi_{k,t}^i + u_{k,t}, \tag{3}$$

where $y_{k,t}^i$ is a scalar that denotes the delay of train i at station k during trip t. The parameters a_m and c_k denote the relationship of train delays among the current and past stations. The term $u_{k,t}$ in (3) is a scalar which denotes the error associated with the model. The predictor $\hat{y}_{k,t}^i$ is given as:

$$\hat{y}_{k,t}^i = a_1 y_{k-1,t}^i + \cdots + a_p y_{k-p,t}^i + c_k + \Phi_{k,t}^i. \tag{4}$$

The term $\Phi_{k,t}^i$ denotes the delays of another trains that may contribute to the delay of train i at station k. This term is modeled as:

$$\Phi_{k,t}^i = \Sigma_{(j,\tilde{k},\tilde{t}) \in \Omega_{i,k,t}} b_j y_{\tilde{k},\tilde{t}}^j, \tag{5}$$

where $\Omega_{i,k,t}$ denotes the set of train–station pairs that contribute to $y_{k,t}^i$. The term $y_{\tilde{k},\tilde{t}}^j$ is the delay of train j at station $\tilde{k}$ during trip $\tilde{t}$. Note that station $\tilde{k}$ is not necessarily the same station as k since the delay of train j at other stations $\tilde{k}$ may also influence the delay of train i at station k (e.g., if train i and train j share the same track, but move in opposite directions). Moreover, trip $\tilde{t}$ must be distinguished from t since it is a trip index for train j. The parameter b_j is the factor that indicates how the delay of train j at station $\tilde{k}$ on trip $\tilde{t}$ impacts $y_{k,t}^i$.

The existence of Φ can be interpreted as follows. If two trains are closely scheduled on a single track line and the front train is delayed at a station, then it is possible for the following train to experience knock–on delay. Note that because Amtrak shares track with freight trains, and freight train positioning data is not publicly available, the knock–on delay caused by freight traffic cannot be captured when this model is implemented with Amtrak data only.

We also consider two variations of the online regression model (3). The first one is a predictor which is constructed based on the assumption that the delay of train i at station k is simply equal to the delay of the same train at station $k-1$. In this case, the model (3) becomes:

$$y_{k,t}^i = y_{k-1,t}^i + u_{k,t}. \tag{6}$$

The second variation of regression model does not consider the delays caused by the interactions among trains. The simplified (interaction free) model is given as:

$$y_{k,t}^i = a_1 y_{k-1,t}^i + \cdots + a_p y_{k-p,t}^i + c_k + u_{k,t}, \tag{7}$$

As a result, model (6) can be viewed as a baseline approach where delay is assumed to propagate from the upstream station to the downstream station. Model (7) captures non–constant delay relationship between stations by training the parameters a_m and c_k, while model (3) adds another component Φ to incorporate the delay caused by interactions among trains.

IV. Implementation and results

The proposed methods are tested with Amtrak passenger train data released by *AmtrakStatusMaps* [2]. The historical regression model (2), the online baseline model (6), the online regression non–interacting model (7), the online regression interacting model (3), and the scheduled time table are tested and compared using data from 282 Amtrak trains (or 120 trains in the case of the interacting model (3) since the other trains have an empty interaction set $\Omega_{i,k,t}$).

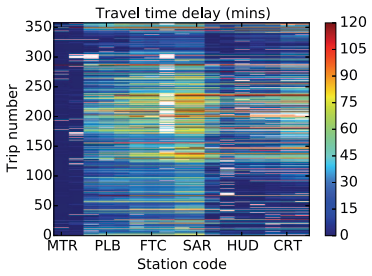


Fig. 1: The figure shows delays for Amtrak train 68 in 2013. Six of the 17 stations along the route are labeled. The color in the figure denotes departure delay at each station for each trip. Missing data are shown in white.

The general procedure to evaluate the models is as follows. First, structural parameters of each regression model must be selected (e.g., the lag p and the interaction set $\Omega_{i,k,t}$). Second, the available data is partitioned into a training dataset and a test dataset. Third, the parameters of each regression model are determined through least squares estimation on the training dataset. Finally, the model is evaluated on the test dataset to determine the accuracy of the predictor.

A. Data description and training data selection

Amtrak data is used as an example of passenger train data to illustrate and test the proposed method. The dataset contains all Amtrak passenger train arrival and departure data at each station from 2006 to 2013. The dataset is released by *AmtrakStatusMaps* [2] and is publicly available. For each train and each trip, the following data are recorded: *station code*, *scheduled arrival day and time*, *scheduled departure day and time*, *actual arrival time*, *actual departure time* and *comments*.

After an exploration of the dataset, it is found that the Amtrak data is coarse and a number of data records are missing. In particular, most stations do not have records for actual train arrival times, and some stations do not have records for scheduled train arrival times. However, nearly all the stations have records for the scheduled departure time and the actual departure time. As a result, the time difference between scheduled departure time and actual departure time is used to denote travel time delay in the following experiments.

A year of delay data of a typical train (train 68 in 2013) is used as an example to visualize the delay (Figure 1). The train travels daily from Montreal (MTR) to New York, and stops at 18 stations. In Figure 1, two data patterns can be observed from the recorded delays. First, some stations are more likely to experience delay compared to the others (e.g., Ticonderoga, NY (FTC) compared to Hudson, NY (HUD)). Second, once delay occurs on a trip, the delay is likely to last for several stations. Such data patterns are also commonly observed on other Amtrak trains.

Data from 282 Amtrak trains from 2011 to 2013 are used to train and test the proposed algorithms, which consists of more than 100,000 train trips. For each train, data from 2011 and 2012 are used as training data, while the first 30 trips of 2013 for each train are used as test data. Note that the first 30 trips may occur over one to several months depending on the frequency at which the train operates (e.g., daily, weekly). We also note that *AmtrakStatusMaps* contains data for more than 450 Amtrak trains from 2011 to 2013, however a regression model cannot be constructed for all trains. The vast majority of excluded trains were subject to a route re-configuration (e.g., adding a station) during the three year period, meaning that a complete set of training or test data is not available. A small subset of trains without schedule reconfigurations were also excluded due to a large amount of missing data. These are practical issues that must be addressed before data-driven methods can be widely deployed.

When models (2), (6) and (7) are tested, data from all 282 Amtrak trains are used as training data. The online regression interacting model (3) is evaluated on a smaller subset consisting of 120 trains, where the interacting set $\Omega_{i,k,t}$ is non-empty.

B. Cross validation

In order to test if the results of the proposed methods are sensitive to the training data, a k–fold cross validation [21] is used. The training data is partitioned to five sets. Each model is run five times, and for each run, a district set composed of four of the five sets are used to construct the training dataset. Different from the standard k–fold cross validation where the test dataset is also changed during each fold, the algorithm is tested with the data in 2013, to avoid the scenario that the model is trained with data from the future and tested on the past.

C. Selection of structural regression parameters

We briefly describe how the order p for models (3), (2), and (7) is selected, and how the set $\Omega_{i,k,t}$ is determined when model (3) is deployed.

When the historical model (2) is implemented, multiple p values have been tested, and it is found that the historical model has the overall best performance when the order p is set as one. Practically, the order p associated with the historical model for each train can also be determined individually by minimizing the final prediction error following the criteria in [19]. When the online regression models (3) and (7) are implemented, the order p is also chosen as one. Because once the train delay at the upstream station is known, the delays of the train from the stations further upstream do not contribute to the estimation accuracy. This assumption was also tested by evaluating larger orders p for the model, which caused slight decreases in the predictive accuracy.

When the online regression interacting model (3) is implemented, the set $\Omega_{i,k,t}$ for each station is constructed according to the following assumption. If train j is scheduled at a the same or neighboring station $\tilde{k}$ within an hour of train i at station k, then the delay of train j at station $\tilde{k}$ is considered as part of the regression. As a final step we prune any trains that are scheduled at the same station but do not

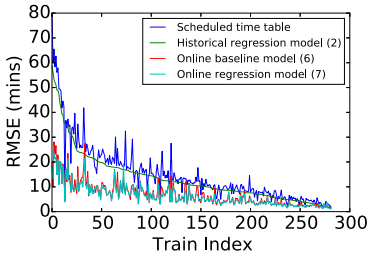


Fig. 2: Ranked average MSE associated with scheduled time table, models (2), (6), and (7) for each train.

share the same track, which is common at major terminals such as Chicago's Union Station.

D. Regression results without interactions among trains

In this section, the historical model (2) and the online models (6) and (7) are trained and tested with the data from the 282 Amtrak trains.

The average RMSE e_i of the proposed models for train i is computed as follows:

$$e_i = \sqrt{\left(\frac{1}{N}\sum_{n=1}^{N}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\frac{1}{K}\sum_{k=1}^{K}u_{n,k,t}^2\right)\right)\right)}, \quad (8)$$

where N denotes the total number of cross validations and T denotes the total number of trips to be estimated. The term $u_{n,k,t}$ is the model error of the t^{th} trip for the n^{th} cross validation for train i. The mean square error is computed and averaged over the T estimated trips, and then averaged over the N cross validations. In this simulation, $T = 30$ and $N = 5$.

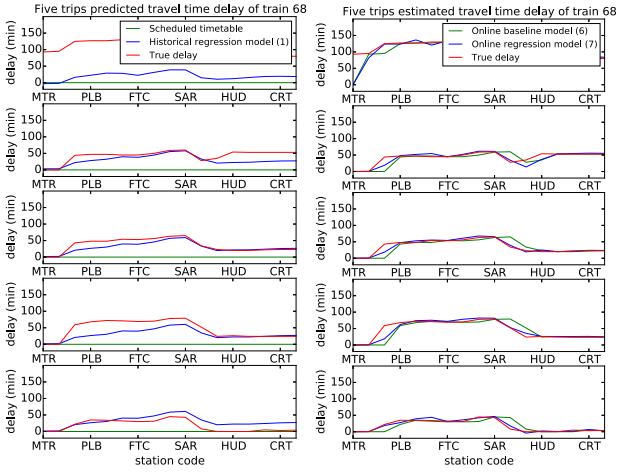
The ranked average RMSE of the proposed methods and the scheduled time table for each train are shown in Figure 2. The average RMSE over all trains for each predictor is summarized in Table I. The historical regression model has better estimation accuracy compared to the scheduled time table, since delays from the past trips are incorporated in the model. Both online algorithms perform significantly better than the historical model, because online delay information from the upstream station are used to estimate the delay for the downstream station. Moreover, the online regression model (7) performs better than the online baseline model (6), because it is able to incorporate the potential delay that may occur between the current station and the next station, by training the parameters a_m and c_k. In summary, compared to the scheduled time table, the historical regression model (2) improves the RMSE by 12%, and the online regression model (7) improved the RMSE by 60%.

Note it is not possible to compactly display the calibrated model parameters for each train and for each model within the space of this manuscript. However, to promote further development, all supporting source code for the developed estimation models are open source and are available for download [22]. In order to provide more details of how

Method	RMSE (min)	Improvement
Scheduled time table	19.4	N/A
Historical regression model	17.0	12%
Online baseline model	8.4	57%
Online regression model	7.7	60%

TABLE I: Average RMSE of the proposed methods. The improvement shown in the table are percentage improvements of proposed methods compared to the scheduled time table with respect to RMSE

the regression models perform, we again use train 68 as an example. The prediction results by the scheduled time table and historical model (2) of the first 5 of the 30 trips for train 68 are shown in Figure 3a, and the estimation results by online models (6) and (7) are shown in Figure 3b. Again, we can conclude the historical model performs better than the scheduled time table, and the online models can further improve the delay estimation accuracy compared to the historical model.



(a) Historical model and time table (b) Online models

Fig. 3: Five trips delay estimation results of train 68 (top to bottom). The left figure shows the results for the scheduled time table and the historical regression model. The right figure shows the results for the online regression models

E. Regression results with interactions among trains

Next, the online regression interacting model (3) is tested. The online interacting model (3) is compared with the online non–interacting model (7) to investigate if modeling the delay caused by interaction among trains may help to improve the estimation accuracy. It is found that the RMSE difference between the two models for most of the trains are less than 2%, and the average RMSE over all trains of the two models are computed as 7.42 and 7.40 min, respectively.

The online regression interacting model (3) tends to capture the knock–on delay effect by including the term Φ. However, the performance of these two models are very close. After an investigation on the trained parameters b_j, it is found the values of b_j are nonzero and they do influence the final estimation, however, it does not outperform the online regression model (7).

One possible explanation is as follows. Once a train is delayed at a station, it is observed that the delay will propagate for several stations. As a result, it is usually the case that both the front train and the following train are delayed for several consecutive stations on a trip. While it is true that the following train is delayed due to an interaction with the leading train, the online regression model (7) is able to capture this knock–on delay by modeling the delay propagation from its upstream station for all stations except the first one, where the delay is initiated. As a result, similar performance is found for the online regression model (7) and the online regression interacting model (3).

V. Conclusion and future work

This paper studies the passenger train travel time delay problem by using data driven approaches. A historical regression model is proposed to predict train delays before the current trip starts, and an online regression model with two variations are developed to estimate train delays using delay information from current trips recorded at upstream stations and other related trains. The proposed methods are tested with Amtrak passenger train data. Compared to the prediction based on the scheduled time table, the historical regression model (2) improves the RMSE of the delay estimation by 12%, and the online regression model (7) improves the RMSE by 60%.

This article is the first work that uses data driven approaches to study passenger train delays in the US. The paper shows standard regression models can significantly improve the travel time delay estimates compared to the scheduled time table even data are coarse and limited, which lays the foundation for the development of more sophisticated algorithms for passenger train delay estimation.

There are several aspects that this work can be further studied. First, in this work it is found the estimation accuracy does not improve when the delays caused by the interactions among trains are modeled. Further investigations are needed to study the knock-on delays for passenger trains. For example, other criteria can be developed to construct the interaction set $\Omega_{i,k,t}$. Second, in this work, the potential delay between the current and previous stations is captured by a single parameter. Various factors may contribute to delays, including geometry and weather. As a result, provided data is available, we can separately model these factors in the regression model to improve the performance of the method. Additionally, alternative data driven approaches such as nonlinear regression models [23] or calibration via robust least squares [24] can be investigated. Finally, the impact of rescheduling decisions on train delay prediction [25] should also be investigated.

References

[1] M. Dingler, A. Koenig, S. Sogin, and C. P. Barkan, "Determining the causes of train delay," in *AREMA Annual Conference Proceedings*, 2010.

[2] [Online]. Available: http://dixielandsoftware.net/Amtrak/status/StatusMaps/

[3] P. Murali, M. Dessouky, F. Ordóñez, and K. Palmer, "A delay estimation technique for single and double-track railroads," *Transportation Research Part E: Logistics and Transportation Review*, vol. 46, no. 4, pp. 483–495, 2010.

[4] O. Frank, "Two-way traffic on a single line of railway," *Operations Research*, vol. 14, no. 5, pp. 801–811, 1966.

[5] E. Petersen, "Over-the-road transit time for a single track railway," *Transportation Science*, vol. 8, no. 1, pp. 65–74, 1974.

[6] B. Chen and P. T. Harker, "Two moments estimation of the delay on single-track rail lines with scheduled traffic," *Transportation Science*, vol. 24, no. 4, pp. 261–275, 1990.

[7] M. Carey and A. Kwieciński, "Stochastic approximation to the effects of headways on knock-on delays of trains," *Transportation Research Part B: Methodological*, vol. 28, no. 4, pp. 251–267, 1994.

[8] S. Özekici and S. Şengör, "On a rail transportation model with scheduled services," *Transportation Science*, vol. 28, no. 3, pp. 246–255, 1994.

[9] A. Higgins and E. Kozan, "Modeling train delays in urban networks," *Transportation Science*, vol. 32, no. 4, pp. 346–357, 1998.

[10] J. Yuan, *Stochastic modelling of train delays and delay propagation in stations*. Eburon Uitgeverij BV, 2006.

[11] E. Petersen and A. Taylor, "A structured model for rail line simulation and optimization," *Transportation Science*, vol. 16, no. 2, pp. 192–206, 1982.

[12] M. M. Dessouky and R. C. Leachman, "A simulation modeling methodology for analyzing large complex rail networks," *Simulation*, vol. 65, no. 2, pp. 131–142, 1995.

[13] Q. Lu, M. Dessouky, and R. C. Leachman, "Modeling train movements through complex rail networks," *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, vol. 14, no. 1, pp. 48–75, 2004.

[14] S. F. Hallowell and P. T. Harker, "Predicting on-time line-haul performance in scheduled railroad operations," *Transportation Science*, vol. 30, no. 4, pp. 364–378, 1996.

[15] M. F. Gorman, "Statistical estimation of railroad congestion delay," *Transportation Research Part E: Logistics and Transportation Review*, vol. 45, no. 3, pp. 446–456, 2009.

[16] P. Kecman and R. M. Goverde, "Online data-driven adaptive prediction of train event times," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 1, pp. 465–474, 2015.

[17] I. A. Hansen, R. M. Goverde, and D. J. van der Meer, "Online train delay recognition and running time prediction," in *Proceedings of the IEEE Conference on Intelligent Transportation Systems*, 2010, pp. 1783–1788.

[18] [Online]. Available: http://tickets.amtrak.com/secure/content/routeatlas/index.html

[19] H. Lütkepohl, *New introduction to multiple time series analysis*. Springer Science & Business Media, 2007.

[20] E. R. Cook, *A time series analysis approach to tree ring standardization*. The University of Arizona, 1985.

[21] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, vol. 14, no. 2, 1995, pp. 1137–1145.

[22] [Online]. Available: https://github.com/Lab-Work/TrainDelayEstimation_IEEEITSC

[23] D. M. Bates and D. G. Watts, *Nonlinear regression: iterative estimation and linear approximations*. Wiley Online Library, 1988.

[24] L. El Ghaoui and H. Lebret, "Robust solutions to least-squares problems with uncertain data," *SIAM Journal on Matrix Analysis and Applications*, vol. 18, no. 4, pp. 1035–1064, 1997.

[25] P. Kecman, F. Corman, A. D'Ariano, and R. M. Goverde, "Rescheduling models for railway traffic management in large-scale networks," *Public Transport*, vol. 5, pp. 95–123, 2013.