Proceedings of CONF-MLA 2026 Symposium: Learning and Decision Making in Multi Agent Software Systems
DOI: 10.54254/2755-2721/2026.BA36232

New York City Taxi Order Volume Forecasts Based on SARIMA and SVR Models

Yirui Mao

School of Computer Science and Technology, Zhejiang University of Water Resources and Electric Power, Hangzhou, China
2023b15025@stu.zuwe.edu.cn

Abstract. Short-term taxi order forecasting is a key reference for dynamic allocation of urban transportation capacity. New York taxi passenger flow exhibits significant intraday cyclical fluctuations, making it difficult for manual dispatch to predict changes in passenger flow. This paper takes hourly taxi order time-series data in New York City as the research object, selecting the Seasonal Autoregressive Integrated Moving Average (SARIMA) model and the Support Vector Regression (SVR) model to conduct a short-term order prediction comparison. Based on AIC and BIC criteria, the optimal SARIMA parameters are traversed and screened, and SVR data are subjected to Min-Max normalization. Both types of models use multi-step rolling forecasting to generate a complete 7-day prediction sequence, relying on MAE, RMSE, MAPE, and R2 indicators to quantify prediction accuracy. Experimental results show that the optimal SARIMA (1,0,1)(1,1,24) model can accurately fit peak and valley fluctuations in passenger flow, with an R2 of 0.815. SVR models can only capture overall trends and lack the ability to characterize extreme passenger flows. The study demonstrates that for taxi time series data with strong intraday cycles, the SARIMA model offers better prediction stability and accuracy, providing a reference for taxi capacity scheduling.

Keywords: Short-term traffic flow prediction, SARIMA model, SVR model

1. Introduction

Taxis are an important part of the urban public transportation system. It meets the needs of flexible, convenient, and personalized travel services, serving as an important basis for traffic operation management, capacity optimization, and service improvement. Against the backdrop of rapid development of big data and intelligent transportation technologies, conducting research on taxi travel demand has become one of the key directions in the field of transportation science.

In the field of traffic demand forecasting, time-series models, machine learning models, and deep learning models are the main research models, each with its own suitable scenarios and strengths and weaknesses [1]. Among them, time series models represented by ARIMA and SARIMA can effectively capture the trend and periodicity patterns of time series data and are widely used in short-term traffic flow prediction [2]. Machine learning models represented by SVR can effectively mine complex nonlinear correlations in data and adapt to multi-feature fusion prediction scenarios.

DOI: 10.54254/2755-2721/2026.BA36232

Currently, scholars have conducted extensive research on forecasting demand for urban passenger transport and ride-hailing services. Pengcheng Li used four traditional machine learning methods—multiple linear regression, support vector regression, random forest, and the XGBoost algorithm—to predict the number of taxi passengers at Kennedy Airport, and analyzed the predictive performance of different algorithmic models [3]. Kuang Dongyu established short-term prediction models for ride-hailing demand based on ARIMA, single-feature LSTM, and multi-feature LSTM for comparison. Experimental results verify that the multi-feature LSTM model delivers the best prediction results for short-term forecasting of ride-hailing demand [4]. Li Xinlu constructed ARIMA, SARIMA, and LSTM algorithm models to predict demand in taxi big data hotspot areas over time series, conducted comparative analysis, and concluded that LSTM model prediction results are more suitable for taxi big data hotspot area demand forecasting [5]. Ni Jie and colleagues applied the ARIMA model to rail transit passenger flow forecasting, verifying its good predictive performance in smooth travel time series data [6]. Zhang Liyan and colleagues selected ARIMA, SARIMA, and LSTM models for traffic flow prediction, compared the prediction effects of the three models, and ultimately demonstrated that the LSTM model has better prediction accuracy and practicality compared to ARIMA and SARIMA [7]. These studies demonstrate that different prediction models have different adaptation scenarios and advantages/disadvantages.

Based on this, this paper selects hourly taxi order time-series data and uses the SARIMA and SVR models for research to compare and analyze the prediction performance of the two models. The research aims to provide references for urban taxi operation management and traffic optimization, as well as for forecasting related urban transportation demand.

2. Methods

2.1. Data source and index selection

This study uses the New York City taxi trip dataset (2025 sample) published on the Kaggle platform, with the original data provided by the New York City Taxi and Limousine Administration (TLC), and has completed the basic cleaning [8]. This study selects order data from August 1 to August 28, 2025 as the research period, focusing on the boarding time to carry out the following preprocessing work. First, aggregate order volume by hourly granularity, that is, count the order volume within each hour. Second, fill the no-order periods with zeros by padding to form a continuous time series. Finally, the 1% and 99% quantiles were truncated to remove extreme outliers to reduce model interference.

The core variable of this study is the number of hourly taxi orders. Based on preprocessed data, it is divided into training and test sets. 504 sets of data from August 1 to August 21, 2025, were used as training samples for model fitting. A total of 168 data sets from August 22 to August 28, 2025, were used as the test set for model performance evaluation.

2.2. Method introduction

To explore the adaptation effects of different algorithm models in hourly taxi order timing prediction, this paper conducts a comparative study combining traditional time-series statistical models with machine learning models, constructing two types of models. Multiple sets of SARIMA seasonal timing series models with different parameters were set up, and an SVR model was introduced to compare the prediction performance of different algorithms on taxi travel data from both seasonal timing and machine learning dimensions.

DOI: 10.54254/2755-2721/2026.BA36232

2.3. SARIMA model

The Seasonal Autoregressive Integrated Moving Average Model (SARIMA) is specifically designed to fit time series with fixed periodic characteristics, making it highly suitable for data with obvious intraday fluctuations in traffic and ride-hailing orders. Its standard expression is SARIMA(p,d,q)(P, D, Q, S). In the formula, p, d, q are the autoregressive order, difference order, and moving mean order of the non-seasonal part, respectively; P, D, and Q correspond to the related order of the seasonal part; s represents the length of the sequence's seasonal cycle. Compared to the ARIMA model, this model adds a seasonal fitting module that can extract the cyclical variation patterns of time series data and capture the characteristics of order volume during day-night peaks and troughs. In this paper, the cycle parameters are determined according to the 24-hour daily cycle law reflected in the initial time sequence diagram of the data set, and multiple groups of different order combinations are set to carry out research to screen the SARIMA model with the best comprehensive performance.

2.4. SVR model

Support vector regression (SVR) is a typical nonlinear machine learning model, which maps data to high-dimensional space with the help of a kernel function and mines the internal nonlinear Association of data. The time series samples are reconstructed by a sliding window during modeling, and the model does not need to set the cycle rule manually. The core advantage of the SVR model is that, based on the principle of structural risk minimization, the SVR model can efficiently process small-sample, high-dimensional, and strongly nonlinear traffic data with core techniques, and has excellent generalization ability and anti-noise robustness.

3. Results and discussion

3.1. SARIMA model prediction analysis

3.1.1. Sequence stability test

The time series model has strict requirements on the stability of the input sequence. First, the original order sequence needs to be tested by the ADF unit root test. The obtained ADF statistic is -4.9025, and the p-value is 0.0000. The test results show that the p-value is less than 0.05, rejecting the original assumption that there is a unit root, indicating that the sequence itself is a stationary sequence and does not need difference processing.

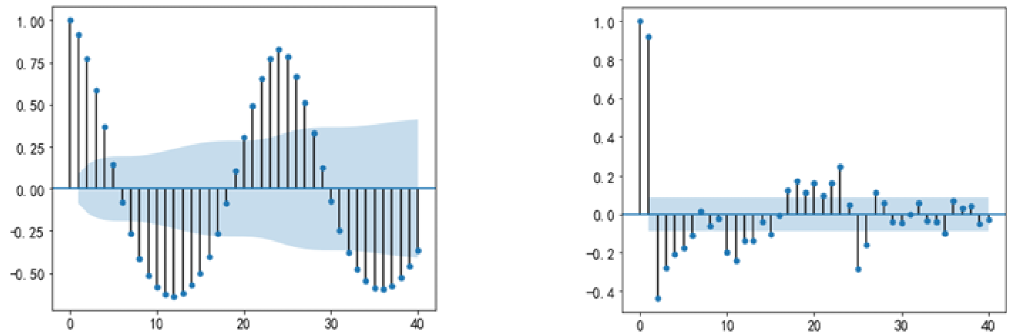
3.1.2. Parameter setting

From the autocorrelation function (ACF) graph of the sequence, it can be observed that the autocorrelation coefficient has a significant peak at the lag of 24 orders, indicating that the order sequence has obvious daily periodicity, and the cycle length is 24 hours. For this kind of time series data with fixed periods, this paper uses the seasonal autoregressive integrated moving average model (SARIMA) to model the intraday periodic fluctuations through the exclusive seasonal term structure.

The general form of the SARIMA model is SARIMA (P, D, q) (P, D, Q, s). The parameters of the non-seasonal term are (P, D, q). Combined with the results of the ADF test, the difference order d=0 is determined. In this paper, the non-seasonal order is searched in the range of 0, 1, 2, 3. The

DOI: 10.54254/2755-2721/2026.BA36232

parameters of the seasonal term are (P, D, Q, s). Combined with the characteristics of the data daily cycle, set the seasonal cycle s=24 and the seasonal difference order d=1 to eliminate the seasonal trend. The characteristics of the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the reference sequence, as shown in Figure 1, limit the seasonal autoregression order P and the seasonal moving average order q to 0 and 1.



(a) Autocorrelation (b) Partial Autocorrelation

Figure 1. Results of ACF and PACF (picture credit: original)

In this paper, AIC and BIC double minimum criteria are used to complete the global traversal optimization within the preset parameter interval, and the other two groups of representative parameter combinations are selected for comparative study. Among them, SARIMA (3,0,3) (1,1,1,24) is the optimal model of the training set selected according to AIC and BIC criteria. SARIMA (1,0,1) (1,1,1,24) uses low-order non-seasonal parameters and is the basic comparison model. SARIMA (2,0,1) (1,1,0,24) adjusts the non-seasonal term and seasonal moving average order to analyze the impact of seasonal structure on the prediction effect.

3.2. Analysis of prediction results

Table 1 shows the predictive performance indicators of each SARIMA model on the test set.

Table 1. Comparison results of prediction performance of different SARIMA models

Model	MAE
SARIMA(2,0,1)(1,1,0,24)	93.38
SARIMA(3,0,3)(1,1,1,24)	65.29
SARIMA(1,0,1)(1,1,1,24)	64.07
Model	MAE
SARIMA(2,0,1)(1,1,0,24)	93.38

Based on the double minimum criteria of AIC and BIC, the optimal training model in the ergodic range is SARIMA (3,0,3) (1,1,1,24), but its generalization prediction performance is not optimal. The test results show that the R2 of SARIMA (1,0,1) (1,1,24) and SARIMA (3,0,3) (1,1,1,24) are more than 0.81, which can effectively capture the intraday periodicity of the order sequence. Among them, SARIMA (1,0,1) (1,1,1,24) has the best comprehensive performance of MAE, RMSE, and R2, and has stronger generalization ability. In contrast to SARIMA (2,0,1) (1,1,0,24), the seasonal

DOI: 10.54254/2755-2721/2026.BA36232

moving average term is missing, and R2 is only 0.545, indicating that the complete seasonal structure is the key factor in fitting the periodic fluctuation.

AIC and BIC are calculated based on the fitting error of the training set and the complexity of the model. Although the high-order SARIMA (3,0,3) (1,1,1,24) is better in the training stage, its parameter complexity is higher, and there is slight overfitting, resulting in a slightly poor generalization effect on the test set.

In this study, SARIMA (1,0,1) (1,1,1,24) was selected as the optimal seasonal prediction model. According to the comparison chart between the actual value and the predicted value of the order quantity finally obtained by the model, the prediction curve is highly consistent with the real sequence, which can accurately capture the intraday periodic fluctuation of the order quantity and has a good fitting effect on the peak and trough, without obvious phase shift or trend deviation. There is only a small error at individual extreme peaks, which does not affect the overall prediction performance.

4. SVR model prediction analysis

4.1. Model training

SVR is a nonlinear regression algorithm based on a kernel function. It does not need data to meet the stationarity assumption and can fit complex nonlinear relationships. In this paper, the 24-hour sliding window method is used to construct the sample; that is, the order volume in the past 24 hours is used as the input feature to predict the order volume in the next hour. At the same time, min-max normalization is used to process the data to eliminate the dimensional influence. The model uses a radial basis kernel function (RBF) to complete the fitting on the training set. For the multi-step prediction task, the rolling iterative prediction strategy is adopted to feed back the prediction results of each step to the feature window, and the prediction sequence from August 22 to 28 is generated circularly. The prediction results are restored to the original value range and limited to non-negative values.

4.2. Analysis of prediction results

The prediction results of the SVR model on the test set were Mae=83.91, RMSE=108.25, MAPE=71.66%, R^2=0.705. The results show that the model can capture the overall trend of the order data, but its ability to describe the cyclical fluctuations in the day is weak, and the prediction deviation is relatively large during the peak period of the order. The MAPE value of SVR is 71.66%, which indicates that the forecasting stability of SVR is not enough in the period of low order volume.

4.3. Comparison of multi-model prediction performance

The comparison results of prediction performance of different models are shown in Table 2. Among them, SARIMA (1,0,1) (1,1,1,24) has the lowest MAE and RMSE, and R2 reaches 0.815, which is the best fitting effect for intraday periodic fluctuations. The performance of the SVR model is second, which can capture the overall trend, but the details are not enough. This result shows that for short-term traffic data with strong intraday cycle characteristics, the SARIMA model considering seasonal effects has more advantages.

DOI: 10.54254/2755-2721/2026.BA36232

Table 2. Comparison results of prediction performance of different models

Model	MAE
SVR	83.91
SARIMA(1,0,1)(1,1,1,24)	64.07

5. Discussion

In this experiment, SARIMA's MAPE was 57.07%, and SVR's MAPE was 71.66%, both of which exceeded 50%. According to the general evaluation standard of traffic prediction, the percentage error performance was poor. This result is mainly caused by the comprehensive effect of the following factors. There are a large number of minimum orders. The number of orders for taxis in New York at night is very small, and the number of orders in some time periods is approaching zero. The number of orders in the morning and evening peak increases rapidly, and the number of orders at night decreases. The number of orders varies greatly day and night. It does not integrate multi-dimensional external features such as weather and holidays, so as to improve the ability of the model to capture sudden passenger flow fluctuations. Existing studies use the SARIMAX model to model and predict urban traffic time series data, effectively taking into account the seasonal laws of time series data and external factors, and accurately identifying the dynamic change law of travel demand [9].

This study still has some limitations and needs to be improved in the future. Deep learning models such as LSTM or hybrid models are not introduced to further enrich the model comparison dimension. Some scholars have fused PCA, SAE, LSTM, and KNN to build a palknn hybrid model, which has lower prediction error than SARIMA and other single models. It is confirmed that multi-algorithm fusion can improve the time series prediction accuracy, and the prediction effect can be optimized by using the idea of multi-model fusion [10]. In the future, minute-level fine-grained prediction research can be added to provide a more comprehensive reference for urban taxi fine scheduling. In addition, this study only uses the SVR model as a univariate baseline model, without giving play to its theoretical advantages of multi-feature fusion, which is one of the follow-up improvement directions.

6. Conclusion

This paper takes the taxi hourly order time series data as the research object, uses the SARIMA seasonal time series model and the SVR machine learning model to carry out short-term order volume prediction research, and completes the prediction performance comparison of the two kinds of algorithm models through quantitative comparison of multiple evaluation indexes.

In this paper, the SARIMA model with seasonal fitting ability is used for prediction modeling. By adding a seasonal structure, the model can effectively mine the periodic evolution law of time series data and can better fit the dynamic characteristics of the day and night peak and trough of order quantity. The experimental results show that the determination coefficient of the optimal SARIMA (1,0,1) (1,1,24) model is 0.815, the error index is the best, and the overall prediction accuracy and stability are excellent.

At the same time, this paper introduces the SVR machine learning model to carry out horizontal comparative research. The SVR model can capture the overall change trend of the order data based on its strong nonlinear fitting ability, and has a certain prediction effect. However, due to the lack of

DOI: 10.54254/2755-2721/2026.BA36232

a special cycle feature extraction mechanism, the model is difficult to accurately depict the details of intraday periodic fluctuations. The prediction deviation is obvious in peak and trough periods, and the overall prediction performance is weaker than the SARIMA model.

The comprehensive research results show that for the traffic time series data with strong intraday periodicity, the SARIMA model with a seasonal term can effectively capture the time series periodicity, and the prediction accuracy and generalization ability are better, which is more suitable for urban taxi short-term order prediction. This study can provide a reference for taxi dynamic capacity scheduling and optimal allocation of traffic resources.

References

[1] Meng, C., Wang, H., Lin, H., et al. (2023). Review of research on road traffic flow data prediction method. *Computer Engineering and Applications, 59*(14), 51–61.

[2] He, J. (2024). Research on traffic flow prediction based on ARIMA model. In *Proceedings of the 1st International Conference on Engineering Management, Information Technology and Intelligence (EMITI 2024)* (pp. 16–21).

[3] Li, P. (2019). *Taxi passenger demand forecasting based on machine learning method* [Master's thesis or doctoral dissertation, Shanghai Jiao Tong University].

[4] Kuang, D. (2019). *Research on short-term forecasting of online car hailing demand based on time series analysis* [Master's thesis or doctoral dissertation, Beijing Jiaotong University].

[5] Li, X. (2021). *Spatial temporal characteristics analysis and demand forecasting research based on taxi data* [Master's thesis or doctoral dissertation, Southwest Jiaotong University].

[6] Ni, J., Yu, L., & Jin, X. (2021). Passenger flow forecast and research of urban rail transit based on ARIMA model. *Intelligent Computer and Applications, 11*(4), 135–138.

[7] Zhang, L., Fan, M., & Ma, J. (2025). Research on traffic congestion prediction based on large language model. *Logistics Technology, 48*(17), 90–93.

[8] Patel, A. (2026). Cleaned NYC taxi trip data (2025 sample) [Data set]. *Kaggle*. https://www.kaggle.com/datasets/aryanpatel212/cleaned-nyc-taxi-trip-data-2025-sample

[9] Paul, B., Biswas, D., Chowdhury, F. T., & Sultana, M. (2025). Geospatial and temporal trends in urban transportation: A study of NYC taxis and Pathao food deliveries. *arXiv*.

[10] Anjaneyulu, M., & Kubendiran, M. (2023). Short term traffic flow prediction using hybrid deep learning. *Computers, Materials & Continua, 75*(1), 1642–1656.