Citation for the original published paper:

Fredriksson, H., Dahl, M., Holmgren, J. (2025)
Freeway Travel Time Estimation Using Sequential Link Regression Modeling
In: *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC* (pp. 303-308). Institute of Electrical and Electronics Engineers (IEEE)
IEEE International Conference on Intelligent Transportation Systems
https://doi.org/10.1109/ITSC60802.2025.11423489

Freeway travel time estimation using sequential link regression modeling

Henrik Fredriksson[1], Mattias Dahl[2], *Member, IEEE* and Johan Holmgren[3]

***Abstract*—Accurate travel time estimations are essential for traffic analysis and enable modern applications such as dynamic route guidance and traffic control. With the growing availability of high-resolution traffic data from GPS-enabled devices and probe vehicles, advanced models have been developed to estimate travel times more precisely. This paper proposes a sequential link estimation method for trip-level travel time estimation. The method exploits how travel times on one link are influenced by the preceding link and influence the subsequent link along a route. The method uses a chain of regression estimation models where each link's estimated travel time depends on the travel time of the adjacent link. Each estimated value is passed as input to the model for the next link, creating a chain of conditional estimates that extends from an arbitrary link to both the beginning and end of a freeway. We evaluate the proposed travel time estimation method using real-world traffic data from freeways in Sweden. The results show an average percentage error as low as 2.38 percent with a standard deviation of 1.88 percent, indicating highly accurate travel time estimates.**

***Index Terms*—Travel time estimation, regression prediction model, sequential link modeling**

I. INTRODUCTION

Travel times can be directly measured or predicted by estimating how long a vehicle will take to traverse a link or route. Accurate travel time estimation is a fundamental part of intelligent transportation systems (ITSs), and information about travel times has a broad range of applications in traffic management [1]. This includes dynamic route guidance, advanced traveler information systems, travel demand modeling, logistics, and infrastructure planning [2]–[5]. Travel times influence trip planning and travel mode choice for individual road users. They are commonly used to evaluate the performance and accessibility of road networks and are used in real-time traffic control and long-term performance monitoring.

In recent years, the amount of traffic data has steadily increased. GPS-enabled devices, e.g., mobile phones and vehicle probe data, have enabled more precise spatial and temporal information about vehicle movement. This large-scale traffic data comes from different sources, enabling different data-driven ways to estimate travel times [6], [7]. These data-driven models aim to inductively identify patterns in the data, learn, and estimate travel times based on identified patterns [7].

A key challenge in traffic prediction is capturing interdependencies between links in a road network. This includes spatial correlations between nearby links [8], [9]. Travel times on one link can directly affect the travel time on downstream links. Likewise, the travel time on one link can be a result of the condition on previous upstream links. For example, a slowdown due to congestion on one link can propagate and influence vehicle speeds on adjacent links.

This paper introduces a sequential link regression approach for trip travel time estimation. The proposed method is designed to model the dependencies between travel times along a sequence of connected links. The central idea is to use regression models to estimate the travel time on each link based on the travel time of the adjacent link. This estimated value is then passed as input to the model for the next link, creating a chain of conditional estimates that extends from an arbitrary link to the beginning and the end of a freeway. This step-by-step modeling captures the sequential nature of vehicle movement through a freeway and enables estimation of cumulative travel time. It should be noted that our approach emphasizes application and interpretability rather than proposing new machine-learning algorithms.

The remainder of the paper is organized as follows. Previous work on regression-based travel time estimation is described in Section II. The proposed methodology is described in Section III, where data processing and sequential link regression modeling are described. The results from a numerical study are presented in Section IV, and conclusions are presented in Section V.

II. PREVIOUS WORK

Data-driven regression methods use historical data, real-time travel time data, or both to estimate travel times. These methods can also incorporate other features such as traffic flow, density, and time of day to improve the estimation accuracy. Regression has been widely applied in the literature to estimate travel times [10]–[13].

Linear regression has been used to estimate bus travel times by modeling driving time, dwell time, and signal delays based on route characteristics [14], [15]. This approach provides an understanding of the factors contributing to travel time variability. Linear regression has also been effective in improving travel time estimates by correcting systematic

---

[1] Corresponding author. Henrik Fredriksson is with the Department of Science and Technology, Linköping University, Norrköping, Sweden, `henrik.fredriksson@liu.se`

[2] Mattias Dahl is with the Department of Mathematics and Natural Sciences, Blekinge Institute of Technology, Karlskrona, Sweden, `mattias.dahl@bth.se`

[3] Johan Holmgren is with the Department of Computer Science and Media Technology, Malmö University, Malmö, Sweden, `johan.holmgren@mau.se`

prediction biases [16]. In addition to modeling travel times at the link or intersection level, linear regression has been applied to estimate travel times between origin–destination (OD) pairs [17].

k-Nearest Neighbors (k-NN) regression techniques have been used to estimate virtual travel times from partial trip data [18]. This method has also been applied to estimate travel times by including features such as OD zones, time of day, and weather conditions [19]. The inclusion of these features enables k-NN models to capture spatiotemporal variability and external influences on travel behavior. Additionally, traffic state variables, i.e., travel speed, traffic density and flow, play a crucial role in characterizing roadway conditions, distinguishing between free-flow and congested states. These traffic states can be used as input features in regression models to further improve the accuracy of travel time estimations [20].

Ensemble methods improve the generalization of predictions by aggregating the outputs of multiple base estimators, all trained using the same learning algorithm, rather than relying on a single predictive model. Random forests and gradient-boosted trees are two commonly used ensemble techniques for estimating and analyzing travel times [21]–[23]. Random forests, for example, have been employed to estimate travel time percentiles, which serve as indicators of travel time reliability and variability [24], [25]. They have also been applied in network-wide settings to predict link-level travel times across different prediction horizons [26]. Gradient-boosted trees, on the other hand, have demonstrated effectiveness in incorporating a wide range of features—such as traffic state variables, time of day, day of the week, and weather conditions—to estimate travel times. These models have been used for estimating travel times on individual freeway links, for OD pairs, and entire trip trajectories [17], [27], [28].

These methods have strengths and weaknesses: linear regression is simple but limited in capturing nonlinearity; k-NN adapts locally but needs similar past cases; ensembles improve accuracy but add complexity.

III. METHODOLOGY

This section presents our travel time estimation method, which uses a single observation to estimate travel times on all links of a freeway. Regression-based models sequentially estimate travel times downstream (or upstream) by using the travel time from the adjacent link.

We consider the following notation. A freeway consists of L ordered links, indexed by $l = 1, 2, \dots, L$. A travel time observation $T_l^{\text{obs}}(t)$ describes the time it takes to traverse link l at timestamp t. An estimated travel time on link l at timestamp t is denoted by $\widehat{T}_l(t)$.

A. Data description and processing

This study is carried out on European roads in the Swedish road network. The data is provided by the Swedish Transport Administration (Trafikverket) and consists of average travel times for each link and minute. The average one-minute link travel times are used to create a dataset of "virtual trips". They may not represent real trips, but they effectively capture both downstream and upstream traffic conditions. Starting from a departure time t_1 at the first link, we retrieve how long it takes to traverse that link using the observed average travel time $T_1^{\text{obs}}(t_1)$ from the data. The time $t_2 = t_1 + \Delta T_1^{\text{obs}}(t_1)$ at which the vehicle exits the first link becomes the estimated arrival time into the second link. At that time t_2, we retrieve the corresponding travel time observation $T_2^{\text{obs}}(t_2)$ for the second link and calculate the arrival time to the third link by $t_3 = t_2 + \Delta T_2^{\text{obs}}(t_2)$. This process is repeated sequentially for each link, always using the travel time observation that corresponds to the estimated time of entry into that link. This generates a sequence

$$(T_1^{\text{obs}}(t_1), T_2^{\text{obs}}(t_2), \dots, T_L^{\text{obs}}(t_L)) \tag{1}$$

where

$$t_{l+1} = t_l + \Delta T_l^{\text{obs}}(t_l), \quad 1 \le l < L - 1. \tag{2}$$

Equation (1) describes travel times along all links of the freeway, conditional on a given departure time. The total travel time T for the entire freeway is then obtained by summing the travel times across all individual links by

$$T = \sum_{l=1}^{L} T_l^{\text{obs}}(t_l). \tag{3}$$

Fig. 1 shows a histogram of estimated travel times (3) from 7883 different departure times along one of the studied freeways. The histogram clearly shows a right-skewed distribution, which aligns with typical travel time behavior.

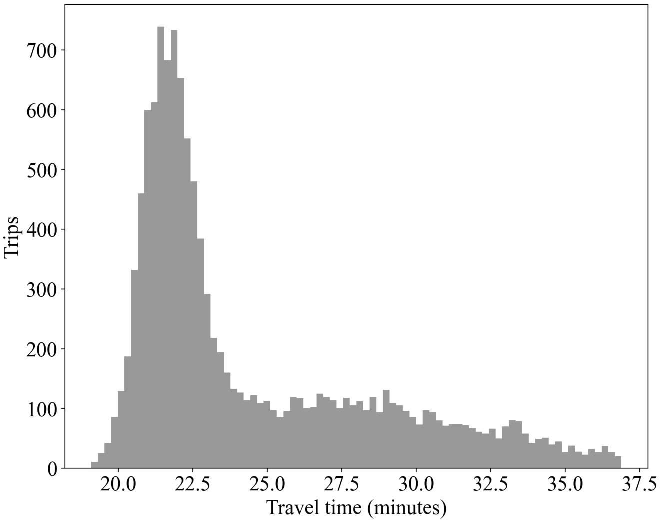


Fig. 1: Histogram of estimated travel times along one of the studied freeways.

B. Travel time correlation between adjacent links

Nearby links often share similar characteristics, and their travel times tend to exhibit strong spatio-temporal correlations [29]. To illustrate this, the heatmaps in Fig. 2 show the pairwise Pearson correlation values (based on timestamp and travel time) between consecutive links l and $l+1$ across a freeway of 44 links. The heatmap reveals statistically significant high correlation (p-values less than 0.01) between most pairs of adjacent links, supporting the motivation for our proposed approach. In contrast, the correlation drops significantly when considering non-adjacent links (e.g., l and $l+2$ or beyond), indicating that the influence of travel time decays rapidly with increasing distance. This observation further justifies the assumption that each link is primarily influenced by its immediate neighbors.

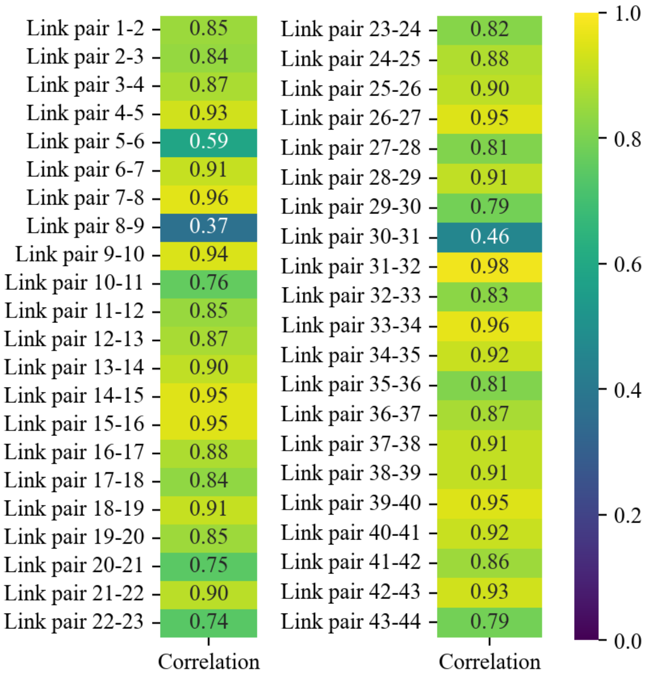


Fig. 2: Heatmap of travel time correlations for each pair of adjacent links along a freeway.

C. Sequential link travel time estimation using regression

The process begins with an observed or estimated travel time T_k^{obs} on some link k of the freeway. The travel time on the adjacent link l is estimated using a model

$$\widehat{T}_l = f_l(T_k^{\text{obs}}, \boldsymbol{\theta}_{k,l}), \tag{4}$$

where f_l is a regression model to determine the travel time link l, parameterized by $\boldsymbol{\theta}_{k,l}$. This feature vector contains historical travel time observations on links k and l and may also include derived temporal features (e.g., hour of day, day of week). In this work, we only considers travel time data from the immediately adjacent link. Travel time estimation proceeds recursively for all links through the entire freeway starting from the observed link k. For all downstream links, we compute the estimated travel time

$$\widehat{T}_{l+1} = f_{l+1}(\widehat{T}_l, \boldsymbol{\theta}_{l,l+1}), \quad k < l < L. \tag{5}$$

For upstream links, we use

$$\widehat{T}_{l-1} = f_{l-1}(\widehat{T}_l, \boldsymbol{\theta}_{l,l-1}), \quad 1 < l < k. \tag{6}$$

This recursive, regression-based approach allows a single observed travel time on an arbitrary link k to be propagated along the freeway, capturing both spatial and temporal dependencies between links, giving a sequence

$$(\widehat{T}_1, \widehat{T}_2, \ldots, \widehat{T}_L), \quad \widehat{T}_k = T_k^{\text{obs}} \tag{7}$$

of estimated travel times for each link with total travel time $\widehat{T}$. By modeling link-to-link transitions, the method captures temporal and spatial dependencies in travel times which makes it applicable in contexts where travel time variability is high—such as during peak periods.

Regression analysis is a statistical process used to model the relationship between a dependent variable and one or more independent variables. In prediction tasks, regression models learn patterns from historical data to estimate future values. In this paper, we consider linear regression, random forest regression, and gradient-boosted trees—eXtreme Gradient Boosting (XGBoost)—to estimate travel times on adjacent links. Linear regression models the relationship between a dependent variable and one or more independent variables. We use historical travel times and hours of day as input features to construct the regression models. The linear regression model is defined as:

$$\widehat{T}_l = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon, \tag{8}$$

where $\widehat{T}_l$ is the estimated travel time on a target link l, x_1 is the historical travel time on an adjacent link, x_2 is the hour of the day. All three models use these same two features as inputs. In 8, β_0 is the intercept, β_1, β_2 are regression coefficients, and ε is the error term. Random forest and gradient-boosted trees are ensemble-based. They combine the predictions of multiple decision trees to improve accuracy. Given a feature vector $\boldsymbol{x}$, the random forest estimate of travel time is defined as the average of predictions from an ensemble of decision trees:

$$\widehat{T}_{l,\text{RF}} = \frac{1}{M} \sum_{m=1}^{M} f_m(\boldsymbol{x}), \tag{9}$$

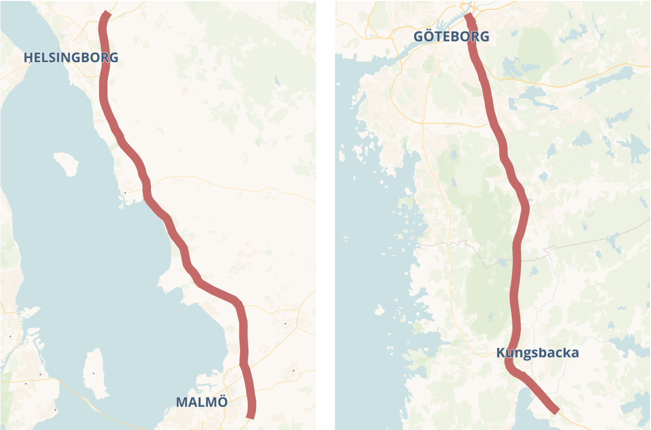
where f_m is the m-th decision tree and M is the total number of trees in the forest.

IV. NUMERICAL STUDY

A. Freeway descriptions

In this study, we consider two freeways in the Swedish road network. The freeways are shown in Fig. 3. These are the European road E6 between Malmö and Helsingborg (a) and the European road E6 between Gothenburg and Kungsbacka (b). We consider the travel direction towards Malmö and Gothenburg, respectively.

Using the process described in Section III-A, we generated trips with departure times with one-minute intervals during the morning peak period (06:00–10:00), from Monday to Thursday. The study covered a three-month period, from March 1 to May 31, 2023. Most days during this period are part of regular school and work weeks. For each freeway, we split the trip data set into a train data set (80% of the data) and a test data set (20% of the data).



(a) European road E6 between Malmö and Helsingborg. The freeway is 64 km long, has 59 links with a free-flow travel time of 34.2 minutes. The posted speed limit is 110 km/h for all links.

(b) European road E6 between Gothenburg and Kungsbacka. The freeway is 35.7 km long, has 49 links with a free-flow travel time of 24.5 minutes. The posted limit varies between 70 and 110 km/h.

Fig. 3: Freeways used in the numerical study.

TABLE I: Descriptive statistics of estimation errors (PE, MSE, MAE) for E6 Malmö, relative to the baseline.

	Statistic	PE (%)	MSE (seconds)	MAE (seconds)
Linear regression	Mean	3.19	3.70	1.90
	Std.	1.89	1.34	0.65
	Min	0.11	1.37	0.88
	25%	1.47	2.13	1.16
	50%	3.09	3.29	1.73
	75%	3.89	4.92	2.36
	Max	7.86	7.42	3.12
Random forest	Mean	13.99	11.76	7.08
	Std.	14.79	9.59	5.39
	Min	0.11	1.55	1.27
	25%	4.53	4.93	3.22
	50%	8.83	8.62	5.52
	75%	17.06	15.03	8.60
	Max	81.62	61.92	30.86
XGBoost	Mean	2.38	3.52	2.04
	Std.	1.88	1.11	0.47
	Min	0.05	2.17	1.36
	25%	0.97	2.91	1.66
	50%	1.69	3.38	1.88
	75%	3.09	3.97	2.36
	Max	10.85	12.32	4.82

TABLE II: Descriptive statistics of estimation errors (PE, MSE, MAE) for E6 Gothenburg, relative to the baseline.

	Statistic	PE (%)	MSE (seconds)	MAE (seconds)
Linear regression	Mean	8.05	5.58	3.20
	Std.	3.52	2.85	1.25
	Min	0.04	1.98	1.37
	25%	5.63	3.69	2.29
	50%	8.12	4.43	2.72
	75%	10.60	5.83	3.87
	Max	15.64	11.59	6.05
Random forest	Mean	23.83	28.90	11.82
	Std.	31.01	41.74	12.24
	Min	0.21	2.46	1.90
	25%	4.60	6.14	4.31
	50%	8.46	8.78	6.51
	75%	23.88	19.27	11.35
	Max	155.49	185.11	58.37
XGBoost	Mean	9.74	6.82	4.08
	Std.	4.01	2.41	1.04
	Min	0.26	2.98	2.15
	25%	6.94	5.00	3.26
	50%	9.69	6.16	3.82
	75%	12.24	7.68	5.07
	Max	17.32	11.27	6.03

B. *Results and analysis*

The proposed method aims to estimate travel times on all links using a single observation from one link. We evaluate the method using two approaches. The first approach compares the predicted travel time to the mean travel time for each link in the test set, which serves as a baseline. In this approach, we evaluate the percentage error

$$\mathrm{PE} = 100 \times \frac{\widehat{T} - \sum_{l=1}^{L} T_l^{\text{test}}}{\sum_{l=1}^{L} T_l^{\text{test}}} \tag{10}$$

between the baseline travel time and the estimated travel time. In Equation (10), T_l^{test} is mean travel time on link l in the test set. We also evaluate the mean squared error

$$\mathrm{MSE} = \frac{1}{L} \sum_{l=1}^{L} \left(\widehat{T}_l - T_l^{\text{test}} \right)^2 \tag{11}$$

and mean absolute error

$$\mathrm{MAE} = \frac{1}{L} \sum_{l=1}^{L} \left| \widehat{T}_l - T_l^{\text{test}} \right| \tag{12}$$

over each link between the baseline and estimated travel times.

We estimated the travel time for each link and hour (6,7,8,9) across both roads. The descriptive statistics of the estimation error are presented in Table I for E6 Malmö and in Table II for E6 Gothenburg.

Overall, the evaluation metrics are lower for E6 Malmö, suggesting more predictable travel times. This can likely be attributed to the homogeneity of the links, which share the same posted speed limit and number of lanes. XGBoost performs better on E6 Malmö than on E6 Gothenburg, though linear regression also yields fairly accurate estimates with significantly faster execution time. Random forest, on the other hand, tends to produce larger and more inconsistent errors—particularly on E6 Gothenburg, where the maximum PE is exceptionally high. Fig. 4, Fig. 5, and Fig. 6 show the estimated link travel times using linear regression, random forest, and gradient-boosted trees for the same observation link during a given hour and the baseline link travel times.

The second approach evaluates how well the method replicates actual trips in the test set. Specifically, it compares each estimated travel time to the most similar observed trip in the test set, based on the lowest MSE or MAE. The descriptive

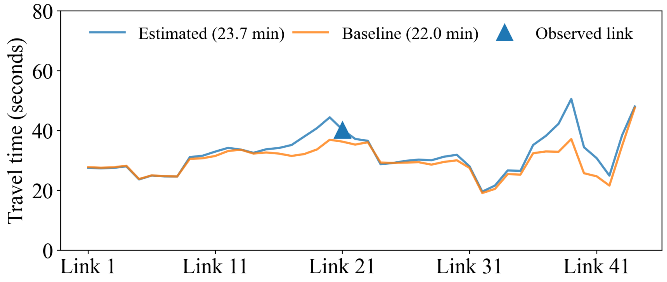


Fig. 4: Linear regression estimated- and baseline travel times for each link. The percentage error is 7.56%, MSE 3.8 seconds, and MAE 2.4 seconds.

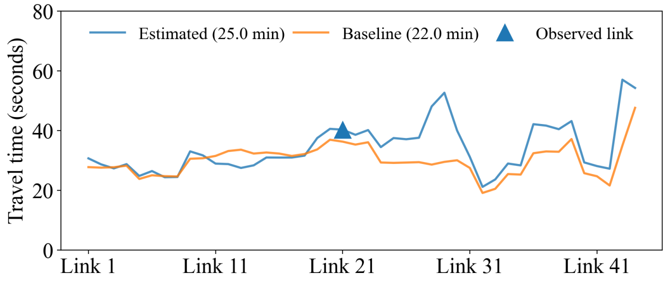


Fig. 5: Random forest regression estimated- and baseline travel times for each link. The percentage error is 13.49%, MSE 7.3 seconds, and MAE 5.0 seconds.

statistics of estimation errors for travel time estimates for each link and hour are given in Table III for E6 Malmö and Table IV for E6 Gothenburg.

Linear regression estimates tend to give the lowest error metrics for both E6 Malmö and E6 Gothenburg. Random forest is the least consistent, with the highest variability in estimation errors, particularly for E6 Gothenburg, where both the MSE and MAE have a wide range. One possible explanation is overfitting.

Fig. 7 and Fig. 8 show link travel time estimates from linear regression, using different observation links as input during a given hour. Each plot also highlights the test set trips that are most similar to the estimated travel time, based on the lowest MSE and MAE (in Fig. 8 these are the same).

V. CONCLUSIONS AND FUTURE WORK

In this paper, we study a sequential link modeling approach to estimate travel times along freeways. The core concept involves applying regression models to estimate travel time on individual links, using the travel time of the adjacent link. Each estimation feeds into the next, forming a sequential chain of conditional estimates on each link. This allows the estimation of an entire freeway from a single link travel time observation. Linear regression performed best on the E6 Gothenburg freeway, while XGBoost gave the highest accuracy on the E6 Malmö freeway in a Swedish case study. This indicates that both a straightforward linear model and a more complex ensemble can excel under different conditions. Future research includes extending and evaluating the method on road networks with heterogeneous roads, rather than single freeways, and examining how prediction errors propagate along link

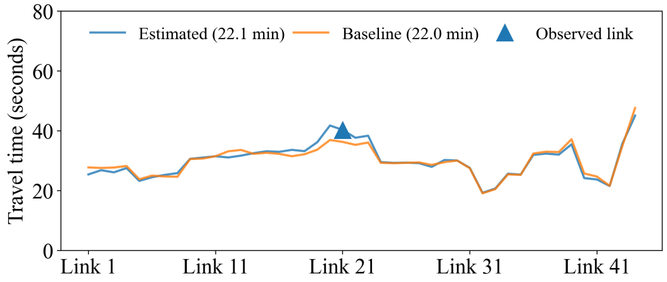


Fig. 6: Gradient-boosted tree estimated- and baseline travel times for each link. The percentage error is 0.4%, MSE 1.5 seconds, and MAE 1.03 seconds.

TABLE III: Descriptive statistics of estimation error metrics (MSE and MAE) for E6 Malmö based on the most similar trips in the test set, identified using the lowest MSE and MAE.

	Statistic	MSE (seconds)	MAE (seconds)
Linear regression	Mean	0.26	1.41
	Std.	0.04	0.18
	Min	0.19	1.13
	25%	0.25	1.34
	50%	0.25	1.36
	75%	0.26	1.41
	Max	0.44	2.54
Random forest	Mean	2.93	10.16
	Std.	4.29	11.64
	Min	0.27	1.59
	25%	1.44	4.63
	50%	1.99	7.33
	75%	3.02	11.54
	Max	22.21	69.45
XGBoost	Mean	0.32	1.69
	Std.	0.09	0.35
	Min	0.19	1.12
	25%	0.26	1.50
	50%	0.34	1.73
	75%	0.34	1.79
	Max	1.15	4.76

REFERENCES

[1] W. Zhang, "Freeway travel time estimation based on spot speed measurements," Ph.D. dissertation, Virginia Polytechnic Institute and State University, 2006.

[2] U. Mori, A. Mendiburu, M. Álvarez, and J. A. L. and, "A review of travel time estimation and forecasting for advanced traveller information systems," *Transportmetrica A: Transport Science*, vol. 11, no. 2, pp. 119–157, 2015.

[3] C. Anda, A. Erath, and P. J. F. and, "Transport modelling in the age of big data," *International Journal of Urban Sciences*, vol. 21, no. sup1, pp. 19–42, 2017.

[4] M. G. Cedillo-Campos, J. Piña-Barcenas, C. M. Pérez-González, and J. Mora-Vargas, "How to measure and monitor the transportation infrastructure contribution to logistics value of supply chains?" *Transport Policy*, vol. 120, pp. 120–129, 2022.

[5] C. Y. Chen, E. W. Sun, M.-F. Chang, and Y.-B. Lin, "Enhancing travel time prediction with deep learning on chronological and retrospective time order information of big traffic data," *Annals of Operations Research*, vol. 343, no. 3, pp. 1095–1128, 2024.

TABLE IV: Descriptive statistics of estimation error metrics (MSE and MAE) for E6 Gothenburg based on the most similar observed trips in the test set, identified using the lowest MSE and MAE.

	Statistic	MSE (seconds)	MAE (seconds)
Linear regression	Mean	0.32	1.99
	Std.	0.08	0.51
	Min	0.22	1.36
	25%	0.27	1.67
	50%	0.29	1.78
	75%	0.35	2.30
	Max	0.62	3.80
Random forest	Mean	0.64	3.88
	Std.	0.60	3.63
	Min	0.10	0.68
	25%	0.28	1.77
	50%	0.38	2.39
	75%	0.75	4.51
	Max	2.96	16.75
XGBoost	Mean	0.42	2.65
	Std.	0.07	0.38
	Min	0.26	1.60
	25%	0.36	2.34
	50%	0.45	2.76
	75%	0.47	2.83
	Max	0.61	3.85

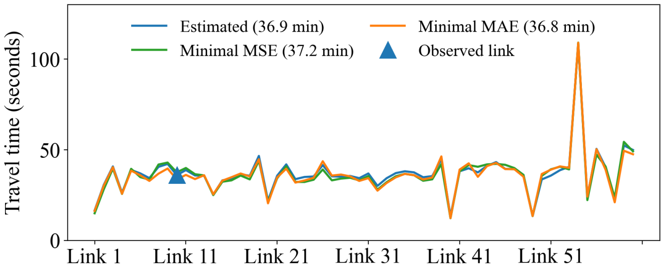


Fig. 7: Linear regression estimated travel times for each link and test set trip (same for MSE and MAE) with minimal MSE (0.21 seconds) and MAE (1.33 seconds).

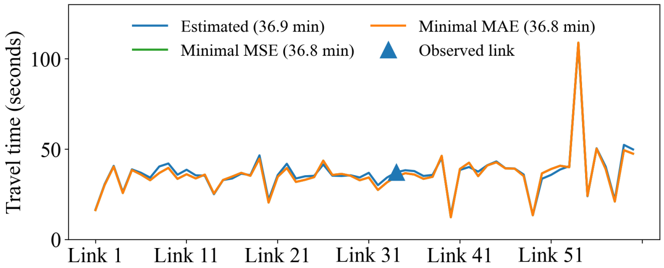


Fig. 8: Linear regression estimated travel times for each link and test set trips with minimal MSE (0.21 seconds) and MAE (1.33 seconds).

[6] D. Bertsimas, A. Delarue, P. Jaillet, and S. Martin, “Travel time estimation in the age of big data,” *Operations Research*, vol. 67, no. 2, pp. 498–515, 2019.

[7] H. Yuan and G. Li, “A survey of traffic prediction: from spatio-temporal data to intelligent transportation,” *Data Science and Engineering*, vol. 6, pp. 63–85, 2021.

[8] L. Bai, L. Yao, S. Kanhere, X. Wang, and Q. Sheng, “Stg2seq: Spatial-temporal graph to sequence model for multi-step passenger demand forecasting,” in *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI-19)*, 2019.

[9] M. Gmira, M. Gendreau, A. Lodi, and J.-Y. Potvin, “Travel speed prediction based on learning methods for home delivery,” *EURO Journal on Transportation and Logistics*, vol. 9, no. 4, 2020, Article 100006.

[10] X. Zhang, J. Rice, and P. Bickel, “Empirical comparison of travel time estimation methods,” University of California, Berkeley, Tech. Rep., 1999.

[11] J. Mendes-Moreira, A. M. Jorge, J. F. De Sousa, and C. Soares, “Comparing state-of-the-art regression methods for long term travel time prediction,” *Intelligent Data Analysis*, vol. 16, no. 3, pp. 427–449, 2012.

[12] Z. Wang, K. Fu, and J. Ye, “Learning to estimate the travel time,” in *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, 2018, pp. 858–866.

[13] P.-Y. Ting, T. Wada, Y.-L. Chiu, M.-T. Sun, K. Sakai, W.-S. Ku, A. A.-K. Jeng, and J.-S. Hwu, “Freeway travel time prediction using deep hybrid model – taking Sun Yat-sen freeway as an example,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 8257–8266, 2020.

[14] B. Yu, W. H. Lam, and M. L. Tam, “Bus arrival time prediction at bus stop with multiple routes,” *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 6, pp. 1157–1170, 2011.

[15] Y. Z. Farid, E. Christofa, and L. Paget-Seekins, “Estimation of short-term bus travel time by using low-resolution automated vehicle location data,” *Transportation Research Record*, vol. 2539, pp. 113–118, 2016.

[16] D. Laoide-Kemp and M. O’Mahony, “Dealing with latency effects in travel time prediction on motorways,” *Transportation Engineering*, vol. 2, 2020, Article 100009.

[17] H. Yuan, G. Li, Z. Bao, and L. Feng, “Effective travel time estimation: When historical trajectories over road networks matter,” in *Proceedings of the 2020 acm sigmod international conference on management of data*, 2020, pp. 2135–2149.

[18] Z. Li, R. Kluger, X. Hu, Y.-J. Wu, and X. Zhu, “Reconstructing vehicle trajectories to support travel time estimation,” *Transportation research record*, vol. 2672, no. 42, pp. 148–158, 2018.

[19] R. Kosasih, A. Fahrurozi, T. Handhika, I. Sari, D. Lestari *et al.*, “Travel time estimation for destination in Bali using KNN-regression method with Tensorflow,” in *IOP Conference Series: Materials Science and Engineering*, vol. 854, no. 1. IOP Publishing, 2020, p. 012061.

[20] J. Cheng, G. Li, and X. Chen, “Developing a travel time estimation method of freeway based on floating car using random forests,” *Journal of Advanced Transportation*, vol. 2019, no. 1, 2019.

[21] H. Taghipour, A. B. Parsa, and A. K. Mohammadian, “A dynamic approach to predict travel time in real time using data driven techniques and comprehensive data sources,” *Transportation Engineering*, vol. 2, 2020, Article 100025.

[22] H. Huang, M. Pouls, A. Meyer, and M. Pauly, “Travel time prediction using tree-based ensembles,” in *Computational Logistics*, E. Lalla-Ruiz, M. Mes, and S. Voß, Eds. Cham: Springer International Publishing, 2020, pp. 412–427.

[23] Y. Zhang and A. Haghani, “A gradient boosting method to improve travel time prediction,” *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 308–324, 2015.

[24] M. Zhao, X. Zhang, J. Appiah, and M. D. Fontaine, “Travel time reliability prediction using random forests,” *Transportation Research Record*, vol. 2678, no. 3, pp. 531–545, 2024.

[25] B. Anil Kumar, G. Chandana, and L. Vanajakshi, “Travel time reliability prediction using quantile random forest regression,” *Transportation in Developing Economies*, vol. 11, no. 1, p. 19, 2025.

[26] Y. Hou, P. Edara, and Y. Chang, “Road network state estimation using random forest ensemble learning,” in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, 2017, pp. 1–6.

[27] Z. Chen and W. Fan, “A freeway travel time prediction method based on an xgboost model,” *Sustainability*, vol. 13, no. 15, 2021.

[28] M. A. Abid, F. Lorig, J. Holmgren, and J. Petersson, “Ambulance travel time estimation using spatiotemporal data,” *Procedia Computer Science*, vol. 238, pp. 265–272, 2024.

[29] F. Alrukaibi, R. Alsaleh, and T. Sayed, “Applying machine learning and statistical approaches for travel time estimation in partial network coverage,” *Sustainability*, vol. 11, no. 14, 2019.