Proceedings of the 10th International Conference on Economic Management and Green Development
DOI: 10.54254/2754-1169/2026.38458

Time Series Regression Forecasting of London Santander Cycles Demand: Nested Modelling with Unequal Weight Rolling Lags and Weekly Lag Correction

Xinyu Zhang[1,*], Yanhong Lin[2], Zixuan Li[3]

[1]Chongqing Nankai Middle School, Chongqing, China
[2]Nansha College Preparatory Academy, Guangzhou, China
[3]U LINK COLLEGE, Guangzhou, China
**Corresponding Author. Email: 18581479591@163.com*

Abstract. Against the background of fast-growing urban shared bike services, supply-demand imbalance has become a key operational challenge. Without reliable demand forecasting, operators face recurring vehicle surplus or shortage, raising logistics costs and harming user experience. This paper constructs time series regression forecasting models using daily Santander Cycles rental data from London spanning 2015 to 2017. We first build the baseline M13 model with an unequal-weight rolling lag term to capture short-term demand continuity. Nested models M1 and M4 are then developed by sequentially adding weekend, holiday, meteorological, and monthly seasonal covariates. To mitigate residual weekly autocorrelation, an extended M4 L7 model is created by introducing the lag7_cnt weekly lag variable. Model performance is assessed via static out-of-sample tests and 365-day rolling window recursive forecasting. Results show that M13 performs optimally for rolling-window high-frequency operational prediction, while M3 delivers the best static forecasting performance for long-term planning. M4 L7 effectively reduces medium- and long-term residual autocorrelation and balances explanatory power and prediction accuracy. This study provides scenario-specific model selection references for different bike-sharing operators.

Keywords: bike-sharing demand forecasting, time series regression, nested linear model, unequal-weight rolling lag

1. Introduction

1.1. Research background and significance

Bike-sharing provides a flexible option for short-distance travel and improves connections with public transportation. The UK government regards bike-sharing as a sustainable form of transport [1].

London is selected as the case study, as Santander Cycles operates a large dock-based bike-sharing network. According to the data from Transport for London [2], Santander Cycles has around

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2754-1169/2026.38458

800 docking stations in the city centre, with more than 12,000 bicycles across London. Users can collect a bicycle from one station and return it to any available docking point. In terms of pricing, the single-ride fare is £1.65 for 30 minutes. Notably, a major operational problem is the imbalance between bicycle supply and demand. Bicycles can be unevenly distributed: some stations are empty while others reach full capacity during peak hours. Operators therefore need to reposition and redistribute bicycles across stations, which pushes up transport costs and labour expenditure [3].

Nevertheless, different stakeholders hold distinct forecasting requirements, so no single model can satisfy all practical demands. For government authorities, the core demand scenario is long-term shared-bike system development and urban transport planning; their forecast horizon covers months to several years, and they place higher priority on model stability and interpretability than short-term prediction accuracy, to support decisions on bicycle procurement and new station construction. For medium-sized bike-sharing operators, the core demand scenario lies in routine daily system management; their forecast horizon ranges from weeks to one month, and stable forecasting performance is their primary concern, supporting regional demand analysis, bicycle redistribution, and regular system maintenance. For large-scale bike-sharing operators, the core demand scenario focuses on real-time operational adjustment and dynamic rebalancing; their forecast horizon corresponds to high-frequency short-term periods, and short-term prediction accuracy is their key priority, enabling timely responses to demand shocks caused by rush hours, adverse weather, and unexpected events and preventing bicycle shortages in high-demand areas. Therefore, there is no universally optimal forecasting model for all conditions. Operators need to select suitable forecasting methods according to their operational needs.

Existing bike-sharing forecasting studies have incorporated weather conditions, holidays, and seasonal variations as explanatory variables, yet these factors alone cannot fully explain all demand fluctuations, which are also driven by complex temporal patterns. Without accounting for such temporal effects, models may produce biased interpretations of demand changes. Accordingly, this paper introduces a lag-7 variable to capture weekly demand patterns. From a practical perspective, this study does not assume that one single model fits all operators. Instead, it compares model performance under different forecasting scenarios and provides guidance for model selection. The results can help bike-sharing operators in both the UK and China adopt appropriate forecasting approaches aligned with their operational requirements.

1.2. Literature review

1.2.1. External factors influencing bike-sharing demand

Early studies mainly focused on external factors that affect bike-sharing demand. Weather conditions, especially temperature and rainfall, greatly influence cycling demand [4]. El-Assi et al. [5] further pointed out that weather factors and urban environments both exert impacts on bike-sharing usage.

Morton [6] used London's bike-sharing system as a case study and showed that changes in weather and air quality affect users' riding behaviour. Besides weather factors, previous studies also found that bike-sharing demand presents time-related features. Faghih-Imani et al. [7] found that station locations, urban environments, and temporal factors shape bike-sharing trip flows. Eren and Uz [8] reviewed existing demand research and indicated that weather, time patterns, spatial conditions, and user behaviour could influence demand. However, most of these studies focus on explanatory determinants of demand, while limited attention has been paid to the inherent time-series continuity of bike-sharing demand.

DOI: 10.54254/2754-1169/2026.38458

1.2.2. Existing forecasting methods for bike-sharing demand

With the development of demand-forecasting research, researchers and operators have adopted different modelling approaches. Traditional time-series models such as ARIMA have been widely used for demand forecasting because they can capture historical patterns embedded in time-series data [9]. Afterwards, multiple regression models have been broadly applied, as they can incorporate weather, holiday, and seasonal variables to interpret how different factors shape demand. In recent years, machine-learning methods have been employed for bike-sharing forecasting and demonstrate strong predictive capacity [10].

Nevertheless, two major limitations can be identified from existing literature. First, most comparative model evaluations are conducted based on static known datasets. Second, many studies fail to simulate real-world forecasting workflows. Under practical operating conditions, operators need to update forecasts continuously with newly arrived daily data.

1.2.3. Weekly cycle patterns and research gaps

Bike-sharing demand contains multiple cyclic patterns, including daily cycles, weekly cycles, and seasonal cycles [11]. Existing forecasting research suggests that time-series forecasting should take these cyclic structures into account. Some scholars combine learning-based models with ARIMA to improve forecasting performance. Prior studies have adopted lag-1, lag-24, and lag-168 variables to capture short-term, daily, and long-term weekly cycles.

However, few studies examine whether weekly cyclic signals still persist in regression residuals after controlling for weather, holiday, and seasonal covariates. This constitutes a notable research gap. Even after incorporating widely-used seasonal and holiday covariates, unmodelled weekly periodicity may remain within model residuals and give rise to residual autocorrelation. To address this issue, this paper introduces lag7, representing bike-rental volumes observed seven days earlier, to account for cross-weekly demand variation, capture leftover weekly cyclical effects and reduce residual autocorrelation.

1.3. Model construction and performance evaluation

This study first processed the shared bicycle rental data, established nesting models with different lag windows, and determined the optimal windows through RMSE. After that, over the weekend and public holidays, variables such as temperature and humidity, and the influence of the month, gradually increased, and the prediction of the external sample and the variable window was implemented. Residue diagnosis was conducted through ACF examination and ljung - box examination, and the periodic user identified the correlation and optimized lag7 [12]. Finally, by integrating RMSE, residual independence, and the complexity of the model, and through multiple index evaluations, the script-based application is executed in combination with the operational requirements of other enterprises.

2. Methodology

2.1. Data source and description

The dataset, which is obtained from Transport for London (TfL), used in this study comprises daily-level total rental volumes for London Santander Cycles, spanning from January 2015 to January

DOI: 10.54254/2754-1169/2026.38458

2017 [13]. The vertical axis represents the daily number of bike rentals, and the horizontal axis represents the date.

The time-series graph shows three important characteristics. First, it has apparent short-term continuity: rental volume fluctuations are minor on successive days, which means that the daily demand for bike rentals has predictive information for the next day. Besides, significant annual seasonal fluctuations can be found in the graph: rental volumes peak during summer months and fall substantially throughout late autumn and winter.

Several isolated sharp spikes may correspond to rare special events, such as strikes or exceptional weather conditions. Lastly, mild weekly oscillations exist, and there's a difference between weekdays and weekends.

2.2. Variable selection and theoretical mechanism

Based on existing bike share time-series literature, weather and temporal conditions (time) are widely recognized as the main factors for daily cycling demand.

To be more specific, weather conditions can directly drive people's willingness to cycle. For instance, higher temperatures often discourage outdoor cycling, and people also choose to go to work by bus on rainy days instead of cycling. Therefore, temperature (t1) and humidity(hum) are included to control weather-related effects.

Temporal factors can also influence people's decisions. On holidays or weekends, people may travel by bike. Furthermore, seasonal shifts produce long-run differences in overall bike share usage. Hence, time-related indicators is_holiday and is_weekend are incorporated into the model.

The dependent variable cnt is daily bicycle hire counts, which is the core outcome we aim to predict.

Besides these raw variables provided in the dataset, two lag indicators are added for this time series research. The weighted lag variable is designed to capture the short-run demand continuity observed in daily rental data. Furthermore, weekend and holiday dummy variables alone cannot fully reflect weekly travel patterns, so introducing lag7_cnt is very important to absorb this seven-day cyclical effect.

2.3. Model construction framework

This study first constructs an unequal-weight rolling lag variable to capture short-term daily continuity, and then uses a series of nested regression models by adding other explanatory factors step by step.

2.3.1. Unequal-weight rolling lag term

Recent days' rental volumes carry more predictive information for time-series forecasting. Therefore, an unequal-weight rolling lag term is constructed, which assigns a higher weight to the most recent daily observation, while distributing the remaining weight evenly across earlier historical days. Its calculation formula is shown in Equation (1):

$$\widehat{cnt_t} = \omega \cdot cnt_{t-1} + \frac{1-\omega}{K-1}\left(cnt_{t-2} + cnt_{t-3} + \ldots + cnt_{t-K}\right) \quad (1)$$

where:

•$\widehat{cnt_t}$ = constructed lag-based proxy for rental volume on day t;

DOI: 10.54254/2754-1169/2026.38458

•cnt_{t-i} = observed actual rental count on day t−i;
•ω = weight assigned to the most recent observation cnt_{t-i}, $0 \leq \omega \leq 1$;
• K = rolling look-back window length, representing the number of historical days used.

To select an appropriate window size K, values ranging from K=2 to K=30 are tested. The in-sample RMSE is computed for each candidate K, and the RMSE trend is examined. The forecasting error reaches its minimum at K=13, so a 13-day window is chosen as the optimal look-back length.

Using this optimised unequal-weight rolling lag term alone yields the simple baseline model M13, which contains no additional control variables.

2.3.2. Step-by-step nested model development

For convenient regression implementation in R, this study adopts a deviation-form variable transformation. Two transformed variables are defined in Equation (2):

$$\mathrm{Yt}^* = \mathrm{cntt} - \mathrm{AVEt} - 1, \quad \mathrm{Xt}^* = \mathrm{cntt} - 1 \ - \mathrm{AVEt} - 1 \tag{2}$$

where $\mathrm{AVEt} - 1$ denotes the average rental volume over the 13-day rolling window.

Regression estimation is performed on these deviation-form variables Yt^* and Xt^* . Taking Model M1 as an illustrative example, its deviation-form regression specification is:

$$Y_t^* = \alpha + \omega X_t^* + \delta is_weekend_t + \varepsilon_t \tag{3}$$

Substituting Equation (2) into Equation (3) and rearranging terms returns the model expressed in the original rental-count scale:

$$cnt_t = \alpha + \omega cntt_{t-1} + (1-\omega) AVE_{t-1} + \delta is_weekend_t + \varepsilon_t \tag{4}$$

This paper applies nested multiple-linear regression to predict daily shared-bicycle rental volume. Instead of feeding all explanatory variables into regression simultaneously, we start from the simple baseline M13 and add groups of explanatory variables sequentially to build a series of nested models:

1. Model M1: M13 + weekend dummy variable
2. Model M2: M1 + holiday dummy variable
3. Model M3: M2 + meteorological variables (temperature, humidity)
4. Model M4: M3 + monthly seasonal dummy variables

Equation (4) gives the prediction specification for Model M1. The same substitution and algebraic rearrangement procedure is applied to derive the original-scale equations for M2, M3 and M4.

2.3.3. Extended model with weekly lag term lag7

The nested baseline model M4 incorporates weekend, holiday, monthly seasonal dummy variables as well as meteorological variables and the unequal-weight rolling lag term. Nevertheless, these explanatory variables mainly capture demand variations within a single week and cannot fully account for cross-weekly periodic fluctuations in shared-bicycle usage.

DOI: 10.54254/2754-1169/2026.38458

Bicycle rental demand typically follows a stable seven-day travel cycle. A preliminary residual diagnostic on Model M4 provides graphical evidence of remaining weekly-driven autocorrelation, as illustrated in Figure 1. Significant periodic spikes are visible at lag-7, lag-14, and lag-21 in the residual ACF plot. This indicates that even after controlling for intra-week seasonal effects, weekly cyclical patterns are not fully captured by M4 and remain in regression residuals.

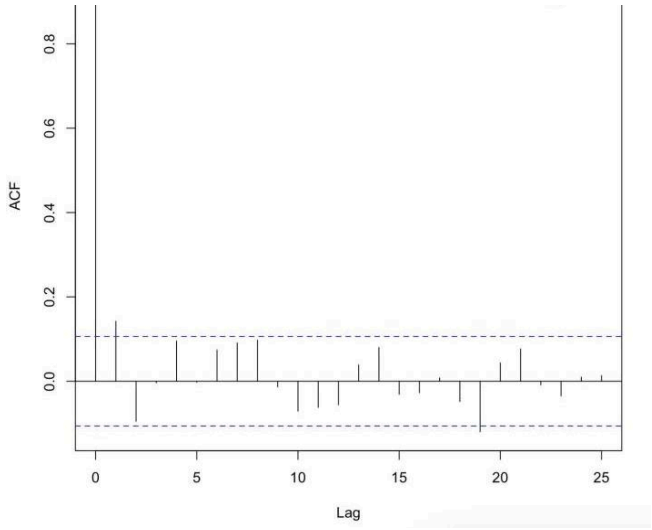


Figure 1. ACF of residuals for M4

Motivated by this pre-diagnostic finding, a weekly lag term, lag7, is introduced into the model specification. This variable corresponds to the rental count observed seven days before the target date, aiming to absorb recurring cross-weekly demand fluctuations. The full specification of the extended model is presented in Equation (5):

$$cnt_t = \alpha + \beta L_t + \gamma T_t + \delta \omega_t + \theta \cdot lag7_cnt + \varepsilon_t \tag{5}$$

L_t = unequal-weight rolling lag term;
T_t = weekend, holiday and monthly dummy variables;
ω_t = temperature and humidity;
α,β,γ,δ,θ = regression coefficients;
ε_t = residual term.

2.4. Model evaluation design

2.4.1. In-sample evaluation

In-sample RMSE measures the model's fitting degree on training data, calculated as the root mean squared deviation between fitted and actual rental counts. For nested linear regression models, adding any explanatory variable mechanically reduces in-sample error, which creates a misleading illusion of improved model performance, and it cannot reflect generalisation ability on unseen data. Hence, in-sample RMSE is only treated as an auxiliary reference rather than the core evaluation standard.

DOI: 10.54254/2754-1169/2026.38458

2.4.2. Out-of-sample evaluation: static & rolling recursive forecast

The out-of-sample test relies on independent, unseen samples to identify overfitting risks; two forecasting frameworks are designed in this paper.

First, static split-sample forecast. To be more specific, the full dataset is segmented by chronological order instead of random sampling to preserve time-series sequence features. All observations before 2 May 2016 serve as the training set, and samples from May 2016 to January 2017 form the static test set. Besides, model coefficients are estimated once and fixed for all prediction periods.

Second, a rolling recursive forecast with a fixed 365-day training window is used. The estimation window slides forward one day per prediction round. In other words, the earliest historical observation is eliminated, and the latest daily rental data is incorporated to re-estimate all coefficients for one-step-ahead prediction. This framework simulates real-time daily forecasting scenarios and adapts to time-varying travel demand characteristics.

2.4.3. Residual diagnostic tools

Three complementary residual diagnostic methods are adopted to test OLS assumption validity:

1. Residual time-series plot: visually judge whether residuals fluctuate randomly around zero and detect periodic oscillation patterns;

2. Normal Q-Q plot & residual histogram: examine residual distribution, identify heavy-tailed deviation from Gaussian normality;

3. ACF graph and Ljung-Box hypothesis test: qualitative + quantitative inspection of residual serial correlation, the core misspecification concern of this study

3. Empirical results and analysis

3.1. In-sample fitting results of nested models

3.1.1. Change trend of in-sample RMSE

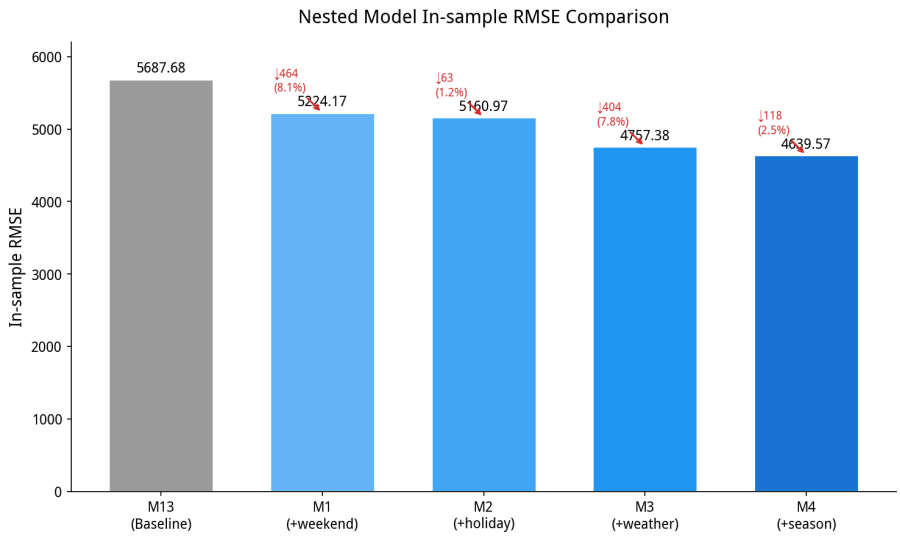


Figure 2. RMSE value for M13, M1, M2, M3, M4

DOI: 10.54254/2754-1169/2026.38458

Figure 2 compares in-sample RMSE values for each model. Beginning with the baseline M13, training-set error falls as weekend, holiday, weather, and monthly seasonal dummy variables are added step-by-step. M13 produces the largest in-sample RMSE at 5687.68, whereas M4, which adds monthly seasonal controls, achieves the lowest error of 4639.57.

For nested linear regressions, adding more explanatory variables will naturally decrease the value of RMSE. In other words, a falling in-sample RMSE does not mean the model can predict new observations accurately. Therefore, it is necessary to assess performance using out-of-sample tests.

3.1.2. In-sample RMSE and coefficient significance analysis

Table 1. Coefficient estimates and significance for M13-M4

Variable	M13	M1	M2	M3	M4
Intercept		1391.20***	1576.60***	1919.30***	21672.00***
Xt*(weighted lag)		0.4157***	0.4001***	0.3351***	0.3078***
is_weekend		-5048.10***	-5260.09***	-5207.10***	-5250.40**
is_holiday			-5500.05**	-5187.50**	-4991.80***
t1(temperature)				-95.59*	196.62***
hum(humidity)				-276.63***	-263.92***
Month dummies	No	No	No	No	Yes
In-sample RMSE	5687.68	5224.17	5160.97	4757.38	4639.57

Note: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

As shown in Table 1, the weighted lag indicator Xt* is highly significant in all five models, which shows that historical daily rental volume can effectively predict current demand.

Weekend and holiday dummy variables always carry negative coefficients. Daily bike orders fall when people do not need to commute regularly.

Humidity has a clear negative influence in Model 3 and Model 4. Higher humidity discourages residents from riding bikes. Besides, Temperature only shows a significant positive correlation after monthly seasonal controls are added.

Most monthly dummy terms in Model 4 pass the significance test, which proves obvious seasonal fluctuation exists in rental data. Even though adding seasonal variables lowers training-set error, it does not guarantee better prediction on new data. The following out-of-sample results will verify model generalization ability.

3.1.3. Static out-of-sample overfitting test

Table 2. Comparison of in-sample and static out-of-sample RMSE

Model	Specification	In-sample RMSE	Static out-of-sample RMSE
M13	Baseline model, only weighted lagged term Xt*	5687.68	5663.75

DOI: 10.54254/2754-1169/2026.38458

Table 2. (continued)

M1	M13 + is_weekend (weekend dummy variable)	5224.17	5502.85
M2	Model 1 + is_holiday (holiday dummy variable)	5160.97	5403.29
M3	Model 2 + temperature and humidity	4757.38	5062.66
M4	Model 3 + monthly seasonal dummy variables	4639.57	5372.07

Table 2 compares in-sample and static out-of-sample prediction errors across nested specifications. As variables are gradually added, in-sample RMSE keeps decreasing, and Model 4 achieves the smallest training-set error. Nevertheless, its static out-of-sample RMSE rises sharply. The monthly seasonal dummies improve fitting on training data but cause obvious overfitting, and the model loses forecasting performance on unseen test data. Model 3 delivers the lowest static out-of-sample RMSE at 5062.66. Without monthly seasonal dummies, it shows better generalization ability for static forecasting. Therefore, Model 3 is selected as the preferred static prediction model.

3.2. Comparative analysis of static vs rolling recursive forecasting

Table 3. Comparison of in-sample, static and rolling-window out-of-sample RMSE

Model	In-sample RMSE	Static Out-of-Sample RMSE	Rolling Window RMSE
M13	5687.68	5663.75	3128.90
M1	5224.17	5502.85	3668.77
M2	5160.97	5403.29	3722.54
M3	4757.38	5062.66	4226.54
M4	4639.57	5372.07	4655.13

Rolling-window forecasting repeatedly re-estimates model coefficients as new observations become available. As shown in Table 3, all models achieve much lower prediction errors under the rolling-window scheme compared with static forecasting.

Unlike static forecasting results, the simplest specification M13 yields the smallest rolling-window RMSE of 3128.90. Adding more explanatory variables increases rolling-window prediction error. Under recursive one-step-ahead forecasting, complex models with more dummy variables do not bring forecasting gains.

This contrast reflects that when coefficients are updated day-by-day, the simple weighted-lag-only model captures time-varying features well. For rolling-window real-time prediction tasks, M13 is the preferred choice.

3.3. Residual misspecification diagnostics for baseline models (M13/M3/M4)

3.3.1. Visual diagnostics: residual time series & distribution characteristics

Models in this paper are constructed following a nested variable-adding procedure. Reporting residual diagnostics for every nested model would produce redundant figures and distract from the

DOI: 10.54254/2754-1169/2026.38458

key empirical patterns. For this reason, this section selects three representative baseline models: M13, M3 and M4.

These three models adopt different sets of explanatory variables. M13 is the simplest specification containing only the unequal-weight rolling lag term. M3 further incorporates weekend, holiday, and meteorological controls. M4 is the full baseline model with additional monthly seasonal dummy variables. Collectively, they represent minimal, intermediate, and full-variable configurations.

Comparing residuals across the three models helps us examine whether expanding conventional explanatory variables can eliminate systematic residual patterns. In other words, we investigate whether systematic residual behaviour persists even as we enrich the model with different groups of control variables.

Models M1 and M2 lie between M13 and M3 in variable composition. Their residual characteristics yield no extra meaningful information, so their diagnostic outputs are omitted for conciseness. The residual properties of these baseline models will be analysed using visual and statistical diagnostic tools in the following part.

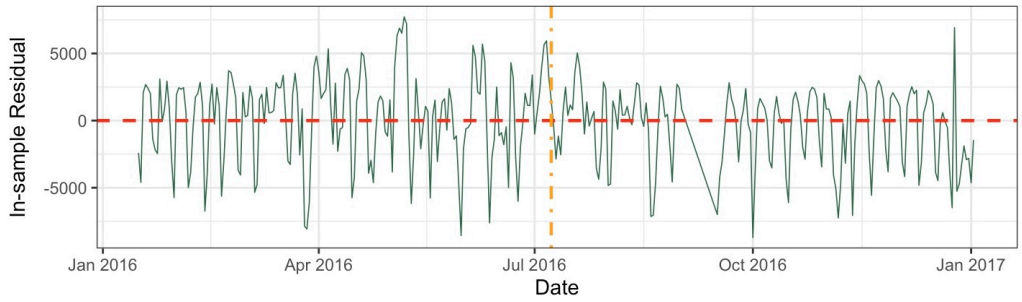


Figure 3. Residual time series graph of M13

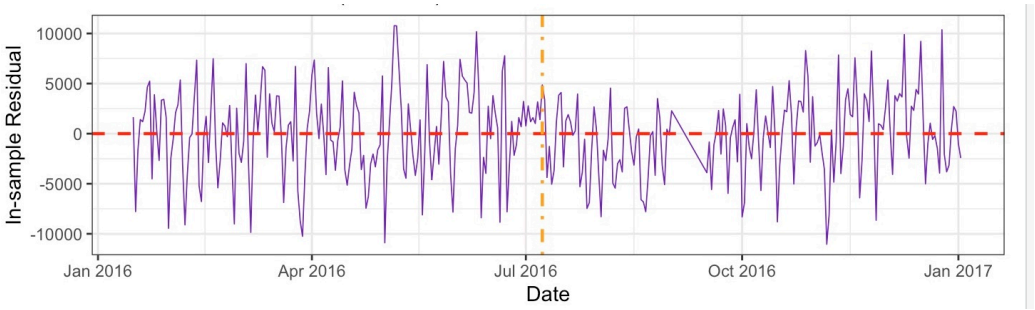


Figure 4. Residual time series graph of M3

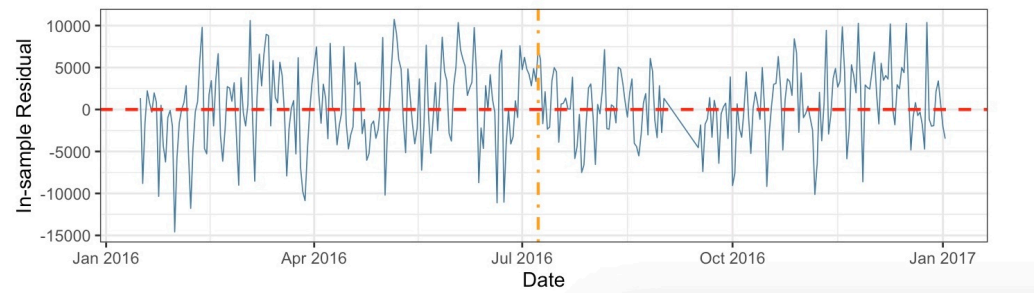


Figure 5. Residual time series graph of M4

DOI: 10.54254/2754-1169/2026.38458

Residual diagnostics are conducted for baseline Model 1, Model 3, and Model 4 . Figures 3, 4,5 indicate that residuals of each model fluctuate around zero without persistent one-sided deviation, which suggests the baseline specification does not generate long-run systematic fitting bias. Nevertheless, visual inspection of the time series cannot identify latent weekly serial dependence. According to the ACF results presented in Chapter 2, baseline residuals exhibit significant 7-order serial correlation. This implies that weekly cyclical components are not fully captured, and residuals fail to satisfy the white-noise property.

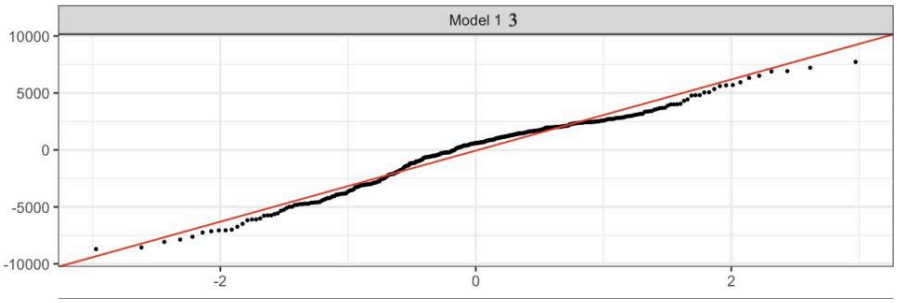


Figure 6. Normal QQ plots for model 13

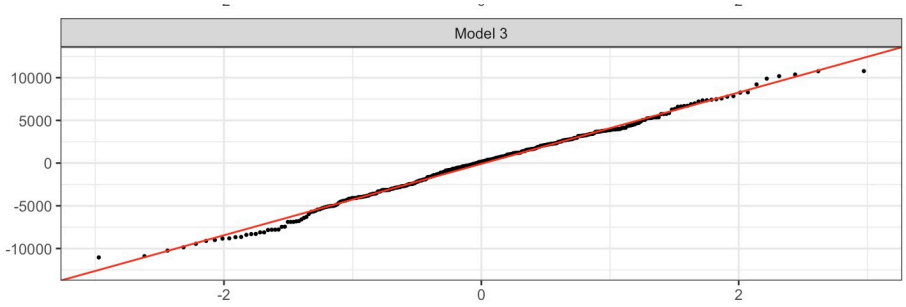


Figure 7. Normal QQ plots for model 3

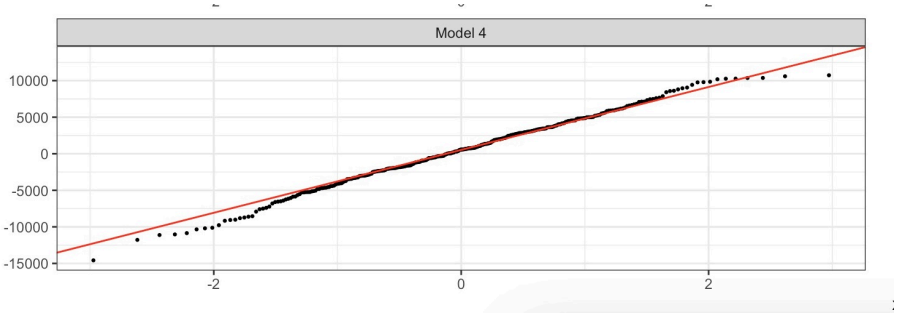


Figure 8. Normal QQ plots for model 4

Figure 7 and Figure 8 present normal Q-Q plots of residuals for Model 3 and Model 4. The horizontal axis corresponds to theoretical quantiles from the standard normal distribution, and the vertical axis represents sample quantiles of model residuals.

Most sample points roughly follow the red reference line in the middle range. However, deviations appear at both tails: points depart from the straight line in the left tail (negative residuals) and right tail (large positive residuals). This indicates that model residuals are approximately but not perfectly normally distributed. Heavy tails exist, meaning the model generates some large positive and negative error values.

DOI: 10.54254/2754-1169/2026.38458

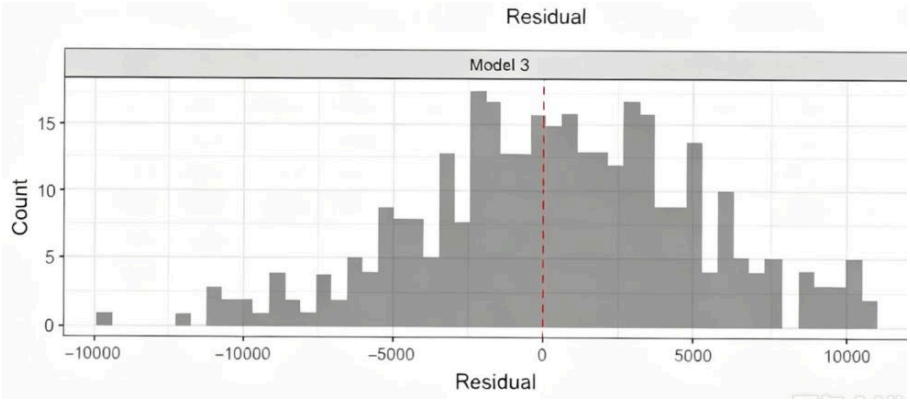


Figure 9. Residual histograms for model 3

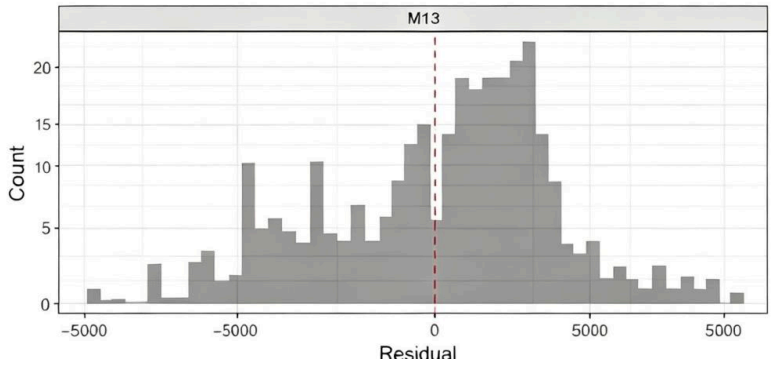


Figure 10. Residual histograms for model 13

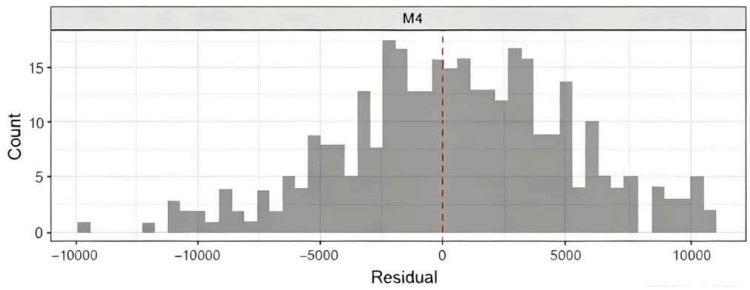


Figure 11. Residual histograms for model 4

Residual histograms complement the information obtained from Q-Q plots. The overall shape approximates a bell-shaped density centred near zero. Nevertheless, comparatively large counts persist in the far-left and far-right bins, confirming the heavy-tail behaviour identified in Q-Q diagnostics. Such tail observations are typical for high-frequency time-series subject to calendar-driven shocks and irregular events.

3.3.2. Core misspecification: persistent weekly serial correlation

Serial correlation and non-normality represent separate model misspecification issues. Serial correlation violates Gauss-Markov assumptions and distorts standard errors as well as t-statistics, which constitutes the major specification concern in this analysis. Under large-sample conditions, non-normal residuals do not invalidate the consistency of OLS estimators, yet confidence interval

DOI: 10.54254/2754-1169/2026.38458

performance may be affected. Combining time-series plots, distribution graphs, and statistical tests, baseline models suffer from misspecification, and lagged terms are required to absorb weekly cyclical variation. As a result, lag7tis added.

3.4. Effectiveness test of weekly lag term lag7

3.4.1. Visual comparison: residual series and ACF after adding lag7

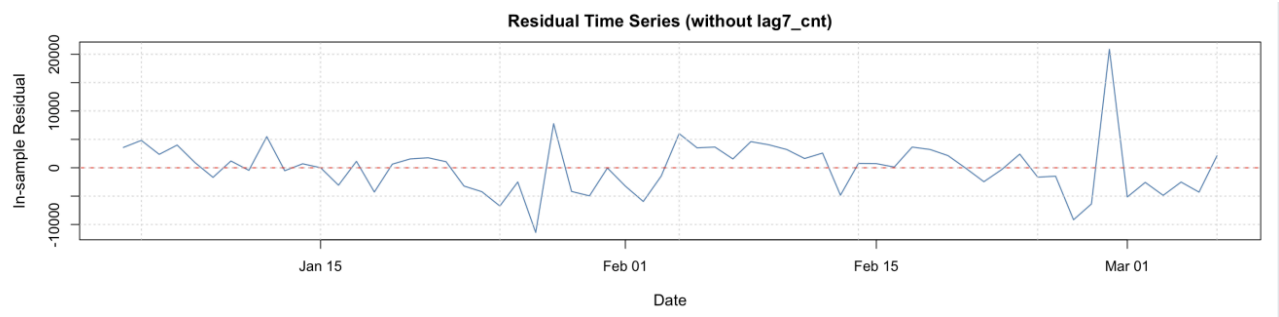


Figure 12. Residual time series for M4(without lag7)

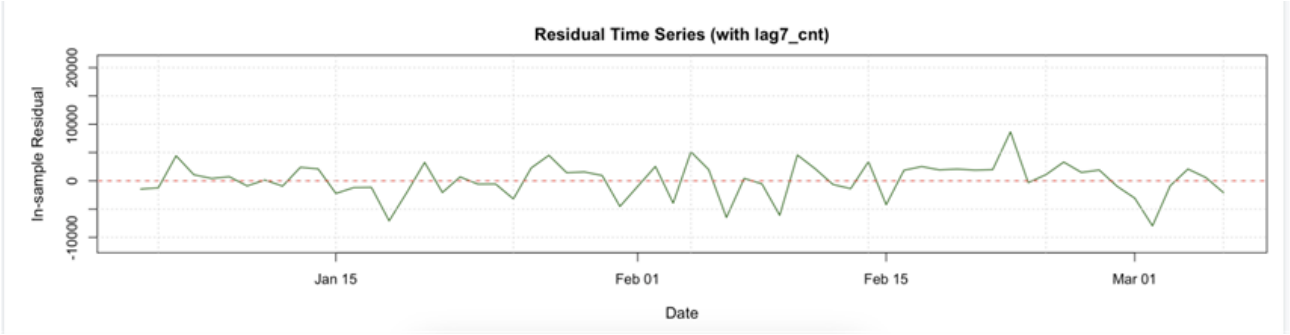


Figure 13. Residual time series for M4 withlag7

Figure 12 exhibits visible periodic oscillating patterns, alongside occasional large-magnitude outliers. After adding lag7 to obtain M4-L7, according to Figure 13, residual fluctuations remain centred around zero. Yet the systematic cyclical movement observed in baseline M4 is substantially reduced. Residuals behave in a more random fashion, and several prominent large-amplitude spikes are attenuated.

This visual evidence reinforces the findings from the ACF diagnostics: the weekly lag term lag7 mitigates the cyclical residual behaviour originating from weekly calendar effects. Nevertheless, occasional large residual deviations still exist, consistent with the heavy-tail feature shown in earlier histogram and Q-Q results.

As documented in Chapter 2, the residual ACF of baseline M4 reveals a notable spike at lag7, indicating leftover weekly cyclical variation. Among all specifications considered, only M4-L7 — the modified variant of M4 — incorporates the weekly lag term lag7. M13, M3, and the original M4 remain baseline set-ups without this lag adjustment.

DOI: 10.54254/2754-1169/2026.38458

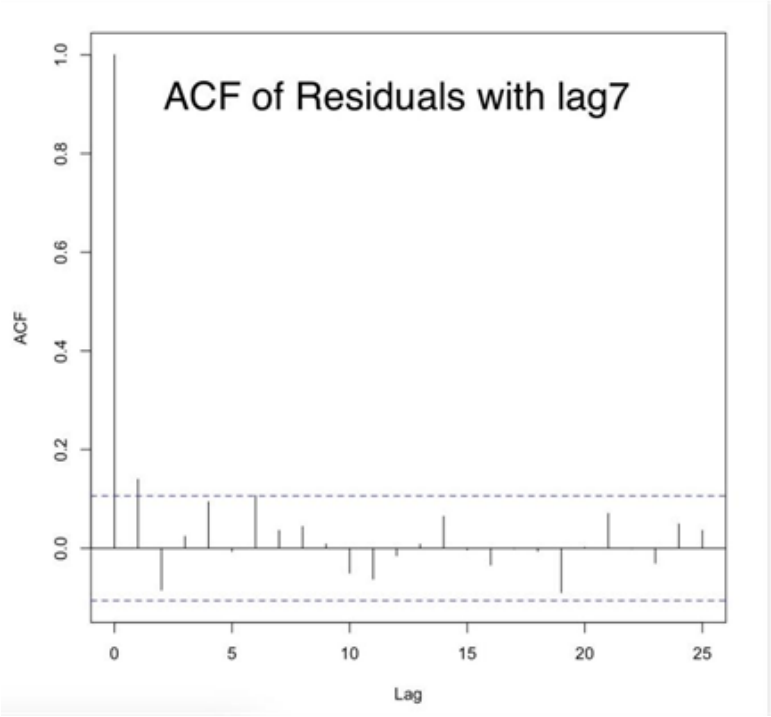


Figure 14. ACF plot of residuals for M4L7 (M4 augmented with lag7)

M4 is the full-variable baseline specification with the richest set of predictors. Adding lag7 solely to M4 keeps its complete predictor set unchanged, which allows us to isolate the marginal effect of the weekly lag term.

After introducing lag7 within M4-L7, Figure 14 shows that the prominent lag7 autocorrelation spike disappears. Autocorrelation coefficients across all displayed lags fall inside the confidence band. This demonstrates that the lag term effectively removes the major serial-correlation problem detected in the baseline M4. The comparison also delivers a key message that although M13, M3 and M4 adopt different explanatory-variable combinations, merely adjusting or expanding predictor sets cannot eliminate weekly residual autocorrelation.

Still, residual heavy-tail features are only partially mitigated. Driven by calendar-related shocks and irregular events inherent in high-frequency observations, full normality cannot be attained even in M4-L7.

3.4.2. Quantitative test: Ljung-Box serial correlation statistics && out-of-sample forecasting improvement via lag7

Table 4. Ljung-Box test statistics && rolling out-of-sample RMSE for M4 and M4-L7

Indicator	Baseline M4	Extended M4-L7
LB(7) Statistic	18.066	16.976
LB(7) p-value	0.0117*	0.0175*
LB(14) Statistic	28.582	21.562
LB(14) p-value	0.0119*	0.0881

DOI: 10.54254/2754-1169/2026.38458

Table 4. (continued)

LB(21) Statistic	38.114	26.750
LB(21) p-value	0.0125*	0.1793
Rolling out-of-sample RMSE	4019.22	3710.44

Note: * indicates significance at the 5 % level ($p < 0.05$). Baseline M4 is the model without lag7, while M4-L7 includes the weekly lag term.

Table 4 reports Ljung-Box test statistics and rolling out-of-sample RMSE for the baseline M4 and its augmented version M4-L7. For baseline M4, Ljung-Box statistics at lags 7, 14, and 21 are all significant at the 5 % level, which rejects the null hypothesis of no residual serial correlation and points to strong serial correlation induced by weekly periodic patterns.

Following the inclusion of lag7, the Ljung-Box statistic at lag 7 remains statistically significant, implying weak residual autocorrelation at the weekly frequency. In contrast, test statistics at lags 14 and 21 turn insignificant, with p-values increasing to 0.0881 and 0.1793, respectively. This finding indicates that longer-range serial correlation stemming from weekly cycles is largely mitigated, even though lag-7 autocorrelation cannot be fully removed.

The rolling out-of-sample RMSE declines from 4019.22 for baseline M4 to 3710.44 for M4-L7. The reduced prediction error shows that the weekly lag term delivers tangible forecasting gains despite the remaining mild autocorrelation at lag-7.

4. Conclusion

This study focuses on the prediction of the daily rental volume of shared bicycles. Based on the combination of various variables and the prediction mechanism, a model is constructed for comparison. Further, the applicability of the model is analyzed from the perspectives of prediction accuracy, residual characteristics, model complexity, and practical application requirements. A comprehensive evaluation is conducted. The research results show that the demand for shared bikes not only has obvious short-term continuity, but also has weekly cycle and seasonal characteristics, and is affected by factors such as dates and holidays. Therefore, merely pursuing the model's fitting effect on past samples cannot guarantee the performance of actual predictions. Evaluation must be conducted in combination with out-of-sample prediction and rolling window prediction [14].

Firstly, the main advantages of Model 13 lie in its simple structure, low computational cost, and strong anti-overfitting ability. This model mainly relies on the application of weighted disparity variables to capture the most recent rental demands, thus enabling rapid predictions using the latest data. In the rolling Windows prediction, Model 13 demonstrates a powerful predictive ability and is suitable for production scenarios where the prediction results need to be updated frequently. The drawbacks are also very obvious. Due to the scarcity of control variables, the marginal impact of weather, holidays, and seasonal changes on rental demand cannot be fully explained, and the results of long-term and static forecasts are relatively fragile.

Therefore, the 13 models include pre-configuring vehicles before morning and evening commutes, arranging vehicles between subway stations and residential areas, and rescheduling emergency vehicles in case of sudden rain, major events, or public transportation suspension, among other large bike-sharing platforms for a day [15]. It is more suitable for regular vehicle dispatching. Meanwhile, optimizing the grid-based maintenance personnel deployment and scheduling routes can also be used to reduce no-load and inefficient scheduling. For large platforms that require rapid

DOI: 10.54254/2754-1169/2026.38458

iteration and real-time decision-making, this "simple and fast" model has high practical value. For instance, it can be applied to large-scale platform scenarios, such as Meituan Bike, HelloBike, and Qiguo Bike [16].

Secondly, the feature of Model 3 is that the variable structure is relatively simple, while maintaining key influencing factors such as weekends, holidays, temperature, and humidity. Compared with Model 13, it has better interpretability and helps operators understand how multiple factors affect the demand for shared bikes. In addition, this model only requires one parameter estimation for long-term demand forecasting, thereby reducing computing and data storage costs. However, the biggest problem with Model 3 is that the residuals still have obvious periodic autocorrelation, which greatly reduces the accuracy of rolling dynamic prediction.

Therefore, it is not suitable for operational tasks that require daily re-forecasting and planning adjustments, but is more appropriate for infrequent demand forecasting, such as annual or quarterly demand forecasting. For example, the company can use Model 3 for rough planning of budget, personnel allocation, and vehicle purchase [17]. This model shows obvious characteristics during peak and off-peak hours, but it also has strong applicability for shared bike services such as tourist attractions where precise daily demand prediction is not required.

Finally, the models with the extension M4 or Lag7 achieved a better balance between their analytical capabilities and short-term predictive abilities. We integrated weather, holidays, and monthly seasonal variables into the existing model, but the residual diagnosis still showed a considerable amount of 7th-order autocorrelation and found that the model failed to fully capture the weekly characteristics of shared bicycle demand. After adding Lag7, the periodicity within the residuals was significantly reduced, the Ljung-Box test results improved, and at the same time, the predicted RMSE also decreased [18]. Therefore, Lag7 is not merely about adding variables; Instead, it captures the repetitive operation patterns during the day from the perspective of time series structure.

From the perspective of practical application, the enhanced M4 is suitable for medium-sized community operators with specific analytical requirements but without the need for high-frequency real-time re-evaluation of the model. In addition to being used for monthly vehicle ownership, recycling, and purchasing based on seasonal demand, it can also be linked with weather forecasts for dynamic pricing and coupon distribution, etc. In addition, companies with limited computing resources can re-estimate their models every 2 to 3 days to balance prediction accuracy and computing costs. The corresponding typical scenarios include regional shared bike services such as campuses and scenic spots.

Overall, this study indicates that there is no "absolutely optimal" model in all cases. Model 13, Model 3, and Extension m4 respectively represent three different modeling concepts: "real-time performance", "long-term planning," and "balance". If an enterprise values real-time scheduling and rapid response the most, it should give priority to Model 13. If the requirements planning is mainly operated at the annual or quarterly level, Model 3 has a lower computational cost and higher interpretability. If it needs to analyze demand drivers such as weather, holidays, and seasons, as well as more reliable short-term predictions, the extended M4 is more suitable.

Therefore, the core conclusion of this study is not simply to seek a unified optimal model, but to select a model suitable for an enterprise based on its scale, prediction frequency, computing resources, and business goals. This model selection method, based on business scenarios, enables the demand prediction of shared bicycles, further assisting vehicle scheduling, resource allocation, and operational decision-making, from simple statistical models to predictive models in practical applications. It can reflect value.

DOI: 10.54254/2754-1169/2026.38458

References

[1] Department for Transport. (2022). Local transport plan guidance. *GOV.UK*. https://www.gov.uk/government/publications/local-transport-plans

[2] Transport for London. (n.d.). Santander Cycles open cycling dataset [Dataset]. https://cycling.data.tfl.gov.uk/

[3] Gammelli, D., Rodrigues, F., Prata, C., & Pereira, F. C. (2022). Predicting bike-sharing demand to improve inventory management. *Transportation Research Part C: Emerging Technologies, 138*, 103602.

[4] Gebhart, K., & Noland, R. B. (2014). The impact of weather conditions on bikeshare trips in Washington, DC. *Transportation, 41*(6), 1205–1225.

[5] El-Assi, W., Mahmoud, M. S., & Habib, K. N. (2017). Effects of built environment and weather on bike sharing demand: A case study of Toronto. *Transportation Research Part D: Transport and Environment, 51*, 336–347.

[6] Morton, C. (2020). The effect of weather and air quality on demand for bike sharing in London. *Transportation Research Part D: Transport and Environment*.

[7] Faghih-Imani, A., Hampshire, R., Marla, L., & Eluru, N. (2014). An empirical analysis of bike sharing usage and rebalancing: Evidence from Barcelona and Seville. *Transportation Research Part A: Policy and Practice, 64*, 79–95.

[8] Eren, E., & Uz, V. E. (2020). A review on bike-sharing: The factors affecting bike-sharing demand. *Sustainable Cities and Society, 54*, 101882.

[9] Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.). Wiley.

[10] Sathishkumar, V. E., Park, J., & Cho, Y. (2020). Using data mining techniques for bike sharing demand prediction in metropolitan city. *Computer Communications, 153*, 353–366.

[11] Li, X., et al. (2023). Bike-sharing demand forecasting based on time-series and deep learning methods. *Transportation Research Part C: Emerging Technologies*.

[12] Bigazzi, A. Y., & Wong, K. I. (2021). How does weather affect bikeshare use? A comparative analysis of forty cities across climate zones. *Journal of Transport Geography, 95*, 103155. https://doi.org/10.1016/j.jtrangeo.2021.103155

[13] Transport for London. (n.d.). Santander Cycles open cycling dataset [Dataset]. Available from: https://cycling.data.tfl.gov.uk/

[14] El-Assi, W., Mahmoud, M. S., & Habib, K. N. (2017). Effects of built environment and weather on bike sharing demand: A station-level analysis of commercial bike sharing in Toronto. *Transportation, 44*, 589–613. https://doi.org/10.1007/s11116-015-9669-z

[15] Gebhart, K., & Noland, R. B. (2014). The impact of weather conditions on bikeshare trips in Washington, DC. *Transportation, 41*, 1205–1225. https://doi.org/10.1007/s11116-014-9540-7

[16] Transport for London. (n.d.). Santander Cycles. https://tfl.gov.uk/modes/cycling/santander-cycles

[17] Transport for London. (n.d.). Our open data. https://tfl.gov.uk/info-for/open-data-users/our-open-data

[18] London Datastore. (n.d.). Number of bicycle hires. Greater London Authority. https://data.london.gov.uk/dataset/number-of-bicycle-hires-2r84d