# Traffic Volume Forecasting Project

Lisa Rigo, Paola D'Amore, Jacopo Malaguti

Exam Project: Forecasting and Predictive Analytics

Professor: Luca Trapin

September 2025

## 1.PROBLEM STATEMENT

We are the Analytics Department at BrightPath Consulting Company, a B2B firm that develops forecasting and data-driven solutions.

Among the several projects we handle, the current assignment comes from EverTrack Logistics, a company specializing in the transportation of perishable goods such as food and pharmaceuticals. EverTrack operates mainly on Interstate 94, the primary route between Minneapolis and St. Paul, here assumed to have no intermediate entries or exits.

Our task is to develop a forecasting model capable of generating accurate <u>hourly forecasts</u> of the <u>expected traffic volume</u> over a <u>72-hour horizon</u>, enabling EverTrack Logistics to plan deliveries effectively: schedule trips in advance to avoid congestion and shift delivery windows when needed, reducing costs and ensuring goods remain fresh while meeting Service Level Agreements (SLA).

## 2.DATA

### 2.1 Dataset

The dataset used in this study was obtained from the UCI Machine Learning Repository and corresponds to the Metro Interstate Traffic Volume dataset.

It consists of hourly observations recorded on Interstate 94 Westbound at the Minnesota Department of Transportation (MnDOT) Automatic Traffic Recorder (ATR) station 301, located roughly midway between Minneapolis and St. Paul.

The data covers the period from January 1, 2016 to September 30, 2018, with a total of 24,096 hourly observations and no missing values.

| Variable Name | Description | Type |
|---|---|---|
| temp | Average temp (kelvin) | Numeric |
| rain_1h | Amount of rain that occurred in the hour (mm) | Numeric |
| snow_1h | Amount of snow that occurred in the hour (mm) | Numeric |
| clouds_all | Percentage of cloud cover | Integer |
| weather_main | Short textual description of the current weather | Categorical |
| weather_description | Longer textual description of the current weather | Categorical |
| time_series | Hour of the data collected in local CST time | Date |
| **traffic_volume** | Hourly I-94 ATR 301 reported westbound traffic volume | Integer |

Table 1: Description of the dataset variables

Table 1: The dataset contains the target variable which is ***traffic_volume***, defined as the number of vehicles passing the ATR station during each hour and other six covariates.

## 2.2 Exploratory Data Analysis (EDA)

To explore traffic patterns over time, we first visualize average traffic volumes, aggregating the data by month, year, or day of the week.

To explore daily patterns over the months:

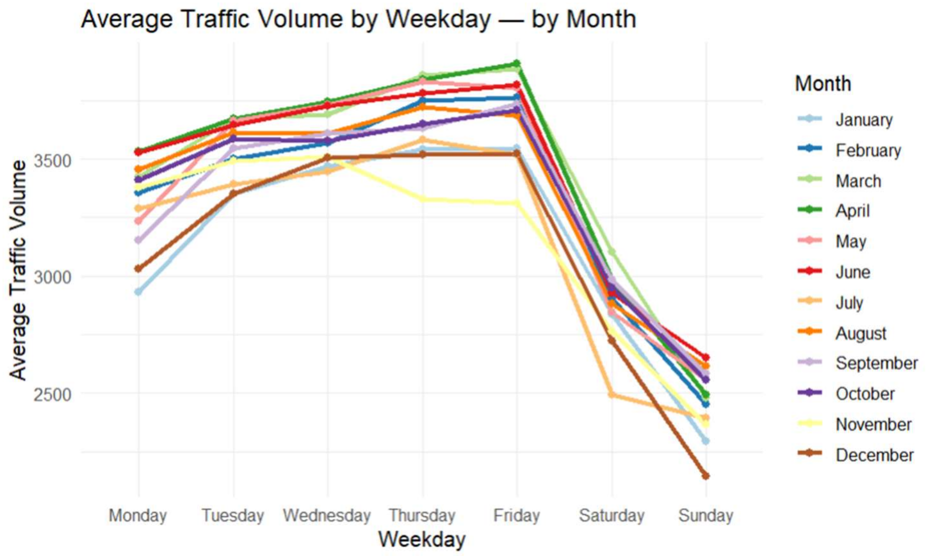



Figure 1: Average traffic per weekday aggregated by month

Figure 1: The plot shows a stable weekly pattern, with traffic showing the highest volumes on weekdays and dropping over weekends due to commuting. Traffic is lower during holiday periods (December, January, and July), but the weekly cycle remains consistent, indicating strong weekly seasonality, a pattern widely documented in the literature, where traffic volumes show pronounced day-of-week effects (higher on weekdays, lower on weekends).

To explore hourly patterns over the years:

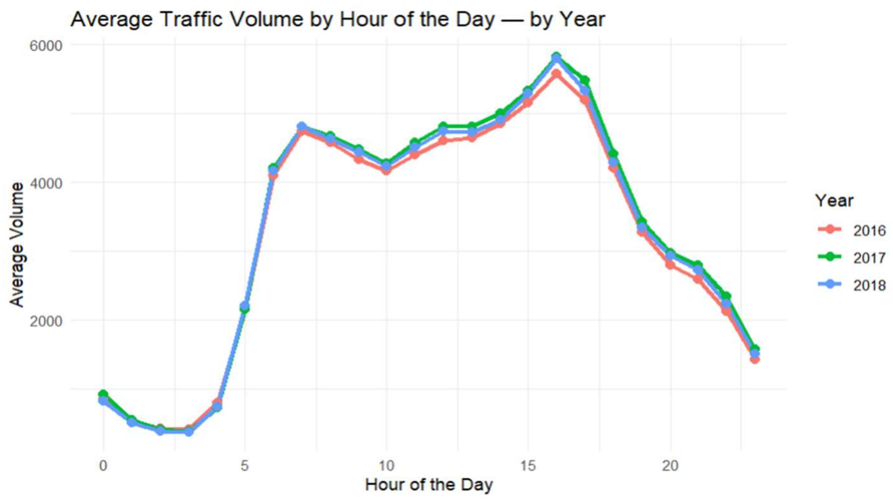



Figure 2: Average traffic per hour, aggregated by year

Figure 2: Daily patterns with morning and late-afternoon peaks are consistent across years, confirming stable daily seasonality.

To explore hourly patterns over the week:

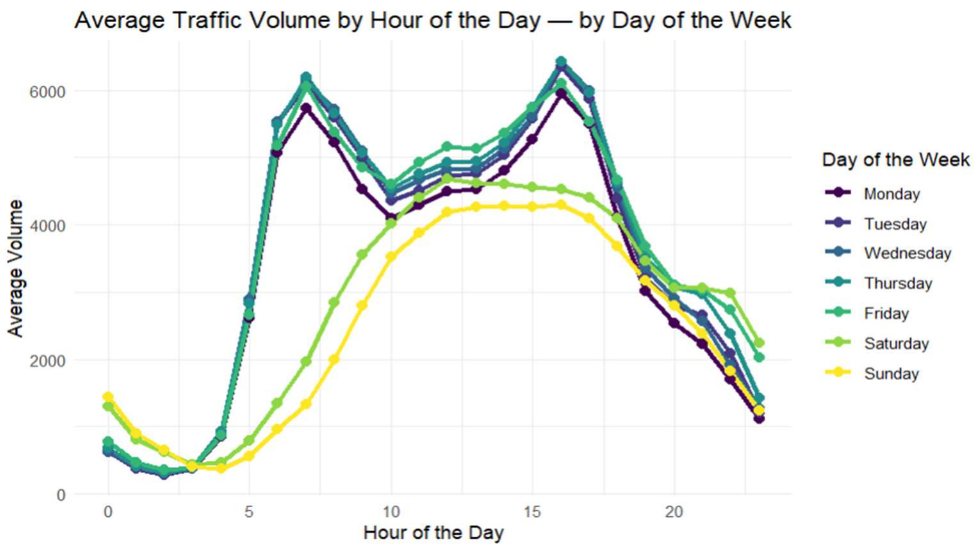



Figure 3: Average traffic per hour, aggregated by day of the week

Figure 3: Weekdays show two pronounced peaks—around 8 AM and 5 PM—reflecting typical commuting behaviour, while weekend traffic is more evenly distributed, driven by leisure and non-work activities. This clear distinction between weekdays and weekends highlights multiseasonality, where traffic follows a daily cycle that differs depending on the day of the week.

The ACF (Figure 4) shows persistent peaks at multiples of 24 hours, reflecting daily and weekly cycles. The sinusoidal profile indicates deterministic seasonality linked to commuting routines and weekday/weekend effects. No long-term trend emerges, as traffic fluctuates around a stable average, consistent with business expectations and with the literature, supporting an explicit seasonal specification for forecasting.

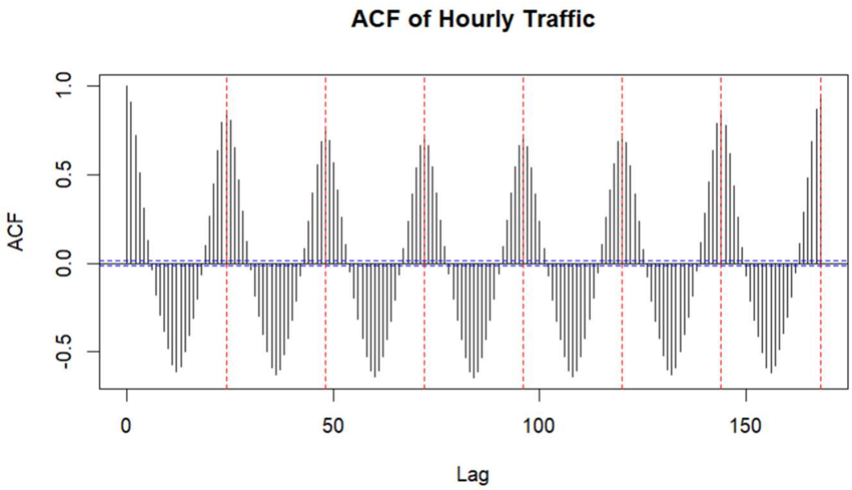


Figure 4: ACF of hourly traffic volume

To further investigate seasonality, we compared traffic volumes at 8 AM on Mondays and Saturdays across all years (Figure 5 and 6). Comparing both perspectives allows us to check whether monthly seasonality is stable across different weekly dynamics:

Figure 5: Average Traffic at 8 AM on Mondays by Month

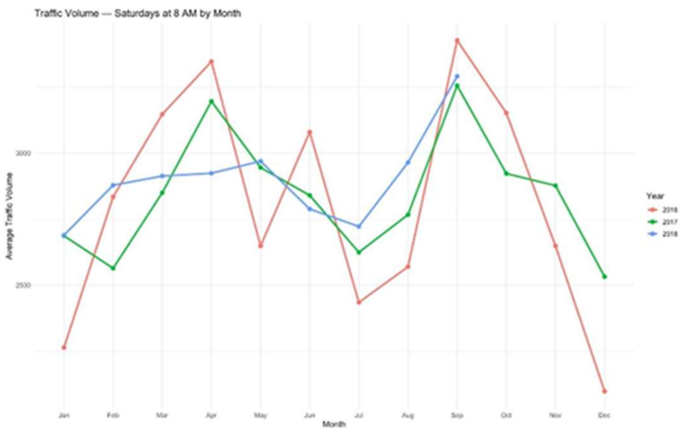


Figure 6: Average Traffic at 8 AM on Saturday by Month

The pattern remains stable across years, with no evident increase in variance, which suggests a deterministic seasonal effect. This interpretation is supported by the Canova–Hansen (CH) test, applied separately to traffic observed at 8 AM on Mondays and on Saturdays, aggregated by month to account for differences between weekday and weekend patterns. The test produced p-values of

0.08 and 0.20, both above the 5% significance level, meaning we do not reject the null of constant beta parameters.

The ADF test with intercept and trend specification indicates that the series fluctuates around a positive mean (significant intercept) but shows no significant deterministic trend ($p = 0.462$). The test statistic ($-58.22 < -3.41$) rejects the null of a unit root, confirming that traffic is stationary around its mean and no differencing is required.

## 3.METHODS

We forecast hourly traffic on I-94 over a 72-hour horizon using three approaches: (i) ETS models with SES and Holt-Winters; (ii) a linear regression with hourly and weekend dummies; and (iii) a SARIMA model to capture autocorrelation in regression residuals.

### (i) Exponential smoothing models

We use the ETS family in state-space form, which becomes not just an algorithm but a statistical model where we adopt additive errors $\varepsilon_t \sim (0,\sigma^2)$. Parameters and initial states are estimated jointly by maximum likelihood, which is implemented by numerical optimization.

Within the ETS family, we selected:

(i.a) Simple exponential smoothing, ETS(A,N,N) with level only.

| Model | Observation | State Equations |
|---|---|---|
| ETS(A,N,N) | $y_t = \ell_{t-1} + \varepsilon_t$ | $\ell_t = \ell_{t-1} + \alpha\varepsilon_t$ |

$\alpha \in (0,1).$

Table 2: Simple exponential smoothing specification in the state-space form.

(i.b) SES + Additive Seasonality, ETS(A,N,A) with no trend and additive seasonality.

| Model | Observation | State Equations |
|---|---|---|
| ETS(A,N,A) | $y_t = \ell_{t-1} + s_{t-m} + \varepsilon_t$ | $\ell_t = \ell_{t-1} + \alpha\varepsilon_t$<br>$s_t = s_{t-m} + \gamma\varepsilon_t$ |

$\gamma \in (0,1),\ m = 24.$

Table 3: SES + Additive Seasonality specification in the state-space form.

We exclude a trend component because EDA and unit-root tests show no evidence of a persistent trend; we use additive seasonality because suitable when seasonal fluctuations are roughly constant. Residual diagnostics indicate residuals not normally distributed so we proceed with 95% prediction interval simulation based (bootstrap).

(ii) Linear Regression with (iii) seasonal ARMA errors: SARIMAX

We modeled daily and weekly seasonality through a linear regression with hourly dummies (Hour00–Hour23) and a weekend dummy, including Hour×Weekend interactions which allows the model to capture distinct hourly profiles for weekdays and weekends.

The regression model is specified as:

$$TrafficVolume_t = \sum_{h=0}^{23} \beta_h \cdot Hour_{h,t} + \gamma \cdot Weekend_t + \sum_{h=0}^{23} \delta_h \cdot (Hour_{h,t} \cdot Weekend_t) + \varepsilon_t$$

where:

- $Hour_{h,t} = 1$ if observation t is at hour h, 0 otherwise
- $Weekend_t = 1$ if t falls on Saturday/Sunday
- $Hour_{h,t} \cdot Weekend_t$ is the interaction, adjusting hourly means for weekends
- $\varepsilon_t$ = residual part

We examine the residuals' ACF and PACF (Figure 7) to assess serial correlation and guide SARIMA specification:

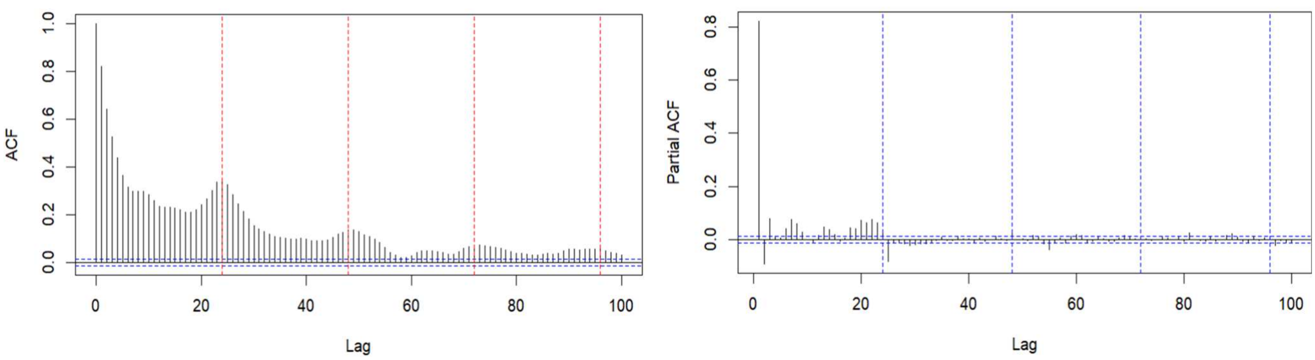
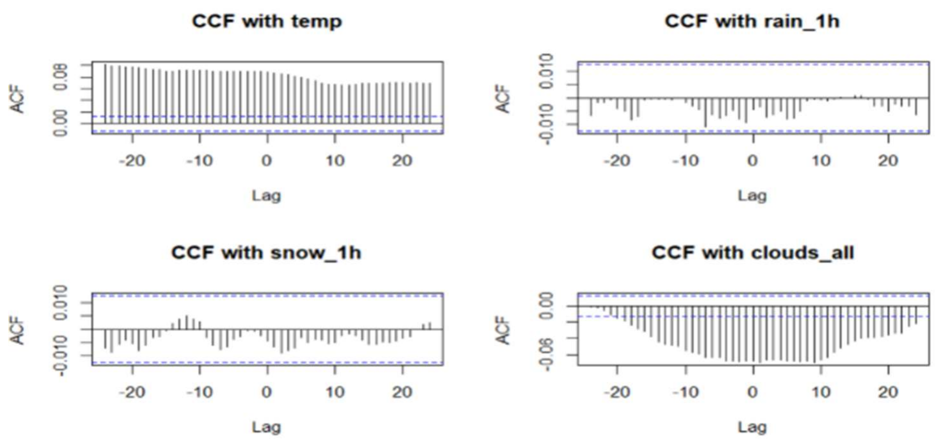



Figure 7: ACF and PACF of residuals from the fitted regression model

Residuals from the regression exhibited significant autocorrelation, so we fitted SARIMA models to account for remaining structure not captured by the regression. Several specifications were compared using AIC and BIC, between which we selected SARIMA(2,0,2)(1,0,1)[24] for its balance between fit and parsimony. Residual diagnostics of the selected SARIMA model were performed, showing no autocorrelation (approximately 0) and QQ-plot indicates residuals not normally distributed.

SARIMAX: Analysis with an exogenous covariate




We examined cross-correlations between regression residuals and weather covariates to assess potential delayed impacts on traffic. The largest correlation, with lagged temperature (≈0.08), was negligible and did not improve the SARIMAX fit.

## 4.RESULTS

We evaluate 72-step-ahead forecasts using a rolling-window with a fixed size (17,544 hours) for the training window that slides forward by one hour across the test period (2018-01-01 00 to 2018-09-30 23, 6,552 hours), re-estimating the model and producing a 72-hour forecast path at each position, for a total of 6,481 rolling windows.

For point forecasts, accuracy is measured by the Mean Squared Error (MSE) computed per horizon across windows, which is the natural measure of performance, since our goal is to estimate the expected traffic volume.
For interval forecasts, we build 95% bootstrap prediction intervals and evaluate them via unconditional coverage and Winkler score.

### 4.1 Point forecast

ETS models are competitive only at very short horizons, with accuracy deteriorating as *h* increases.

LM and LM+SARIMA achieve the lowest MSE on average across horizon. Between those two LM delivers lower MSE for most horizons. To test whether this difference in MSE is significant, we used Diebold Mariano test. At short horizons (h = 1–2) and at h = 15, the LM model exhibits a statistically significant advantage, as indicated by the negative Diebold–Mariano (DM) statistic with $p < 0.05$. Although SARIMAX achieves marginally lower MSE values at some horizons (e.g., h = 4–7), the corresponding DM tests fail to reject the null of significant difference. Overall, the findings favour LM for point forecasts.

| | StepAhead | DM_stat | DM_pvalue | Significant | Preferred_by_DM | MSE_LM | MSE_LM_SARIMAX | Preferred_by_MSE |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | -15.6619236 | 2.700824e-54 | TRUE | LM | 273741.3 | 380887.2 | LM |
| 2 | 2 | -4.4962601 | 7.036354e-06 | TRUE | LM | 273644.6 | 310926.6 | LM |
| 3 | 3 | -0.7726985 | 4.397290e-01 | FALSE | No significant diff | 273493.2 | 280008.7 | LM |
| 4 | 4 | 0.9341880 | 3.502417e-01 | FALSE | No significant diff | 273435.9 | 265489.9 | LM_SARIMAX |
| 5 | 5 | 1.6555929 | 9.785271e-02 | FALSE | No significant diff | 273430.0 | 259177.3 | LM_SARIMAX |

Figure 8: MSE & DM for the first 5 step ahead across window— LM vs. LM+SARIMA

### 4.2 Interval forecasts

The 95% prediction intervals are assessed using unconditional coverage (target = 0.95, binomial test) and the Winkler score. ETS models (SES and HW-additive) fail to reach the target coverage, making them unsuitable.

We therefore focus on LM and LM+SARIMA, which both satisfy the requisite of unconditional coverage.

In terms of sharpness, LM is clearly better at $h = 1$ (DM<0, p<0.05), LM+SARIMA is better at $h = 3$-6 (DM>0, p<0.05), and from roughly $h = 16$ onward LM has the lower Winkler on most horizons, with significant wins at 15 horizons versus 4 for LM+SARIMA; elsewhere differences are not significant. Overall, LM is the preferred.

| | StepAhead | UC_LinReg | UC_SARIMAX | UC_pvalue_LinReg | UC_pvalue_SARIMAX | Winkler_LinReg | Winkler_SARIMAX | DM_Winkler_stat | DM_Winkler_pvalue |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0.9532480 | 0.7295170 | 0.24255155 | 0.000000e+00 | 3261.007 | 5093.846 | -19.2015567 | 5.782497e-80 |
| 2 | 2 | 0.9526308 | 0.8967752 | 0.34694395 | 3.115625e-67 | 3232.081 | 3342.915 | -1.3748014 | 1.692406e-01 |
| 3 | 3 | 0.9532480 | 0.9356581 | 0.24255155 | 3.772765e-07 | 3230.211 | 3069.548 | 2.3746839 | 1.759293e-02 |
| 4 | 4 | 0.9524765 | 0.9510878 | 0.37695828 | 7.110429e-01 | 3242.698 | 3025.429 | 3.3133901 | 9.267783e-04 |
| 5 | 5 | 0.9518593 | 0.9540194 | 0.51216855 | 1.460037e-01 | 3244.471 | 3023.113 | 3.3224061 | 8.973913e-04 |

Figure 9: Coverage & Winkler of PI for the first 5 step ahead - LM vs. LM+SARIMA

## 5.DISCUSSION

### 5.1 Key findings

Our objective was to develop a forecasting model capable of generating accurate hourly forecasts of the expected traffic volume over a 72-hour horizon to support delivery planning. The data exhibit intra-day (24h) and intra-week (168h) seasonality with no evidence of a persistent trend. In this setting, the LM model delivers the lowest average MSE across horizons. Diebold–Mariano comparisons against the SARIMAX regression indicate **statistically significant** improvements for many horizons.

In our setting, the **single-seasonal** ETS specifications (ETS(A,N,N) and ETS(A,N,A) with s=24) are competitive at short horizons but deteriorate as the forecast horizon h increases, largely due to **unmodelled multi-seasonality** (24-hour and 168-hour patterns).

For **prediction intervals**, simple additive bootstraps around ETS produced **under-coverage** relative to the 95% level (Figure 9), whereas bootstrap intervals constructed on top of the LM were **better calibrated**.

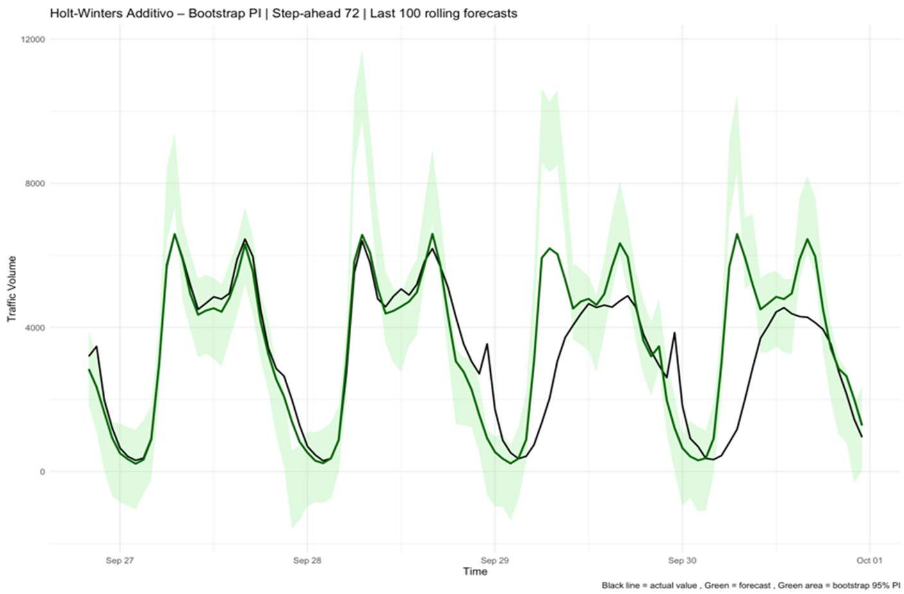



Figure 10: Holt–Winters additive model with 72-step-ahead rolling forecasts (last 100 windows)

The black line represents the observed traffic volume, the green line the point forecasts, and the shaded area the 95% bootstrap prediction intervals.

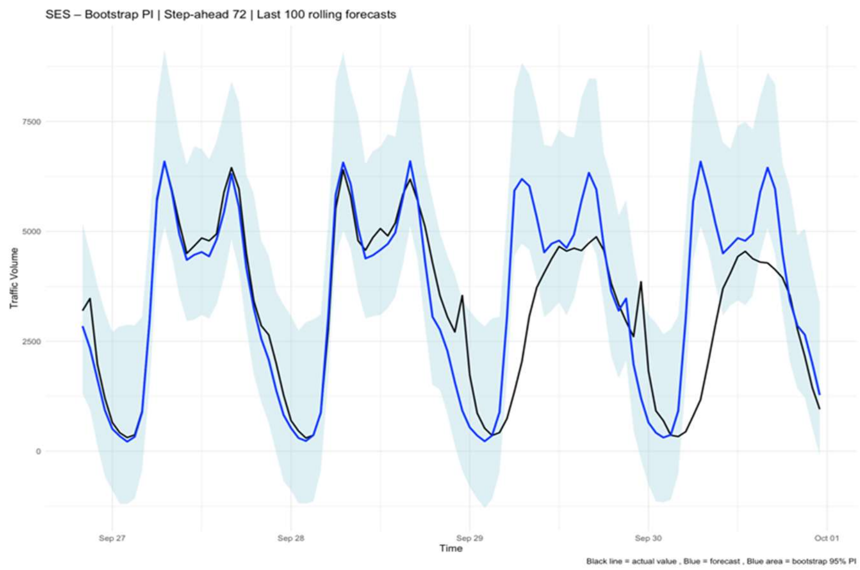



Figure 11: Simple Exponential Smoothing (SES) model with 72-step-ahead rolling forecasts (last 100 windows)

The black line represents the observed traffic volume, the blue line the point forecasts, and the shaded area the 95% bootstrap prediction intervals.

The Winkler loss favoured Holt-Winters over SES—narrower, more informative intervals—at the cost of lower coverage; overall the LM delivered the best trade-off between sharpness and calibration.

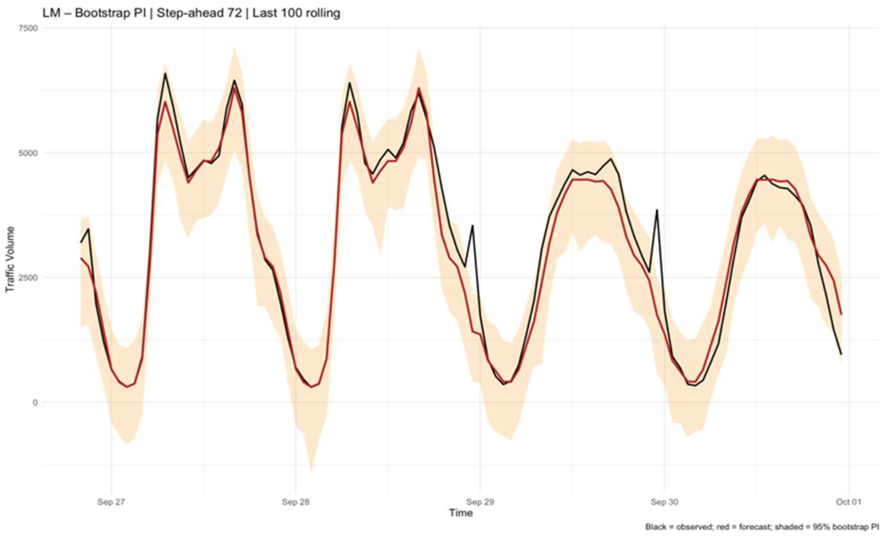



Figure 12: Linear Model (LM) with 72-step-ahead rolling forecasts (last 100 windows).

The black line represents the observed traffic volume, the red line the point forecasts, and the shaded area the 95% bootstrap prediction intervals.

In conclusion, the LM model provided the most accurate forecasts in our setting.
A possible improvement of the model could be achieved by adding calendar or event features (public holidays, sport, concerts, school terms) and incidents, allowing the model to adapt to irregular traffic conditions and further enhance forecast accuracy, especially during atypical days or periods of disruption.