



IMPROVING PUBLIC TRANSPORT PREDICTION ACCURACY THROUGH MULTI-SOURCE DATA INTEGRATION

A CASE STUDY OF LJUBLJANA

Robert Medvedec¹, Špela Verovšek¹, Miha Moškon²

¹ University of Ljubljana, Faculty of Architecture, Slovenia

² University of Ljubljana, Faculty of Computer and Information Science, Slovenia

KEYWORDS

Public transport,
Travel time,
Congestion,
Machine learning,
LOOCV;

ABSTRACT

This study proposes a data acquisition, preprocessing and machine learning pipeline to enhance the accuracy of travel time estimates for the public bus network in Ljubljana, Slovenia. Existing prediction services often rely on vehicle location and may insufficiently capture dynamic congestion and contextual conditions. Using data from five Ljubljana Passenger Transport (LPP) routes, we evaluate LPP and Google Routes estimates against observed travel times. We develop a multisource model combining real-time variables with contextual predictors, including bus line, time of day, day type, and weather. The approach demonstrates how multisource integration of local transit feeds, external routing, and environmental indicators can improve prediction accuracy, reliability, and overall performance.

1. Introduction

Reliable travel-time information is an important component of public transport quality affecting journey planning, perceived consistency, and system usability. From the passenger perspective, real-time information extends beyond short-horizon vehicle-arrival predictions at boarding stops, commonly provided through real-time passenger-information displays and journey-planning applications; over longer prediction horizons it also supports estimation of arrival at the final destination and the temporal coordination of transfers and daily activities (van Ardenne et al., 2025). Accurate predictions allow passengers to allocate their time more efficiently by reducing unnecessary waiting and uncertainty associated with travel. This function has gained importance and relevance as journey-planning applications routinely integrate estimated arrival times into scheduling decisions, making prediction accuracy a critical component of system accessibility and modal competitiveness (Liu, et al. 2023; Henriquez-Jara et al., 2025).

For bus services, however, accurate predictions remain particularly challenging. Unlike rail-based modes operating on dedicated infrastructure, buses share road space with general traffic and are therefore directly exposed to congestion variations and dynamic road conditions (Parslov et al., 2023; Serin et al., 2022). Travel times can consequently vary considerably even along the same line and at the same scheduled service frequency. Such variability affects both prediction accuracy and the reliability experienced by passengers. Travel-time uncertainty is particularly relevant in congested urban environments, where unreliable journey times can influence passenger satisfaction, mode choice, and public transport use (Chen et al., 2026).

Bus travel-time prediction has increasingly been addressed using data-driven and machine-learning approaches that combine historical and real-time information, including Random Forest, support vector machine, and neural-network models (e.g., Xu & Ying, 2017; Yu et al., 2018; Parslov et al., 2023; Serin et al., 2022). Comparative studies across different cities show that prediction accuracy is closely related to route-level journey-time variability, with more consistent routes allowing simpler, less computationally demanding models to perform comparably to more complex approaches (Dunne et al., 2023). The challenge of high variability has also been addressed explicitly in bus travel-time prediction, with Bachu et al. (2021) applying support vector machines to improve prediction under heterogeneous and highly variable traffic conditions. Other work has examined which temporal, and contextual features contribute most to accuracy, showing that explicitly modeling temporal correlation can reduce error beyond standard feature sets (Kandiri et al., 2024). Deep-learning approaches have also been applied to short-horizon bus arrival prediction using vehicle-location, route, and temporal information (Rashvand et al., 2024). At the same time, passengers and operators increasingly rely on travel-time estimates already provided by routing platforms and local transport services. Whether these existing estimates themselves can serve as inputs to a predictive model and be complemented by temporal and contextual information to improve accuracy remains comparatively unexplored. This raises the possibility that existing operational estimates may be further improved through integration with additional predictors, without replacing the underlying estimation approach.

This study addresses this question using the bus,public transport network in Ljubljana, Slovenia. We first evaluate travel-time estimates provided by the local public transport service (LPP) and Google Routes against observed bus travel times, after which we examine whether combining these approximations with temporal and weather-related variables in a Random Forest model can reduce prediction error. In the study we therefore consider both the performance of existing travel- time estimation services and the potential added value of integrating multiple real-time and contextual predictors.

Ljubljana is a medium-sized European city with 302.826 inhabitants and a population density of around 1100 inhabitants/km² (SURS, 2026). Due to its importance as the nation's capital, it attracts a high number of daily commuters, which often puts additional pressure on the city's traffic infrastructure. The historical center has been transformed into a large pedestrian zone over the years with limited vehicular access, while traffic flows radially from the center to the outskirts. Ljubljana's public transport system is primarily bus-based, and its network largely reflects this spatial pattern through predominantly diametral routes that pass through the city center and connect opposite parts of the city (Bole, 2004; Trček & Rijavec, 2026). Heavy rail provides an additional transport layer but primarily serves suburban and regional connections; its limited station coverage within the city restricts its role in everyday urban travel. The mobility patterns remain strongly car oriented. Among residents of MOL (*Mestna občina Ljubljana*; City Municipality of Ljubljana), the modal split is 44.2% car, 33.8% walking, 12.7% cycling, 8.0% public transport, and 1.3% other modes. In the surrounding municipalities, the corresponding shares are 55.9% car, 32.9% walking, 4.0% cycling, 6.0% public transport, and 1.2% other modes (Trček & Rijavec, 2026, p. 2). These figures indicate a strong reliance on private cars, particularly in the surrounding municipalities, contributing to congestion and less predictable travel times in the city area.

The combination of substantial commuting flows, a predominantly bus-based public transport system, and exposure of buses to general road traffic make Ljubljana a relevant setting for examining travel-time variability and prediction performance.

2. Methods

2.1 Study design and data sources

LPP bus network serves Ljubljana and the surrounding municipalities through 49 daily lines, including city and integrated suburban services. We selected five city bus lines for the analysis: four principal city lines (1, 2, 5, and 7), each operating at least three trips per hour during peak periods of a day, and Line 13, which operates at up to two trips per hour (Ljubljanski potniški promet, 2025). While selected routes represent different route types, with both central and suburban characters, all are traversing through the city center. The selection focused on common and widely used lines that are strongly exposed to peak-hour traffic conditions. Repeated observations across these routes allowed the identification of temporal variations and time periods in which traffic most strongly affects public transport performance. The routes were analyzed unidirectionally. Selected route segments were excluded in cases where incorrectly geocoded bus stops produced unrealistic travel-time estimates by introducing additional walking time. Table 1 summarizes the selected lines and route segments and Figure 1 shows their spatial coverage within the wider LPP network.

We collected data from two main sources – Google Routes API and Ljubljana Passenger Transport (LPP) internal service. Google Routes API is a publicly available service, also used by Google Maps, which allows for quick and easy trip time approximation. Trips can be customized using a variety of parameters which allow us to adjust it to our case scenario. Google Routes API also uses real-time GTFS data from LPP to pinpoint each individual bus at all times, therefore adding an additional layer of accuracy to its predictions. LPP internal service is a semi-public API which requires a token that can be obtained for research purposes from the City of Ljubljana (Ljubljanski potniški promet, 2026). LPP API provides access to General Transit Feed Specification (GTFS) data, as well as information regarding individual buses, stations, routes, and timetables.

Data collection was conducted over a three-week period, from May 4 to May 24, 2026, with observations recorded daily between 04:00 and 24:00. The period was selected to exclude national holidays and major public events and thereby capture predominantly regular traffic and

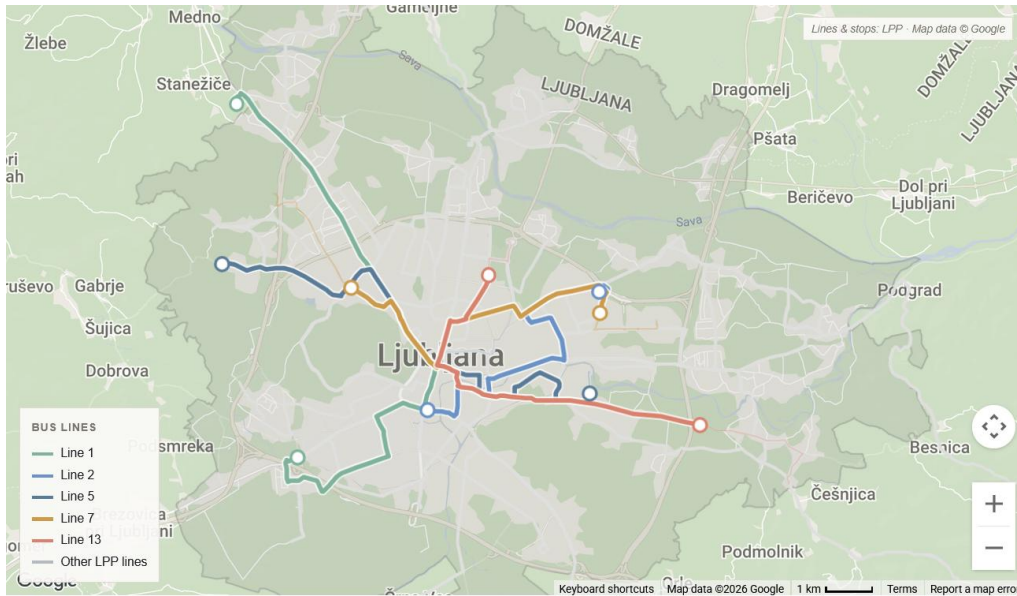
public transport conditions. Hourly historical weather data for Ljubljana was additionally retrieved from Open-Meteo Historical Weather API (Open-Meteo, 2026).

Table 1. Selected bus lines and route segments used in the analysis

	Route name	Origin station (ordinal number)	Destination station (ordinal number)	LPP line coverage (%)	Average observed route time (min)
Line 1	STANEŽIČE P+R - DOLGI MOST P+R	STANEŽIČE P+R (1)	DOLGI MOST P+R (29)	100%	47
Line 2	ZELENA JAMA - NOVE JARŠE	KRIŽANKE (11)	NOVE JARŠE (28)	64%	26
Line 5	ŠTEPANJSKO NASELJE - PODUTIK	NUSDORFERJEVA (2)	PODUTIK (23)	96%	35
Line 7	LETALIŠKA P+R - PRŽAN	BRATISLAVSKA (4)	ZGORNJA ŠIŠKA (21)	63%	31
Line 13	CENTER STOŽICE P+R - SOSTRO	KARDELJEVA PLOSCAD (3)	DOBRUNJE (20)	69%	29

Source: LPP, 2026.

Figure 1. Selected bus routes in Ljubljana. Bold color: study line segments; faded color: same line, off segment; grey color: remaining LPP network for reference. Interactive map available [here](#).



Source(s): Basemap - CARTO Positron, © OpenStreetMap contributors, © CARTO.

2.2 Data collection workflow and travel-time derivation

The data retrieval process was automated in Python, including API querying, error handling, and initial data filtering and preprocessing. Google Routes and LPP APIs were queried simultaneously at three-minute intervals, resulting in approximately 2,000 requests per day and around 40,000 requests over the three-week study period. Reducing the sampling interval to one minute would have increased API expenditure without yielding a significant reduction in the

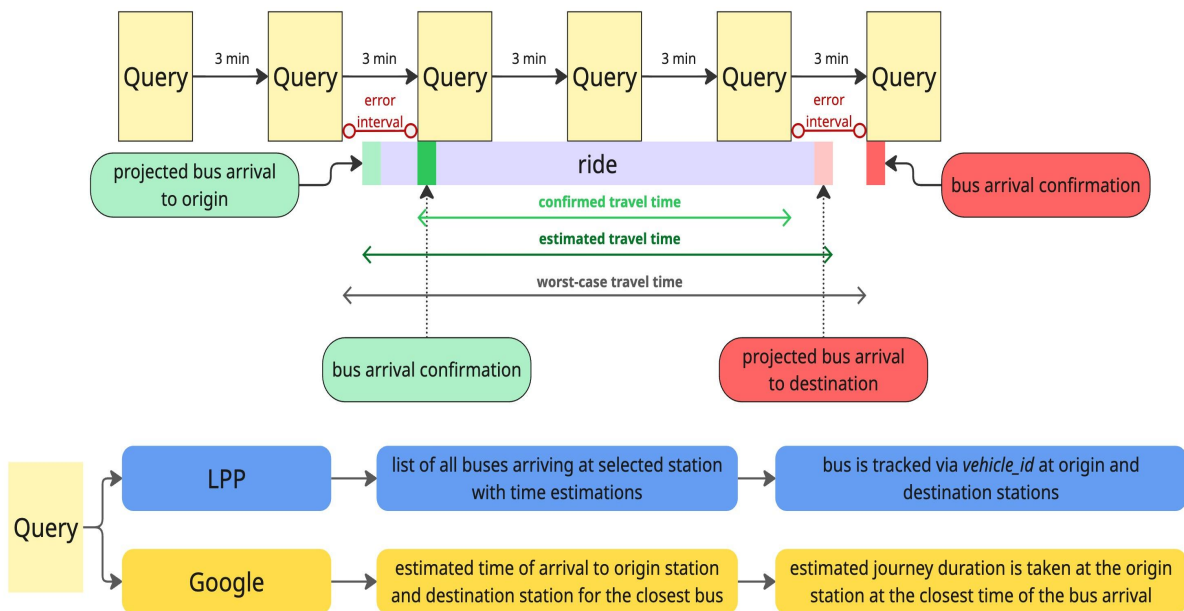
margin of error. Therefore, the three-minute interval provided an optimal balance of data granularity and resource efficiency. To contextualize the transit data, meteorological conditions were retrieved retroactively at one-hour intervals (Open-Meteo, 2026).

The two data sources were queried and processed as described below, with observed travel time derived from consecutive LPP vehicle observations at the origin and destination stops as an approximation of the realized travel time.

LPP querying. Using the publicly available LPP API, specific transit stations (station-code) and individual vehicles (vehicle-id) were continuously monitored. Each LPP API query of the station returned the following information: a) buses currently at the station; b) buses arriving at the station in the next 60 minutes; c) estimated time of arrival for incoming buses. These observations were used both to derive approximations of the observed bus arrival times and to obtain LPP travel time predictions. Observed travel time was calculated as the elapsed duration between a vehicle's arrival at the origin station and its subsequent arrival at the destination stop. As LPP API was queried at three-minute intervals, arrival times were approximated from consecutive observations. If by the time the next query was made, the bus was not at the station and not in the queue for arriving buses, it signified that the bus had already left the station. The most recent approximate time from LPP service was then used to calculate the exact moment at which the bus would reach the station. The same procedure was applied at the destination, allowing the origin-to-destination travel time to be derived. In order to obtain the LPP API prediction value (LPP estimated time of arrival, ETA), which is separate from the calculation of observed travel times via LPP API, an API call was made to the destination station at the exact moment of the bus arrival to the origin station (with previously defined limitations). With the use of vehicle-id parameter to map the desired vehicle, arrival approximation value was provided by LPP API.

Google Routes querying. For the same origin–destination stop pairs, the Google Routes API provided a direct estimate of public-transport travel time between the origin and destination stops. Unlike the LPP procedure, no additional derivation from consecutive vehicle observations was therefore required. The data collection workflow for both APIs is shown in Figure 2.

Figure 2. The data collection workflows using LPP and Google Routes APIs.



Source: Own elaboration, 2026.

2.3 Models for travel time approximation and their validation

The primary aim of this study was to assess the predictive accuracy of travel-time models currently used for public transport and to develop an alternative with improved performance. The complete pipeline of models' development and their assessment is described in this section.

2.3.1 Features and labels

The observed origin-to-destination travel time derived from the LPP observations described in Section 2.2 was selected as a target variable (label). Predictors as features in our model included bus line ID, hour of day, day of week, weather category, and the corresponding Google Routes and LPP travel-time predictions. The latter two labels provided real-time predictive information in addition to the temporal and contextual predictors.

2.3.2 Data preprocessing

Each observation was assigned to one of the five bus lines analyzed (1, 2, 5, 7, or 13). Days were grouped into weekdays (Monday–Friday) and weekends (Saturday–Sunday). No national holidays occurred during the data collection period. Weather was included as a contextual predictor to capture variation in traffic conditions and travel times that may not be explained by the temporal variables and API predictions alone.

Weather conditions were classified into three categories: good, mild, and adverse. Each day was assigned to a weather group based on average weather during that specific day, aggregated from the historical data which provided hourly weather accuracy. 'Adverse' weather was defined by low ($\leq 8^{\circ}\text{C}$) or high ($\geq 28^{\circ}\text{C}$) thresholds, precipitation ($\geq 1\text{mm}$), or high wind speeds ($\geq 30\text{km/h}$); 'good' weather was defined by no precipitation, moderate temperatures ($> 15^{\circ}$ and $< 26^{\circ}$), and low wind speeds ($< 20\text{km/h}$); all remaining conditions were classified as 'mild' weather. The classification and thresholds were informed by previous transport-weather studies addressing weather-related travel behavior (Hansson et al., 2022; Gössling et al., 2023; Schmitt et al., 2023) and were additionally adapted to typical May climatological conditions in Ljubljana (Mestna občina Ljubljana, 2022; ARSO, 2025).

Trips that did not reach the destination (e.g. buses failing to reach their destination due to mechanical malfunctions), as well as observations involving unmapped route deviations, were excluded during preprocessing. Completed trips with the observed travel time that exceeded the expected journey time by more than double were also excluded. No missing values remained in the resulting dataset; therefore, imputation was not required. Feature scaling was not applied because the selected models did not require it.

2.3.3 Regression models

Two different models were employed in this study. Firstly, an initial model created with a dummy regressor (scikit-learn developers, n.d.) predicted the output as the mean target value of training instances having the same input feature values as the query instance. Furthermore, a random forest regression model was applied to enhance prediction accuracy.

Random forest regression from 'scikit-learn' Python library (scikit-learn developers, n.d.) was used with the default value of 100 estimators (total number of individual decision trees). Increasing the number of estimators to 1000 did not provide statistically different results. Final models were trained with all features sets available (see Section 2.3.1).

2.3.4 Model validation

Leave-One-Out Cross-Validation (LOOCV) is a common approach when analyzing smaller datasets and trying to use as much data as possible to train the model, while keeping the remaining for validation. The data collected in our case has a temporal structure. We thus

employed a leave-one-day-out strategy, where all observations from a single day were held out for validation and the model was trained on the remaining days. This approach preserves the chronological dependencies within each day, prevents information leakage between training and validation sets, and provides a more realistic assessment of the model's ability to generalize to unseen future days. Since our use case featured two specific day groups, weekdays and weekends, LOOCV was performed separately, under the assumption that these two specific day groups provided vastly different traffic flows.

2.3.5 Metrics for prediction quality estimation

The results from all leave-one-day-out iterations were averaged to obtain the final evaluation metrics, namely Mean Absolute Error (MAE), Mean Squared Error (MSE), and Normalized Mean Absolute Error (nMAE) were used to assess the predictive quality of the trained model on the testing data. These metrics allow us to summarize the overall performance of the model by using aggregated values of all iterations, with each iteration having the same weighted value in the result. While MAE treats all errors proportionally, MSE squares the errors before calculating the average, putting more emphasis on the outliers in the dataset. Two models can have the same MAE but much different MSE, meaning that the one with the lower MSE value will have a more consistent error. On the other hand, nMAE helps with comparison of MAE values across different scales, as average trip duration is different for each individual route.

3. Results and discussion

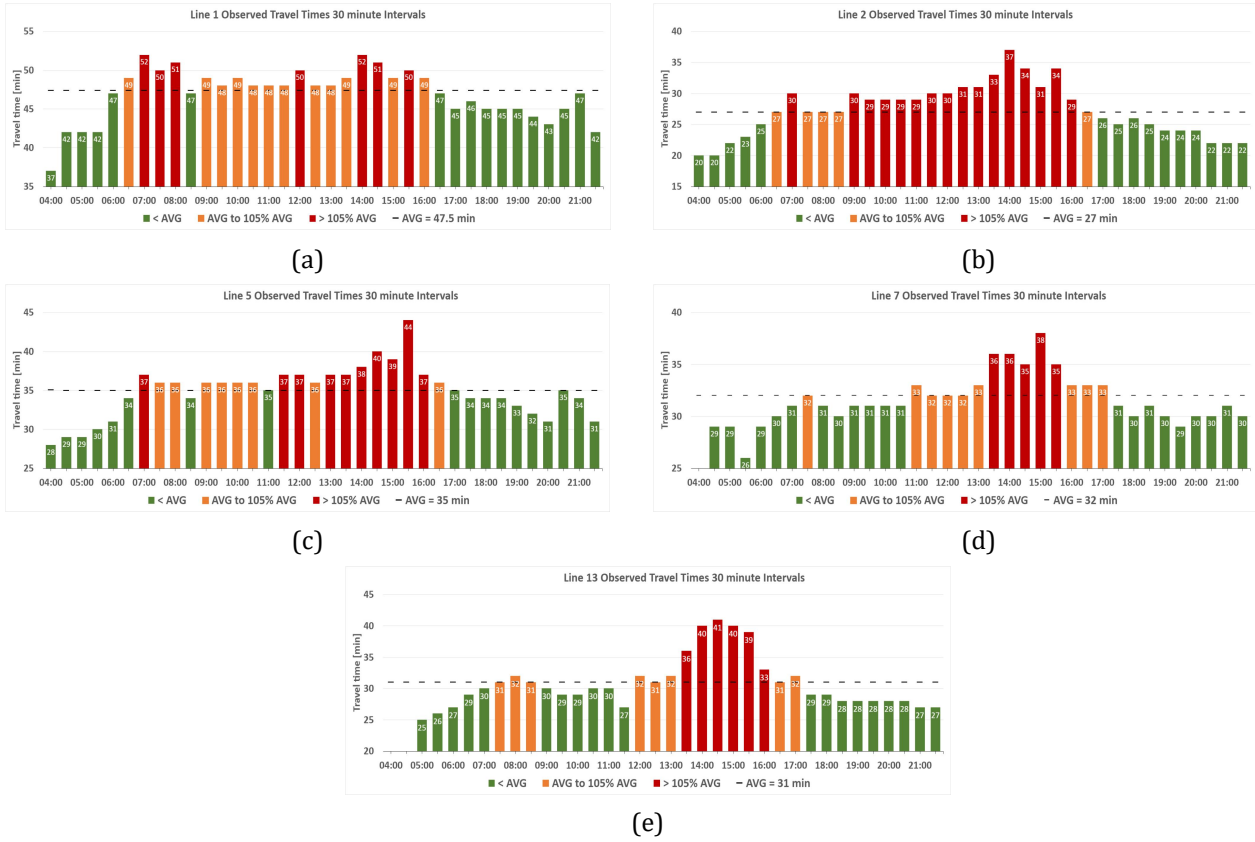
3.1 Exploratory data analysis

After data set cleanup, the final dataset comprised 4,702 individual datapoints (1,244 for Line 1; 913 for Line 2; 1,022 for Line 5; 948 for Line 7; 575 for Line 13). 87% of trips account for weekday travel, with the remaining 13% accounting for weekend travel.

Firstly, we used the collected data to determine key characteristics of congestion that affects public transport travel times in Ljubljana. Data was analyzed in three phases: individual datapoints, 30-minute intervals, and 60-minute intervals, which allowed us to get more accurate results on a few different levels. Analyses were conducted separately for weekdays and weekends, with weather conditions considered as an additional explanatory factor. A direct way to identify specific congestion intervals is to compare average observed travel times across 30-minute intervals which is presented in Figure 3.

As expected, the largest deviations from the average travel times occur during the afternoon peak period (14:00 to 17:00) across all lines. In addition, Lines 1, 2, and 5 exhibit substantially longer travel times during the morning peak period (07:00 to 09:00). The difference between the shortest and longest average travel times ranges from 41% on Line 1 to 85% on Line 2.

Figure 3. Distributions of average travel times throughout the day for each route. Red bars represent 30-minute intervals with more than 105%, yellow bars with 100-105%, and green bars intervals below 100% of the average travel time.



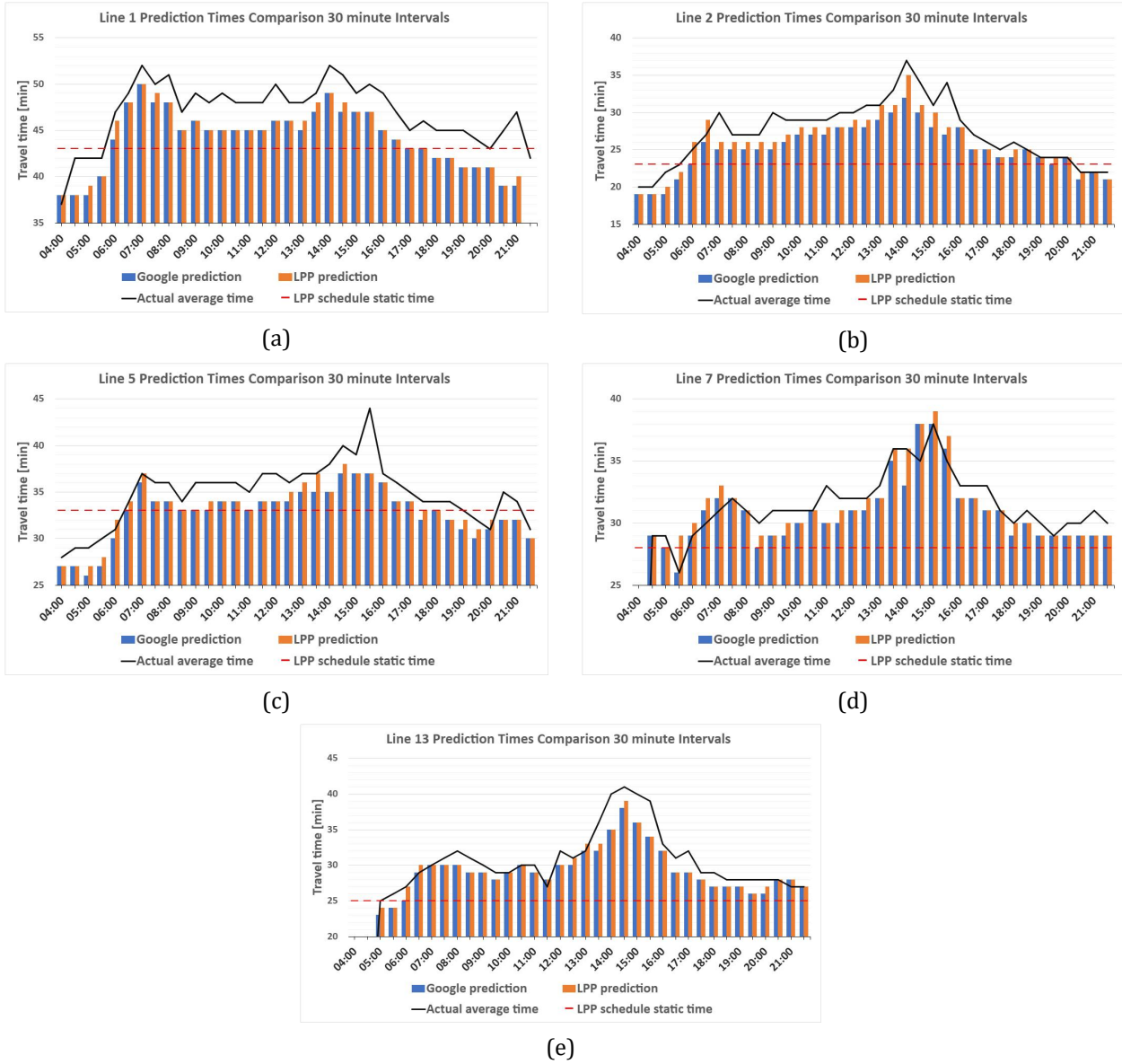
Sources: Own elaboration calculated from LPP API data, 2026.

These discrepancies indicate that the public transport system is strongly influenced by road traffic conditions, particularly on Line 2, one of the network's busiest routes. As a result, travel times are highly variable, making accurate predictions and reliable journey planning challenging. Figure 4 compares travel time estimates from Google Routes API (blue bars), LPP API (orange bars), and observed travel times (black line). The red line represents the static timetable-based prediction provided by LPP, which passengers commonly use as a reference estimate of travel duration; it serves as the public transport equivalent of the driving free-flow reference.

The results show a considerable temporal variation in observed travel times, especially during peak periods. While both prediction services partially capture these fluctuations, their performance varies across routes and times of day.

Both Google and LPP predictions follow travel time trends in a similar manner, which is expected given that Google Routes builds on LPP service data while applying its own routing and prediction procedures (Google, 2024). On the other hand, both show significant prediction errors, especially during rush hours on lines 1, 2, and 5.

Figure 4. Distributions of predicted travel times in comparison with observed travel times throughout the day for each route. Google Routes API (blue bars), LPP API (orange bars), observed travel times (black line), and static LPP prediction (red line).



Sources: Own elaboration, Google Routes API and LPP API data, 2026.

3.2. Analysis of travel time prediction services

Data collected during the data collection phase allowed us to assess the accuracy of the prediction models used by the Google Routes API and the LPP API. While the technical details of these models are not publicly available, travel time prediction systems generally rely on a combination of historical observations and current operating conditions (Lau, 2020). Therefore, both sources of information are assumed to contribute to the travel duration estimates analyzed in this study.

Services' prediction accuracy was evaluated against the observed travel times derived from LPP vehicle observations, following the procedure described in Section 2.2 and illustrated in Figure 2. Aggregate results were computed by combining observations from all bus lines and day

types, with each category weighted according to the number of available observations. Because weekday observations substantially outnumber weekend observations, the aggregate metrics are more strongly influenced by weekday travel patterns. The performance results for Google Routes API and LPP API are presented in Tables 2 and 3, respectively.

Table 2. Google Routes API travel prediction time error with regards to observed travel times. Cumulative values combine weekday and weekend predictions weighted by the number of trips

	Line 1	Line 2	Line 5	Line 7	Line 13	All lines
Weekdays	4.13	3.55	2.91	2.43	2.64	3.25
MAE & MSE	32.14	29.37	24.93	15.46	22.54	25.67
Weekends	3.32	2.27	2.72	2.60	2.06	2.63
MAE & MSE	22.05	9.21	21.53	12.58	8.65	15.42
Cumulative	4.03	3.23	2.87	2.47	2.56	3.13
MAE & MSE	30.87	24.31	24.12	14.79	20.80	23.66

Source: Own elaboration, 2026.

Table 3. LPP API travel prediction time error with regards to observed travel times. Cumulative values combine weekday and weekend predictions weighted by the number of trips.

	Line 1	Line 2	Line 5	Line 7	Line 13	All lines
Weekdays	3.98	3.15	2.76	2.49	2.56	3.10
MAE & MSE	30.64	25.12	23.62	15.54	22.06	24.15
Weekends	3.29	2.22	2.64	2.48	1.99	2.56
MAE & MSE	21.81	8.94	20.96	11.94	7.88	14.95
Cumulative	3.89	2.92	2.73	2.49	2.49	3.00
MAE & MSE	29.53	21.06	22.98	14.70	20.28	22.34

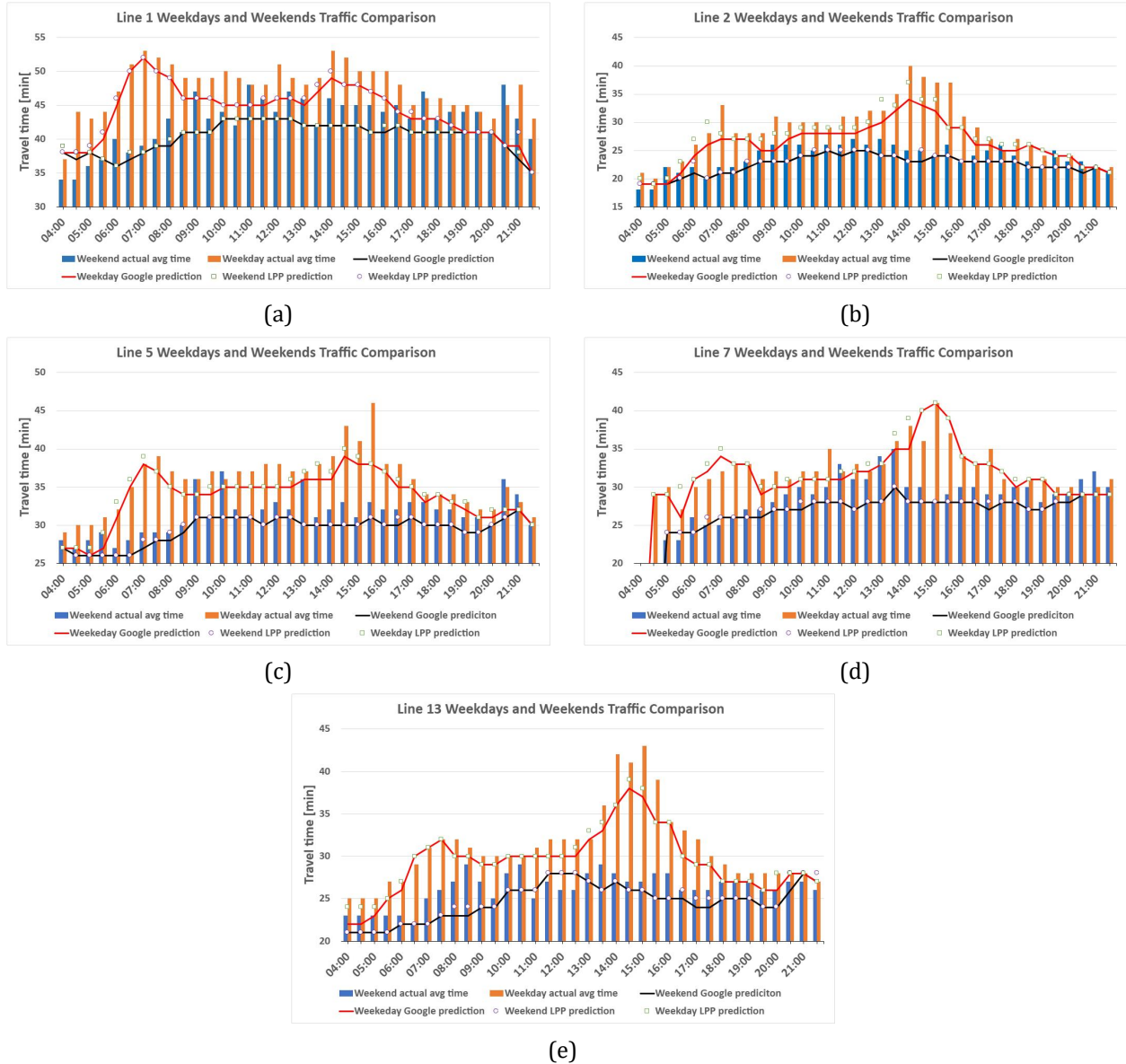
Source: Own elaboration, 2026.

Services' prediction errors are significantly larger during weekdays compared to weekends. Additionally, lines 1, 2, and 5 have considerably higher MSE values than Lines 7 and 13, indicating a greater occurrence and/or magnitude of large prediction errors thus higher variability in the accuracy. Overall, the models are relatively stable with particularly small errors observed on weekends for Lines 2 and 13. Additional analysis of 30-minute intervals provides more granular insights into the model's accuracy, which is shown in Figure 5.

While Google's predictions generally follow the observed trends in observed travel times, several substantial discrepancies were identified. For Lines 1, 2, and 5, prediction errors are most pronounced during weekday afternoon peak periods, whereas differences during other time intervals and weekends are comparatively small. Line 2, with a free-flow travel time equivalent of approximately 23 minutes, is the most affected by congestion and shows the largest prediction errors, with average deviations reaching 7 minutes during both morning and afternoon peaks. Across all the lines analyzed, travel time estimates are considerably more accurate during morning peak hours than during the afternoon peak.

A comprehensive analysis of Line 13 is limited by the small number of observations, particularly on weekends, due to the line's relatively low service frequency. Nevertheless, the available weekday data reveal the same general pattern observed on the other lines: the largest discrepancies occur during the afternoon peak period, when Google's predictions consistently underestimate observed travel times by approximately 15%.

Figure 5. Distributions of Google Routes API predicted travel times in comparison with observed travel times throughout the day for each route – weekdays and weekends. Google predictions for weekends (red line) and weekdays (black line) are mapped together with observed travel times (blue bar for weekends, orange bar for weekdays) and LPP predictions for weekends (purple circle) and weekdays (green square).

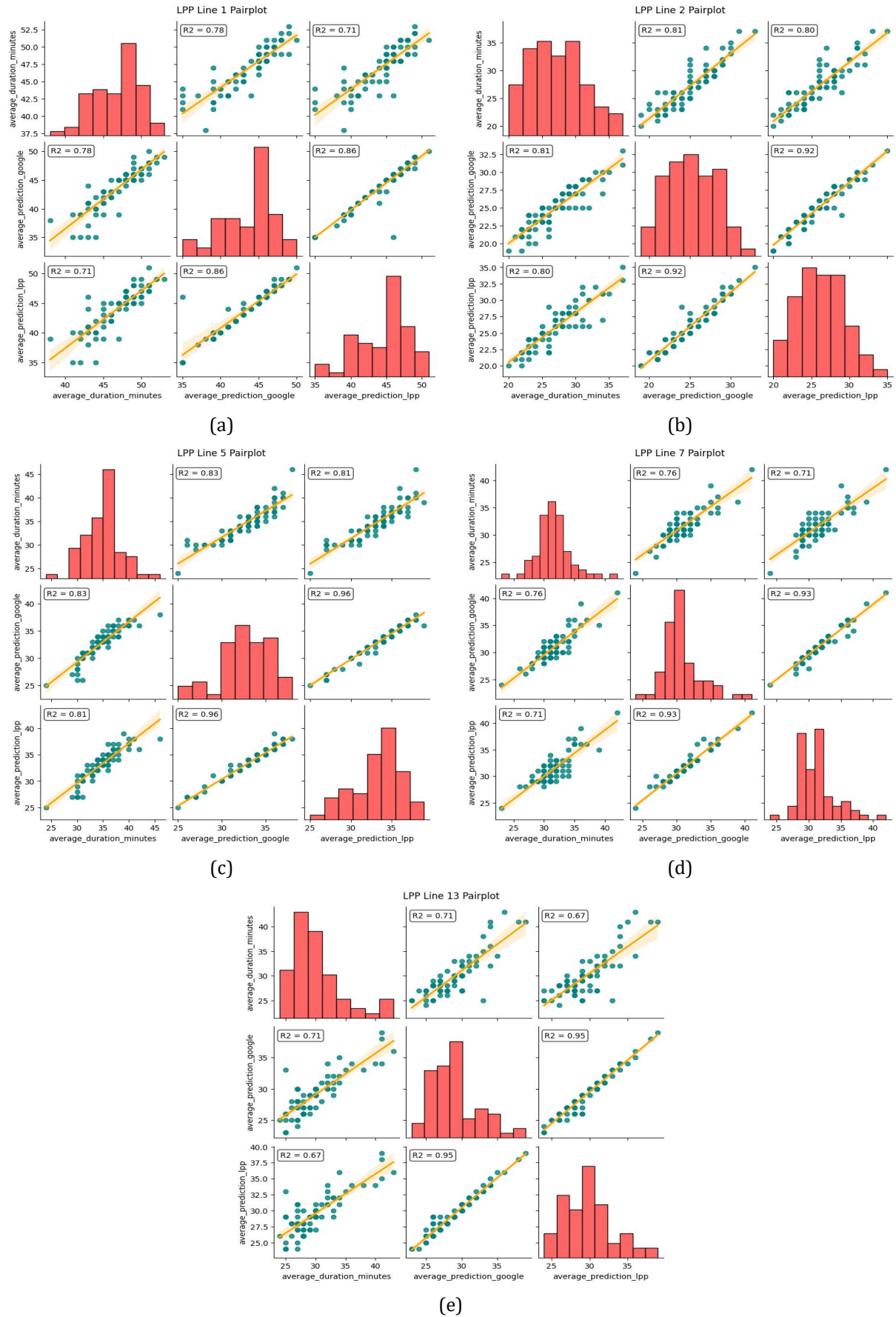


Source(s): Own elaboration, Google Routes API and LPP API data, 2026.

3.3. Comparison of travel time prediction services

Prediction performance was quantified using Pearson's correlation coefficient (R), which measures the strength of the linear relationship between predicted and observed travel times. Comparisons among observed trip durations, Google Routes predictions, and LPP predictions are presented in Figure 6. The diagonal panels (red histograms) depict the distributions of the respective variables across the analyzed time intervals. Greater similarity between these distributions indicates closer agreement between predicted and observed travel times. The off-diagonal panels display individual 30-minute observations (blue scatter points), the fitted linear regression line (orange), and its corresponding confidence interval (yellow shaded area).

Figure 6. Pair plots of relationships between trip duration, Google Routes prediction, and LPP prediction.



Source: Own elaboration derived from Google Routes and LPP predictions, 2026.

Overall, Google Routes predictions show a stronger correspondence with observed travel times than those produced by the LPP model, as evidenced by the tighter clustering of observations around the regression line and higher correlation values

The largest differences between the two services are observed for Lines 1 (Google: $R = 0.78$; LPP: $R = 0.71$) and 7 (Google: $R = 0.76$; LPP: $R = 0.71$). In contrast, the correlation coefficients for Lines 2 and 5 are nearly identical, suggesting comparable predictive performance by the two models. Across all five lines, the correlation between predicted and observed travel times ranges from $R = 0.67$ to $R = 0.81$. Although these values indicate a moderate to strong association, they also suggest that both prediction models have potential for further improvement.

3.4. Regression of travel times

The established dummy regressor and random forest models were both applied to the collected dataset. The initial analysis was done by creating five individual models for each line. We ran the process for each line separately. The cross-validation results of the Dummy Regressor and Random Forest models are presented in Table 4.

Table 4. Cross-validation results for the Dummy Regressor and Random Forest models. Cumulative values combine weekday and weekend predictions weighted by the number of trips.

	Line 1	Line 2	Line 5	Line 7	Line 13	All lines
Weekdays						
Dummy Regressor	4.04	4.61	3.60	3.06	4.06	3.86
MAE & MSE	29.10	28.42	21.93	21.43	37.58	31.26
Weekends						
Dummy Regressor	3.50	2.50	2.78	2.72	2.00	2.76
MAE & MSE	23.69	10.65	28.19	13.13	7.01	16.16
Cumulative						
Dummy Regressor	3.97	4.08	3.40	2.98	3.80	3.65
MAE & MSE	28.42	33.97	30.16	19.50	33.76	28.30
Weekdays						
Random Forest	3.53	3.29	2.77	2.47	2.48	2.99
MAE & MSE	23.19	27.28	22.48	15.23	20.21	21.86
Weekends						
Random Forest	3.31	2.27	2.75	2.40	1.96	2.58
MAE & MSE	19.98	8.78	24.58	11.09	8.06	15.36
Cumulative						
Random Forest	3.51	3.04	2.77	2.46	2.41	2.91
MAE & MSE	22.79	22.64	22.98	14.27	18.69	20.58

Source: Own elaboration obtained through model CV, 2026.

The Dummy Regressor provided a naïve mean-based reference rather than a competitive prediction model. As expected, the model provides very poor predictions even when compared to the original Google (Table 2) and LPP (Table 3) predictions. However, Its results establish a baseline for assessing whether the temporal, contextual, and service-based predictors incorporated into the Random Forest reduce prediction error.

Comparing the Random Forest results to Google (Table 2) and LPP (Table 3) predictions, we noticed a significant improvement in some bus lines. Most of the lines showed a minor improvement, especially compared to LPP internal prediction model, which even outperformed our new model in Lines 2 and 5. The new model, however consistently outperformed Google's

predictions across the board. The biggest improvement is shown in line 1 – 0.52 minutes (12.9%) compared to Google, and 0.38 minutes (9.77%) compared to LPP.

Table 5 shows normalized MAE values across all analyzed lines and models using the base MAE values from Tables 2, 3 and 4. Relatively Line 2 is the most unpredictable line across the board as it is also the shortest among analyzed segments.

Table 5. Normalized MAE (%) results across prediction models

	Line 1	Line 2	Line 5	Line 7	Line 13	All lines
Google Routes predictions nMAE (%)	8.57%	12.42%	8.20%	7.97%	8.83%	9.17%
LPP predictions nMAE (%)	8.28%	11.23%	7.80%	8.03%	8.59%	8.59%
Our model Dummy Regressor nMAE (%)	8.45%	15.69%	9.71%	9.61%	13.10%	10.45%
Our model Random Forest nMAE (%)	7.47%	11.69%	7.91%	7.94%	8.31%	8.35%

Sources: Own elaboration, Google Routes API and LPP API data 2026.

To create a single prediction model that can be used on all lines, we merged the dataset of all individual lines into one big dataset of 4,702 individual datapoints. While this approach potentially sacrifices the accuracy of individual models, it utilizes the strength in numbers since the training and validation dataset is 4-5 times larger, which allows the model to create a much deeper understanding of the traffic behavior on the city level. The results of this analysis are reported in Table 6.

Table 6. LOOCV results for the Random Forest Model (Combined dataset)

	All lines
Weekdays	2.97
MAE & MSE	21.74
Weekends	2.55
MAE & MSE	15.13
Cumulative	2.89
MAE & MSE	20.44

Source: Own elaboration, 2026.

Another advantage of this approach is that it creates a model that can be used on all public transit lines in the city, rather than only ones for which the test and validation data were collected. Having more lines will create a more accurate city-wide model.

Final results of our models show increased precision across the board. Initially, both Google's and LPP's predictions had much lower errors for weekend days than for weekdays. This was most notable in MSE value, which is around 40% lower, signifying that those have quite large discrepancies when unexpected delays occur.

Our Dummy Regressor individual model performed the worst across the board and was not assessed for the Combined dataset due to poor initial performance. Our Random Forest models' prediction values remained on par with Google's and LPP's models during the weekend trips. However, we managed to achieve much better accuracy during the weekdays. Individual lines dataset model managed an 8% improvement compared to Google, and 4% improvement compared to LPP, while combined dataset model managed 9% and 4% respectively. When it

comes to MSE, the differences are even larger – individual lines dataset model managed a 15% improvement compared to Google, and 10% improvement compared to LPP, just like the combined dataset model.

Table 7. Unified results table - MAE & MSE for Google, LPP and LOOCV models

	Google Routes API	LPP API	LOOCV Dummy Regressor Individual lines dataset	LOOCV Random Forest Individual lines dataset	LOOCV Random Forest Combined dataset
Weekdays	3.25	3.10	3.86	2.99	2.97
MAE & MSE	25.67	24.15	31.26	21.86	21.74
Weekends	2.63	2.56	2.76	2.58	2.55
MAE & MSE	15.42	14.95	16.16	15.36	15.13
Cumulative	3.13	3.00	3.65	2.91	2.89
MAE & MSE	23.66	22.34	28.30	20.58	20.44

Sources: Own elaboration, Google Routes API, LPP API, 2026.

4. Conclusions

Accurate public transport travel predictions are increasingly important for trip planning, as they can influence departure time, mode choice, and route selection. This study examined the accuracy of existing travel-time predictions and assessed whether their integration with additional temporal and contextual information can improve prediction performance.

We evaluated the existing most notable public transit travel times prediction services available in Ljubljana, Slovenia. These included Google Routes (used through Google Maps) and the official city public transit provider LPP. The analysis revealed that while the models provide a good level of precision during weekends and off-peak hours, they are quite inconsistent in rush hours and on heavily congested lines. Additionally, during the analysis of observed public transit travel times, we noticed that discrepancies between the shortest and the longest trip during the day on individual lines span from 41% (for Line 1) to 85% (for Line 2), signifying that lack of dependability and consistency of public transit which puts even more emphasis on having accurate prediction models.

We subsequently developed a prediction model combining data from Google Routes and LPP routing systems with weather and time-stamp information to improve travel-time estimates for a certain public transit route. LOOCV with Random Forest was used on a three-week dataset to obtain the model that showed increased prediction accuracy by being between 4% and 9% more accurate (MAE) than Google Routes API and LPP API, with much lower MSE values (15% and 10% respectively) signifying lower values of extreme errors. Even with a very limited dataset of only three weeks, we managed to create not only more precise, but much more stable model, which is less prone to large errors in prediction. The method of data collection, analysis, and model creation is relatively straightforward and can be transferred to other transit systems where multi-source data and observed public transport travel times are available, either from historical records or through real-time data collection or calculation. This enables individual cities to develop locally adapted models that can improve the accuracy of travel-time estimates.

Several methodological limitations must be acknowledged when interpreting the findings of this study. First, data collection was restricted to a three-week period and focused on a limited subset of five transit routes, with partial coverage of certain segments. Consequently, the dataset may not fully capture seasonal variations or network-wide traffic dynamics. Future research should expand both the spatial and temporal scope of the study to evaluate a more comprehensive array of routes over a longitudinal period.

Second, the proprietary nature of Google Routes and LPP APIs restricts a deeper deconstruction of their predictive mechanics. We proceeded with the assumption that both algorithms integrate historical baselines with real-time traffic conditions, consistent with Google's documented methodology for live routing services (Lau, 2020; Google, 2026). However, the extent of algorithmic evolution remains undisclosed. Notably, there is no public indication that either API incorporates weather into its calculations, thus our inclusion of weather data introduces an explanatory dimension that these models likely omit.

Third, regarding the data acquisition constraints discussed in Section 2.2, arrival times at origin and destination stations were approximated from LPP API queries at three-minute intervals, introducing timing uncertainty into individual observations. Since the associated errors are expected to be randomly distributed, their influence diminishes through aggregation across repeated observations. Reducing the query interval to one minute or less could theoretically reduce this uncertainty, however, we contend that such an adjustment would yield negligible analytical improvements while significantly increasing computational overhead and cost.

Finally, our methodology required the removal of extreme outliers (i.e. incomplete trips, journeys exceeding twice the expected duration) to ensure model stability. While statistically necessary for this study, a dedicated analysis of these events represents a highly valuable avenue for future research. Transit dependability is as critical to urban mobility as operational speed. Investigating the frequency and underlying causes of these service failures would provide a more holistic assessment of the Ljubljana public transport network's reliability. Furthermore, integrating service-failure probabilities into future predictive models could substantially refine their real-world accuracy. Future iterations of this framework could be scaled and deployed as an official, city-wide predictive tool for public transit travel times. To ensure the model remains robust year-round and consistently improves, we recommend focusing on the following developmental areas.

The framework could be enhanced with incremental learning capabilities, enabling it to continuously incorporate newly acquired data without requiring complete retraining. This approach would allow the model to adapt efficiently to evolving traffic conditions, changing passenger behavior, and long-term route modifications, thereby improving prediction accuracy, maintaining model relevance over time, and extending its operational lifespan. Future iterations must expand upon the current weather parameters by incorporating deep seasonal modeling. Because specific weather conditions severely impact traffic flow, a higher-resolution meteorological model is required to provide accurate seasonal predictions. The proposed model's predictive capacity must account for irregular traffic disruptions caused by national holidays, major city events, school vacations, and tourist seasons. We propose developing an anomaly-aware calendar system that adjusts predictive baselines during these events. Additionally, since the intensity and timing of peak-hour congestion fluctuate across different seasons, capturing these longitudinal trends is essential for improving accuracy during highly unpredictable traffic periods.

The ongoing research with the study pilot demonstrated in this paper therefore provides a basis for developing city-specific public transport prediction models with improved accuracy over existing prediction systems. Applicable to cities and regions with access to comparable public transport data, the approach shows how available prediction services can be treated not as final outputs, but as inputs that can be locally recalibrated and enriched with contextual information. This opens a practical path toward prediction systems that are more responsive to the specific operating conditions of individual public transport networks.

Acknowledgements

The authors acknowledge Ljubljanski potniški promet (LPP) for providing access to GTFS-RT and GeoJSON public transport data. This research was funded by the Slovenian Research and Innovation Agency (ARIS) under project J5-70190 “Integrative urban analytics: Advancing interoperability of open-source data and methods to inform urban insights and decision-making”, and programmes P5-0068 “Sustainable planning for the quality living space” and P2-0359 “Pervasive computing”.

Data availability

Public transport travel-time data were obtained via the Google Routes API (Google, 2026) and the LPP API (Ljubljanski potniški promet, 2026). Weather data were obtained from Open-Meteo (2026). Documentation for these services is available at: Google Routes API: <https://developers.google.com/maps/documentation/routes>; LPP API: <https://data.lpp.si/doc/>; Open-Meteo: <https://open-meteo.com/>. Access to Google Routes and LPP APIs is subject to provider-specific access conditions and usage limits. The processed dataset generated for this study is available from the corresponding author upon reasonable request.

References

- Bachu, A. K., Reddy, K. K., & Vanajakshi, L. (2021). Bus travel time prediction using support vector machines for high variance conditions. *Transport*, 36(3), 221–234. <https://doi.org/10.3846/transport.2021.15220>
- Bole, D. (2004). Geografija javnega potniškega prometa na primeru Ljubljane. *Geografski vestnik*, 76(2), 21–32.
- Chen, R., Zhang, J., Xu, X., Han, T.-D., & He, M. (2026). Comparative analysis and case study of travel time reliability measures for public usage of transportation information. *Journal of Transport Geography*. <https://doi.org/10.1016/j.jtrangeo.2026.103928>
- Dunne, L., Rocco Di Torrepadula, F., Di Martino, S., McArdle, G., & Nardone, D. (2023). Bus journey time prediction with machine learning: An empirical experience in two cities. In M. A. Mostafavi & G. Del Mondo (Eds.), *Web and wireless geographical information systems* (pp. 105–120). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-34612-5_7
- Google. (2024, October 16). *GTFS Realtime Overview*. Google for Developers. <https://developers.google.com/transit/gtfs-realtime>
- Google. (2026). *Routes API*. Google for Developers. <https://developers.google.com/maps/documentation/routes>
- Gössling, S., Neger, C., Steiger, R., & Bell, R. (2023). Weather, climate change, and transport: A review. *Natural Hazards*, 118, 1341–1360. <https://doi.org/10.1007/s11069-023-06054-2>
- Hansson, J., Pettersson-Löfstedt, F., Svensson, H., & Wretstrand, A. (2022). Effects of rural bus stops on travel time and reliability. *Public Transport*, 14, 683–704. <https://doi.org/10.1007/s12469-021-00281-1>
- Henríquez-Jara, B., Arriagada, J., & Tirachini, A. (2025). Impact of real-time information on passenger satisfaction across varying public transport quality levels in 13 Chilean cities. *Transportation Research Part A: Policy and Practice*, 200. <https://doi.org/10.1016/j.tra.2025.104408>
- Kandiri, A., Ghiasi, R., Nogal, M., & Teixeira, R. (2024). Travel time prediction for an intelligent transportation system based on a data-driven feature selection method considering temporal correlation. *Transportation Engineering*, 18, Article 100272. <https://doi.org/10.1016/j.treng.2024.100272>
- Lau, J. (2020, September 3). Google Maps 101: How AI helps predict traffic and determine routes. The Keyword. <https://blog.google/products-and-platforms/products/maps/google-maps-101-how-ai-helps-predict-traffic-and-determine-routes/>

- Liu, L., Porr, A., & Miller, H.J. (2023). Realizable accessibility: Evaluating the reliability of public transit accessibility using high-resolution real-time data. *Journal of Geographical Systems*, 25(3), 429–451. <https://doi.org/10.1007/s10109-022-00382-w>
- Ljubljanski potniški promet. (2025). *Shema dnevnih linij LPP: Mestni javni linijski prevoz in integrirane potniške linije* [Public transport network map].
- Ljubljanski potniški promet. (2026). *LPP API documentation*. <https://data.lpp.si/doc/>
- Mestna občina Ljubljana. (2022). Lokalni energetske koncept Mestne občine Ljubljana za obdobje od 2022 do 2030: I. del – Analiza stanja.
- Open-Meteo. (2026). *Weather Forecast API*. <https://open-meteo.com/en/docs>
- Parslov, A., Petersen, N.C., & Rodrigues, F. (2023). Short-term bus travel time prediction for transfer synchronization with intelligent uncertainty handling. *Expert Systems with Applications*, 232, 120751. <https://doi.org/10.1016/j.eswa.2023.120751>
- Rashvand, N., Hosseini, S., Azarbayjani, M., & Tabkhi, H. (2024). Real-time bus arrival prediction: A deep learning approach for enhanced urban mobility. In *Proceedings of the 13th International Conference on Operations Research and Enterprise Systems* (pp. 123–132). SCITEPRESS. <https://doi.org/10.5220/0012365500003639>
- Schmitt, A. U., Burgemeister, F., Dorff, H., Finn, T., Hansen, A., Kirsch, B., Lange, I., Radtke, J., & Ament, F. (2023). Assessing the weather conditions for urban cyclists by spatially dense measurements with an agent-based approach. *Meteorological Applications*, 30(6), e2164. <https://doi.org/10.1002/met.2164>
- scikit-learn developers. (n.d.). *sklearn*. scikit-learn. Retrieved August 28, 2026, from <https://scikit-learn.org/stable/api/sklearn.html>
- Serin, F., Alisan, Y., & Erturkler, M. (2022). Predicting bus travel time using machine learning methods with three-layer architecture. *Measurement*, 198, 111403. <https://doi.org/10.1016/j.measurement.2022.111403>
- Slovenian Environment Agency (ARSO). (2025). Ljubljana Bežigrad: Climate statistics and normals 1991–2020.
- SURS, Statistical Office of the Republic of Slovenia. (2026). *Selected data on municipalities, Slovenia, annually*. SiStat Database. <https://pxweb.stat.si/SiStatData/pxweb/en/Data/Data/2640010S.px/>
- Trček, I. S., & Rijavec, R. (2026). Strategic evaluation of cycling infrastructure in emerging contexts: A case study of high-capacity corridors in Ljubljana, Slovenia. *European Transport Research Review*, 18, Article 67. <https://doi.org/10.1186/s12544-026-00833-1>
- van Ardenne, M. T., Cebecauer, M., Cats, O., & Ma, Z. (2025). Personalised passenger information systems in public transport: A review and a 5-level personalisation taxonomy. *Transport Reviews*, 45(6), 1016–1047.
- Xu, H., & Ying, J. (2017). Bus arrival time prediction with real-time and historic data. *Cluster Computing*, 20(4), 3099–3106. <https://doi.org/10.1007/s10586-017-1006-1>
- Yu, B., Wang, H., Shan, W., & Yao, B. (2018). Prediction of bus travel time using random forests based on near neighbors. *Computer-Aided Civil and Infrastructure Engineering*, 33(4), 333–350.