

# Experimental Analysis of Time Series Models for Traffic Flow Prediction

Nivedita Tiwari<sup>1</sup>, Lalji Prasad<sup>2</sup>

<sup>1</sup>Sage University, Indore, India

## ABSTRACT

Traffic flow prediction is a research topic signified by several researchers in a league span of disciplines. In this context, one of the most in-demand techniques of Machine Learning, especially time series-based techniques, helps in predicting traffic flow forecasting and increases the accuracy of the prediction model. In order to deliver extremely precise traffic forecasts, it is crucial that we put the prediction system into practice in the actual world. We aim to perform computations related to traffic on the traffic datasets and determine the accuracy of each model. For this purpose, we are using three distinct time series models: Long Short Term Memory (LSTM), the Autoregressive Integrated Moving Average (ARIMA), and the Seasonal Autoregressive Integrated Moving Average (SARIMA). From the results obtained, it is concluded that the proposed model achieves the highest prediction accuracy with the lowest root mean squared error.

**Keywords:** Time Series Forecasting, LSTM, ARIMA, Seasonal ARIMA, RMSE.

*SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology* (2023); DOI: 10.18090/samriddhi.v15i03.05

## INTRODUCTION

Accurate and timely traffic flow prediction is necessary for intelligent traffic management; hence, rerouting and rescheduling can be implemented and achieve smooth traffic flow. The vital parameters are vehicle movement, relieving congestion, choosing the best route, prediction of traffic flow, volume and density of traffic flow. Traffic flow in real time has a strong temporal connection and periodicity, weather conditions and irregularity make the task more difficult to predict traffic flow.

For traffic prediction, many models has been proposed. The two major categories of the current traffic flow prediction models are traditional models and statistical models.<sup>[1]</sup>

The previous study depicts variety of methods for traffic forecasting that are still essential and integral part for effective traffic control. According to this viewpoint, the following study uses three time series models to forecast future traffic data from real-world daily traffic streams: Long Short Memory (LSTM) as traditional model, Autoregressive Integrated Moving Average (ARIMA) as a statistical model, and Seasonal Autoregressive Integrated Moving Average (SARIMA) as a classic RNN model.

Traditional approaches like long short-term memory (LSTM) and recurrent neural network (RNN) are examples of traditional techniques that exhibit greater function fitting abilities in complex and nonlinear traffic flow prediction issues. A unique variety of recurrent neural network that can recognize long-term dependencies is the LSTM network. In 1997 by Hochreiter & Schmidhuber<sup>[2]</sup> underwent much

---

**Corresponding Author:** Nivedita Tiwari, Sage University, Indore, India, e-mail: nivedita.8584@gmail.com

**How to cite this article:** Tiwari, N., Prasad, L. (2023). Experimental Analysis of Time Series Models for Traffic Flow Prediction. *SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology*, 15(3), 312-320.

**Source of support:** Nil

**Conflict of interest:** None

---

investigation and refinement, and concluded that LSTMs were created to prevent the long-term reliance issue. Practically speaking, their greatest advantage is their ability to retain information for a long time.

Recurring module in conventional RNNs will have a very straightforward structure, such as a single hyperbolic tangent layer. While the repeating module in LSTM also has this chain structure, it is distinct.<sup>[3]</sup> The authors of<sup>[4]</sup> suggest using neural networks to exhibit forecasting skills with good accuracy. The most famous techniques are two Deep learning techniques, convolutional neural networks (CNNs) and recurrent neural networks (RNNs). CNNs are typically used for learning spatial structure,<sup>[5]</sup> while RNNs (like LSTM<sup>[6]</sup> and GRU<sup>[7]</sup>) are frequently used for learning temporal and sequential information. In<sup>[8]</sup> <sup>[9]</sup> <sup>[10]</sup> deep learning- a type of brain structure with self-association and the ability to execute nonlinear auto-relapse used to perform traffic forecasts. However, the majority of the aforementioned approaches model each traffic time series independently, failing to recognize their spatial dependence.

For the flow prediction technique, Vijayalakshmi *et al.*<sup>[11]</sup> developed the attention-based CNN-LSTM model, which analyzes the data with deep neural architecture and incorporates the spatial correlation of data from the closest station. The outcomes show that the technique performs well for peak and nonpeak timing prediction.

Statistical methods include the autoregressive integrated moving average in comparison to conventional methods. The former method relies on the historical average value at a particular point in the past to determine the expected value, whereas the latter involves constructing a mathematical model based on the time series data. This approach has been extensively utilized for a considerable period due to its ability to illustrate the cyclical variations in traffic flow statistics. A well-liked time series analysis method that has been applied to predict traffic is the auto-regressive integrated moving average to forecast web traffic<sup>[12]</sup> and road traffic.<sup>[13, 14]</sup>

The processing of complicated data is still a limitation for the single model to process. Therefore, combination models are more preferred over single models to enable more precise traffic flow forecasts. In our earlier research,<sup>[15]</sup> seasonal ARIMA was used to capture the periodicity of traffic flow on different days and during rush hours. Chai employed a sparse seasonal ARIMA in<sup>[16]</sup> to properly predict aviation traffic from January to July of 2020 with a 95% confidence interval and to recognize the COVID-19 outbreak, financial crisis, and political unrest. Combining ARIMA with additional techniques could increase the forecasting precision of hybrid models.<sup>[17]</sup>

DBNs were used by Koesdwiady *et al.*<sup>[18]</sup> combining traffic flow and weather data at the decision level can lead to more accurate predictions based on traffic and weather data. However, achieving good performance in decision-level data fusion can be difficult since some studies have only focused on traffic prediction based on weather data, without considering other weather factors. Jian *et al.*<sup>[19]</sup> studied the characteristics of microscopic traffic flow in wet weather conditions, and although they partially revealed a link between traffic volume and weather, their findings did not address the specific prediction problem at hand. To understand the connection between weather and traffic movement, Zheng *et al.*<sup>[20]</sup> presented a revolutionary SAE and RBF combination architecture based on weather and traffic flow information that extracted the weather disturbance using embedding components but did not analyze or manipulate the weather parameters. For greater accuracy, it is essential to carry out research on traffic flow prediction using both traffic and weather data.

The efficacy of the proposed model was assessed through an experiment analysis involving two groups of datasets and several time series models. The performance of these models was then evaluated using specific criteria across different simulations for comparison. The current paper is

organized as follows: Section 2 reviews the literature on traffic flow forecasting and presents several traffic forecasting techniques. Section 3 which covers data exploration, scope,

and the study's context. The subsequent part is presented in Section 4 dedicated to modeling by depicting an investigation of the model's performance on a real dataset according to predefined criteria. Section 5 provides the conclusion and future recommendations.

## LITERATURE REVIEW

By studying the available literature on previous methods, their challenges and areas of improvement, we can better understand the time series analysis to tackle traffic forecasting.

Zheng and Huang,<sup>[21]</sup> developed a traffic flow forecast model based on the LSTM network. They proposed the LSTM network by comparing two classic forecast models ARIMA model and the BPNN model. The author's experimental results show that the proposed model can accurately predict the traffic flow based on their natively stable time series under normal conditions. MOU, ZHAO and CHEN<sup>[22]</sup> mainly emphasize temporal information for short-term traffic flow prediction. The temporal information includes day-wise information for traffic flow prediction, improving the prediction performance of the LSTM model. Leong, Lee *et al.*,<sup>[23]</sup> explored the impact of rainfall intensity on low-resolution speed band data. They tested rainfall may improve the prediction accuracy of data-driven models for individual roads by using support vector machine (SVM) prediction algorithms. The Authors have done the numerical analysis for 616 road segments in Singapore on a prediction horizon of 5 min. Numerical results show that for a certain number of links rainfall information enhances the prediction accuracy.

Noor Afza Mat Razali *et al.*,<sup>[24]</sup> offer a thorough investigation into the use of ML and DL methods to enhance traffic flow prediction, improving ITS in smart cities. For this purpose, the author provided a thorough and methodical review of the literature that included some articles released from 2016 onward and pulled from four major databases. According to the methodology and algorithms used to predict traffic flow, their findings include the gaps, approaches, evaluation techniques, variables, datasets, and outcomes of each research that was examined.

CNN and LSTM were compared to existing baseline models to determine their effectiveness. Shah, Almazah and Rezami,<sup>[25]</sup> analyze how well functional time series modeling predicts traffic flow one day in advance. Additionally, researchers compared the developed model FAR (1) with the conventional ARIMA Model. As a result the proposed model forecasting errors are considerably lower than the conventional model.

Rooke *et al.*,<sup>[26]</sup> proposed a framework for evaluating feature importance in time series prediction over extended time horizons and with complex feature interactions that is applied to sequential models for multivariate time series prediction. In the actual world, the proposed framework performs 2.47 times better than other feature-importance methods. Congestion-based Traffic Prediction Model (CTPM),

introduced by Nagy and Simon,<sup>[27]</sup> is a novel traffic prediction model that can be used in conjunction and forecasts traffic that could be improved by the developed model by an average of 9.76%. Wang, Zhanga *et al.*,<sup>[28]</sup> compare the proposed model in which deep polynomial neural network GMDH combined with SARIMA to the LSTM based model in order to forecast the short-term traffic flow.

Fang Ma,<sup>[29]</sup> studies the storage capacity of existing road space and seeks the optimal control method for local regional road network congestion. He evaluates the function of regional road networks and proposes a traffic performance quality evaluation system based on the traffic volume of road networks in a city to reduce traffic congestion, improve coordination of urban traffic, and reduce the negative impact of traffic congestion.

The hybrid prediction technique proposed by Luo, Niu, and Zhang,<sup>[30]</sup> combines an improved seasonal autoregressive integrated moving average (ISARIMA) model with a multi-input autoregressive (AR) model by genetic algorithm (GA) optimization. Test results using actual traffic data supplied by TDRL reveal that the suggested method outperforms other methods in terms of prediction performance. Javier Del and Vlahogianni's,<sup>[31]</sup> primary objectives were together and summarize prior work related to data-driven traffic forecasting techniques, to suggest a framework for comparison, and review the most recent literature in light of the updated criteria.

## IMPLEMENTATION

The interaction of a huge number of vehicles determines the complicated, nonlinear, and unpredictable traffic behavior. In order to comprehend and create the best possible transportation system with an effective traffic flow and a minimum amount of congestion issues, traffic flow is the study of interaction between users and infrastructure with respect to time.

The concept of breaking down a time series into deterministic and non-deterministic components is used in the theory of time series analysis (or predictable and unpredictable components). There are two basic steps to go before the experiment analysis.

- **Traffic Flow Detection:** The techniques for detecting traffic movement are based on intelligent video. The first step is to identify the vehicle item in the video or image sequence. The recognized vehicle is then followed to establish a link between subsequent video frames. Finally, it produces the detection result after counting the number of vehicles that pass the relevant lane within a predetermined amount of time.<sup>[32]</sup>
- **Traffic Flow Prediction:** The intelligent transportation and management system (ITMS) needs a traffic prediction model and thus manages according to possible future events. This is a precise forecast of traffic in a particular location at a particular time in the future.<sup>[33]</sup>

## Data Experimentation

Traffic data comprises of both historical data and current traffic information. Some of the parameters are a number of vehicles passing at a certain point, their speed, time, vehicle type, diversion, seasonal traffic load and climatic conditions.

## Data Source: Weather, Climatic and Traffic Dataset Analysis

This section depicts that traffic flow and weather datasets are in direct proportion change in one causes a change in other as well. Many factors influence traffic, and all the factors need to be considered carefully for accurate predictions. As a result, there are several major groups of data that you must collect. You must also keep in mind that ML algorithms function optimally when there is enough data to train the models and fine-tune them to obtain the highest level of accuracy. As a result, the larger datasets you are able to collect, the better the outcomes.

## Traffic Data

The dataset used in this research considers two vital parameters traffic volume and speed. The data consist of two columns, one for the dates and the other for the total number of vehicles flown between 2015 and 2017.

## WeatherData

Environmental factors influence traffic and speed limitations throughout the roadway. Therefore their historical and current and forecasted traffic information has to be analyzed. The dataset used for the analysis consists of five columns, one for the dates and the other for the weather factors that affect the traffic flow.

## Data Visualization

Data visualization once we are sure that the dataset is complete and does not contain erroneous values, we proceed to the visualization. Figure 1 represents the traffic data flow with respect to total count of vehicles. Figure 2 shows a timing diagram of a number of vehicles with respect to time in hours.

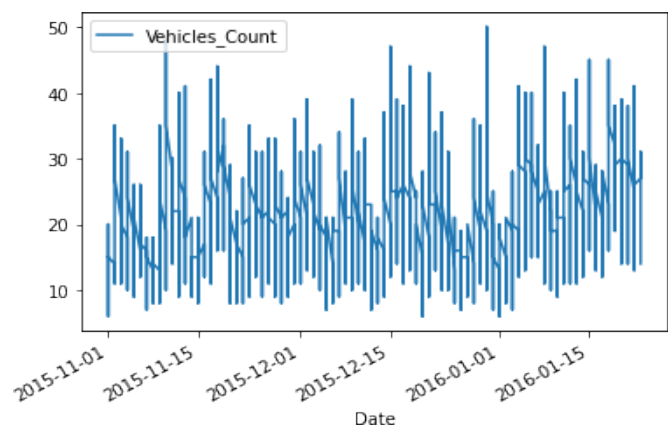


Figure 1: Traffic Data showing Vehicle Count over Date



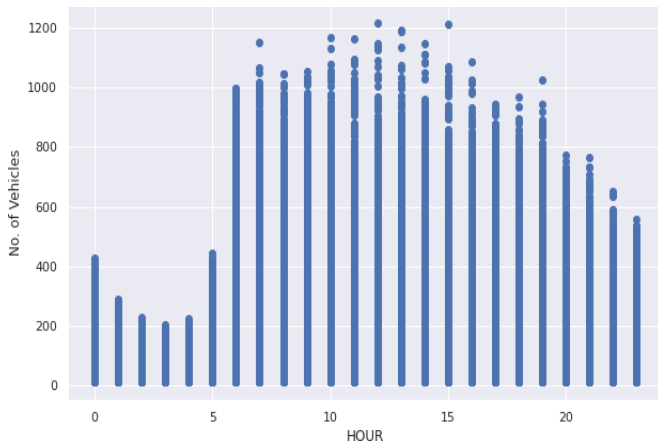


Figure 2: Time diagram showing trajectories of vehicles over Time (Hour)

## MODELING

Machine learning (ML) can be used to construct prediction models incorporating vast quantities of diverse data from various sources. There has been extensive research on using ML models for traffic forecasting, and selecting an appropriate model for time-series data will depend on specific data characteristics, such as trend and seasonality [10]. Therefore, it's crucial to choose the model that is most suitable for your data. Below are some examples of effective models:

### RNN: LSTM

The purpose of recurrent neural networks (RNNs) is to handle time-series data or observations gathered over a predetermined period of time. A good example of such observations is traffic patterns. RNN is a form of Artificial Neural Network. It works very well, especially for sequences of data that are important for doing time series forecasting.

Long Short-Term Memory (LSTM) networks are a unique type of recurrent neural network that enables the retention of information over time. In Python, LSTM can be implemented using various libraries. One of the significant benefits of LSTM is its ability to circumvent long-term dependency issues, which is a common problem in other types of neural networks.

### LSTM Model Structure

The structure of an LSTM model comprises three components, commonly referred to as gates. The first gate, known as the Forget gate, determines whether information from the previous timestamp should be disregarded. The second gate, called the Input gate, attempts to learn new information from the input. Lastly, the Output gate in the third component transmits the updated information from the current timestamp to the next timestamp. LSTM is suitable for processing time sequences with uncertain durations since it enables such sequences' classification, analysis, and prediction.

```

Model:"sequential"
Layer(type)           OutputShape           Param#
=====
lstm(LSTM)            (None,100)           40800
dense(Dense)         (None,1)             101
=====
Totalparams:40,901
Trainableparams:40,901
Non-trainableparams:0
    
```

### LSTM Model Training

The LSTM model will have to perform the following steps: Firstly, the activation function called Relu is used to train the model. Then, use optimizer Adam to define the loss as the mean square error. Finally, the model will be trained with 16 sequences per batch for 50 epochs as shown in Figure 3.

### LSTM Results

We test our model on the test set after the creation and training phases to see how it performs. Figure 4 illustrates the results that were achieved. The LSTM recurrent neural networks perform well, but they have a significant drawback affecting the findings' quality. In particular, they struggle to handle time series that have several variations in the passage of time.

### Statistical Approach: ARIMA

The auto-regressive integrated moving average (ARIMA) model is widely regarded as one of the most effective models

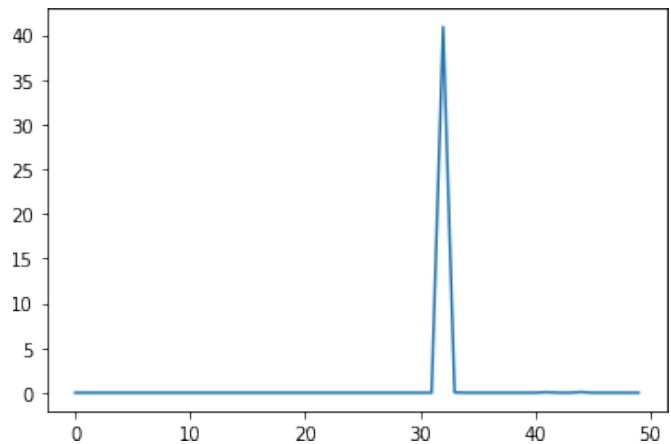


Figure3: Loss Per Epoch

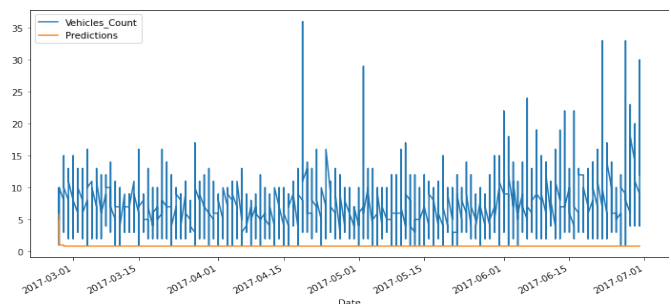


Figure 4: Time series prediction with LSTM

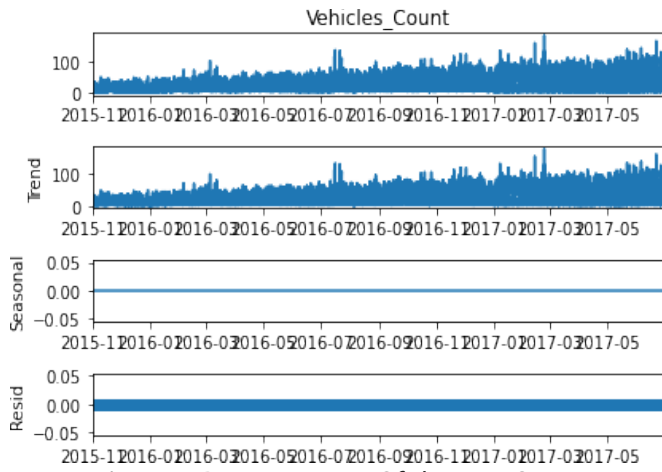


Figure 5: Decomposition Of The Time Series

for time-series analysis, especially for traffic forecasting. Since the 1970s, ARIMA models have been extensively used due to their simplicity and high accuracy compared to other statistical methods. It uses a traditional statistical methodology to analyze the past and forecast the future. It gathers data from a series of regular time intervals and makes the assumption that historical patterns will continue in the future. However, the univariate ARIMA models are unable to handle the complicated structure and large number of variables that make up traffic flow.

### Seasonality and Trend

To better understand the nature of the series and examine the time series' trend and seasonality in order to break it out and apply forecasting models. The goal of this essay is to introduce time-series forecasting.

The 'stats models' package includes a valuable Python function called `seasonal_decompose` that can be used to perform decomposition based on the mentioned patterns. After examining the four parts of the decomposed graphs, as depicted in Figure 5, two fundamental patterns of any time series are discussed: trend and seasonality. Then, a time-series decomposition method based on these patterns is explained. Based on the decomposition results, it can be concluded that the time-series data exhibit a pronounced annual seasonality component and an increasing trend over time.

### Stationarity Test

To apply the ARIMA model it is necessary to perform a test for stationarity to transform non-stationary time series into stationary ones. One common test for stationarity is the Augmented Dickey-Fuller (ADF) test.<sup>[34]</sup> According to the results of the ADF test, if the obtained p-value is less than 0.05, the time series is stationary.<sup>[35]</sup>

### Check for Stationarity

Results of ADF Test

ADF:-25.41162919829304

P-Value:0.0

NumOfLags:23

NumOfObservationsUsedFor ADFRegression and Critical Values Calculation: 32051

CriticalValues:

1%:-3.43055404429831

5%:-2.8616301822764014

10% :-2.566818001236828

The statistical results shows an effect of the stationarity of the time series. While the ADF test indicates that time-series is stationary (p-value > 0.05).

### Autocorrelation

Autocorrelation is interpreted using the ACF and PACF graphs.<sup>[36]</sup> These graphs represent correlation between the original data and the second order shift. The Correlation Coefficient is always computed between two variables. If only one variable is available for computation, then we must use autocorrelation. Figure 6 presents the ACF and PACF graphs.

### Train and Test

In order to separate training and testing, the original dataset is split into two components: train dataset and test dataset. The train data is used to train the model and when new data is generated, the test data is issued to report the square error.

### Evaluation Criterion with Result

In order to evaluate the prediction performance, the implementation uses the root-mean-square error (RMSE) as a core metric to assess the model's error. It was chosen to

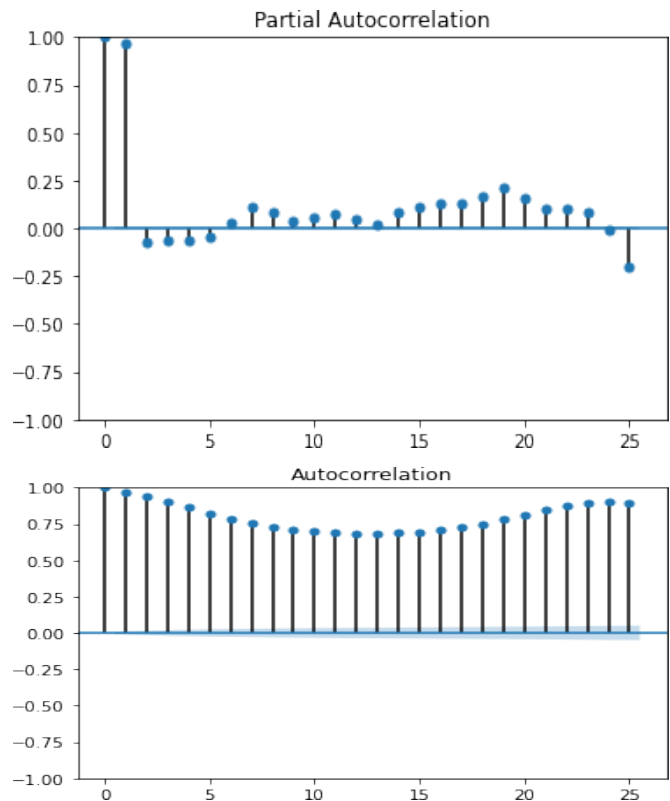


Figure 6: Interpreting ACF and PACF Plots for Time Series Forecasting using ARIMA



evaluate the difference between actual and predicted values.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - x_i)^2}{n}}$$

Where n is the length of prediction data,  $y_i$  and  $x_i$  are the measured value and predicted for the validation sample, respectively. Figure 7 represents the time series prediction with the ARIMA model. The performance is shown in Table 1 with error 7.49595603806666 and accuracy 92.97%.

### SARIMA

The SARIMA model is a powerful tool for time series analysis and shares similarities with ARIMA. SARIMA stands for seasonal autoregressive integrated moving average and can also incorporate exogenous factors in its forecasting. The seasonal components of the model are precisely described by an additional set of parameters (P, D, Q) that SARIMA accepts. Here, m stands for the length of a season in the series, and P, D, and Q stand for the seasonal regression, differentiation, and moving average coefficients:

P represents the seasonal autoregressive order,

D represents the seasonal difference order,

Q represents the seasonal moving average order, and

m represents the number of time steps within a single seasonal period.

The indication is shown as (p,d,q)×(P,D,Q)m. The goal of this modeling is to forecast the daily average number of vehicles flown incorporating the external factors.

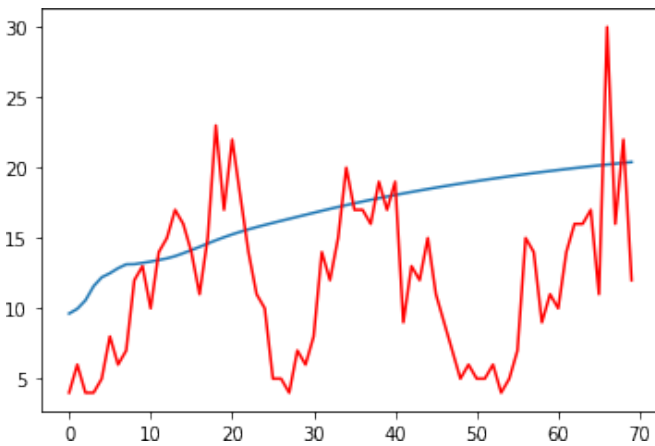


Figure 7: Time series prediction with ARIMA

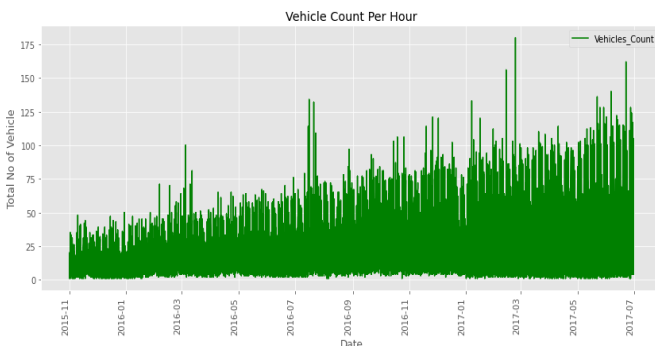


Figure 8: Dataset Visualization

### Visualizing dataset

Before we examine the time-series trends, let's show the data so that each vertical dashed line signifies the start of the year. Figure 8 describes us as a representation where each point in the entire dataset is present with respect to vehicle count.

### Testing the Stationarity of the Dataset

The Dickey Fuller test can be used to determine whether the data are stationary before implementing the SARIMA model. [34] The time series can be considered stationary only if the obtained p-value is less than 0.05. [35]

AugmentedDickey-Fullertest (ADFTest)

A D F S t a t i s t i c : - 7 . 5 1 8 2 7 8 1 1 1 9 3 5 7 2 2 p - value:3.851060199915586e-11

### Differencing

After verifying the stationarity of the time series, differencing is a common technique used to transform it. Differencing is used to remove temporal dependencies that may include trends and seasonality.

Once all temporal dependencies have been removed, differencing can be applied multiple times. Therefore, we can repeat the Dickey-Fuller test. The number of times the differencing is performed is referred to as the difference order (D). After one order of differentiation (D=1) the series becomes stationary.

AugmentedDickey-Fullertest (ADFTest)

ADFStatistic:-35.78125992349637

p-value:0.0

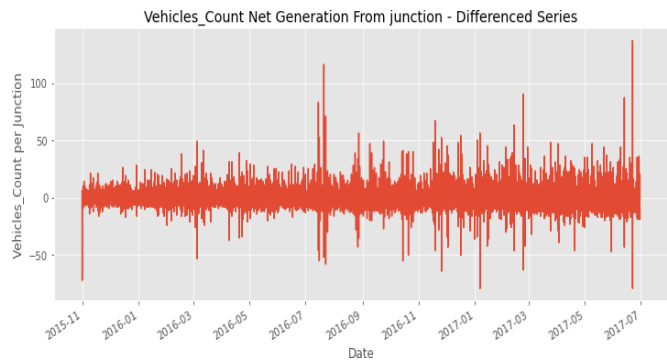


Figure 9: Visualize Differenced Series

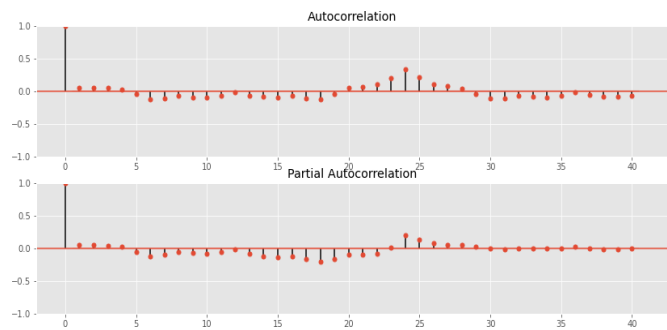


Figure 10: Interpreting ACF and PACF Plots for Time Series Forecasting using SARIMA

### Autocorrelation

Autocorrelation in a time series involves comparing a variable's current data to its past data. To determine the order of a SARIMA model, we use the ACF and PACF plots. Figure 10 shows the ACF and PACF graphs.<sup>[36]</sup>

### AIC

The Akaike Information Criterion (AIC) is a commonly used measure of the quality of a statistical model.<sup>[37]</sup> It assesses the model's goodness of fit and complexity and produces a single statistic that summarizes both aspects. When comparing two models, the one with the lower AIC value is generally considered to be the better choice.<sup>[38]</sup> As a result of the AIC function, the order of SARIMA is (2,1,2)x(2,1,2,12).

### SARIMA Results

```
=====
Dep. Variable: Vehicles_Count No.Observations: 48119
Model: SARIMA(2,1,2)x(2,1,2,12)LogLikelihood-136501.482
Date: Mon,02 Jan2023 AIC273020.964
```

```
Time:09:02:21 BIC273099.994
Sample: 0 HQIC273045.764
-48119
```

### Diagnosing the model residuals

Each observation in a time series can be forecast using all previous observations. The residuals of the model in a time series analysis are useful for determining if a model has fully captured the information in the data and what remains after model fitting. The residuals are typically equal to the difference between the observed values and the corresponding fitted values for time series models.

#### #Diagnosingthemodel residuals

We generate the diagnostic model residuals as shown in Figure 11. Figure 11 (upper left) provides standardized vehicle count (V) residuals. From the QQ plot, we see the normality assumption between the sample quantities and theoretical quantities. On the other hand, Figure 11 (upper right) presents a histogram with estimated density. The correlogram (lower right) of residuals indicates that they are stationary in nature and have no pattern.

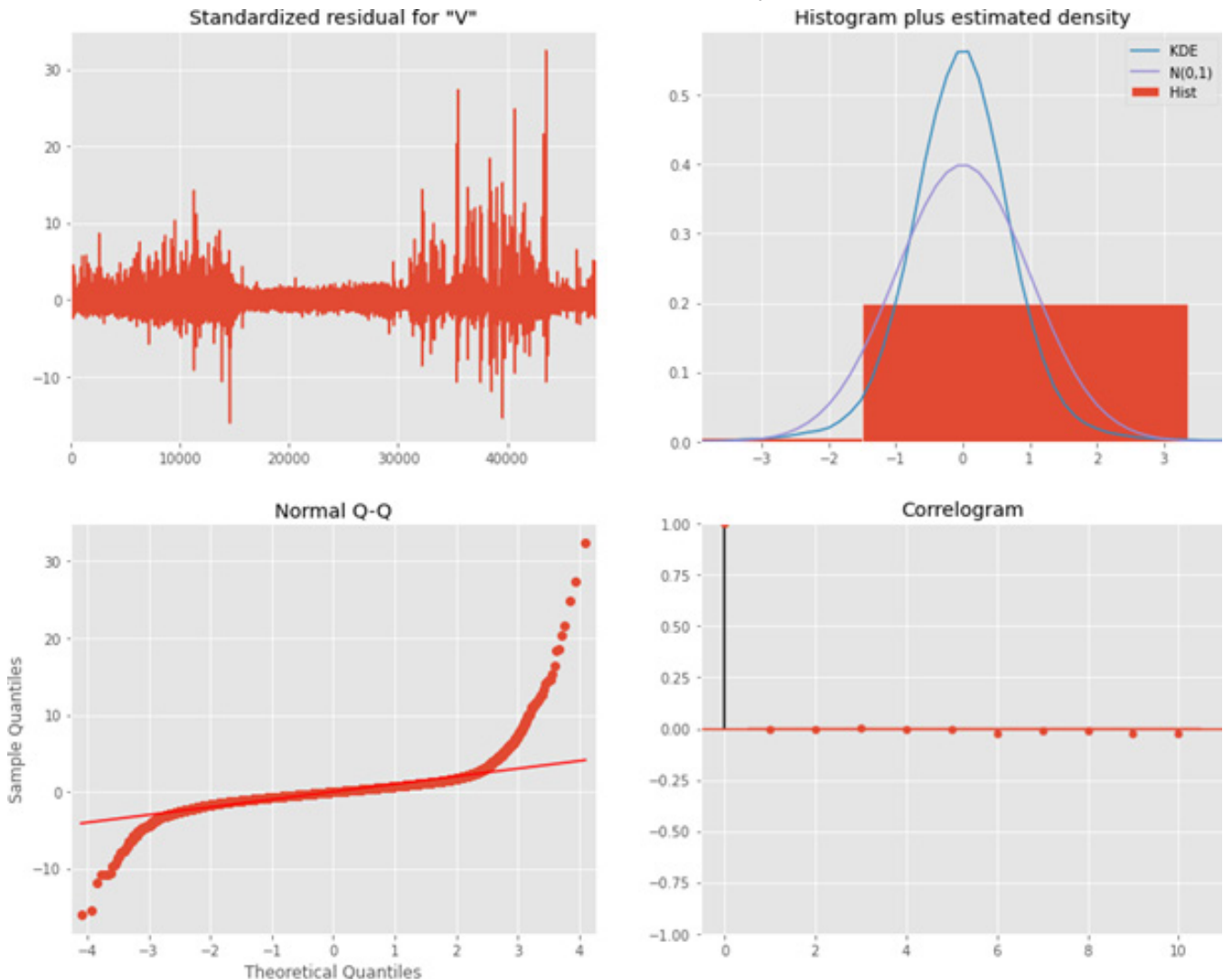


Figure 11: Absolute residuals versus indices of corresponding observations in the SARIMA model



**Table 1:** Experimental Result of Implementation on the Traffic Datasets

Data Set	Performance Criteria	LSTM	ARIMA	SARIMA
Traffic	RMSE	7.03	7.49	4.17
	Accuracy%	92.97	92.51	95.83

### Model Forecast and Validation through Evaluation

For a result evaluation of the SARIMA model, performance measures RMSE are taken into account. The RMSE is a widely used performance measure for models, including those used in traffic forecasting. It provides an indication of how much error there is between the actual and predicted values, with lower RMSE values indicating better model performance. RMSE is calculated as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (|y_i - x_i|)^2}{n}}$$

The performance obtained from the model by using Root Mean Square Error (RMSE): 4.175097. Therefore, the SARIMA model can be used to produce traffic forecasts accurately (see Table 1).

## RESULTS: COMPARISON AND DISCUSSION

The experimental results confirm that the proposed model gives the best forecasting results. Indeed, an example comparing the root mean square error of the proposed models LSTM, ARIMA and SARIMA is given in the table where the minimum root mean square error (4.17) with 95.8% of accuracy is obtained by using the SARIMA model for the traffic data.

## CONCLUSION

In this research work, we have adopted the Time series models for real-time traffic flow prediction. We have used three different time series models (LSTM, ARIMA, SARIMA) for prediction and investigated these three model performances. Using RMSE evaluation, Metrics models are compared. The results of the experiment illustrated that LSTM is most suitable for the sequential data weather in terms of prediction accuracy and reliability. The ARIMA Model performs best. ACCORDING TO THIS RESEARCH, the SARIMA model performs better when Akaike's Information Criterion (AIC) method is used. The utilization of different time series models with weather and periodic features introduced as additional information will make the model more powerful in temporal feature extraction. Our future work could be analyzing the impact of the additional information in normal regular days such as weather, holidays, and events on the traffic flow.

## CONFLICTS OF INTEREST

There is no conflict of interest between the authors of this paper.

## DATA AVAILABILITY STATEMENTS

The datasets analyzed during the current study are available in the Kaggle. <https://www.kaggle.com/datasets/fedesoriano/traffic-prediction-dataset>.

## REFERENCES

- [1] A.M.Nagy,V.Simon,“Survey on traffic prediction in smart cities”, Pervasive and Mobile Computing(2018), <https://doi.org/10.1016/j.pmcj.2018.07.004>.
- [2] S.Hochreiteret, J.Schmidhuber,“LSTM can solve hard long time lag problems”.1997.
- [3] B.-J.HouetZ.-H.Zhou, “Learning With Interpretable Structure From Gated RNN”, IEEE Trans. Neural Netw. Learn. Syst.,p1-13,2020, doi:10.1109/TNNLS.2020.2967051.
- [4] N. Slimani, I. Slimani, N. Sbiti, et M. Amghar, “Traffic forecasting in Morocco using artificial neural networks”,ProcediaComput. Sci.,vol.151,p.471-476,2019,doi:10.1016/j.procs.2019.04.064.
- [5] B. Yu, Y. Lee, et K. Sohn, “Forecasting road traffic speeds by considering area-wide spatio-temporal dependencies based on a graph convolutional neural network (GCN)”,Transp. Res. PartC Emerg.Technol.,vol.114,p.189-204,mai2020,doi:10.1016/j.trc.2020.02.013.
- [6] B.Yang, S. Sun, J. Li,X.Lin,etY. Tian, “Traffic flow prediction using LSTM with feature enhancement”, Neurocomputing,vol.332,p.320-327,mars 2019,doi:10.1016/j.neucom.2018.12.016.
- [7] M.Jiang, W.Chen, et X.Li,“S-GCN-GRU-NN:A novel hybrid model by combining a Spatiotemporal Graph Convolutional Network and a Gated Recurrent Units Neural Network for short-term traffic speed forecasting”, J.DataInf.Manag.,janv.2021,doi:10.1007/s42488-020-00037-9.
- [8] M.X.Dai, M.R.Fu, M.Y.Lin, F.-Y.Wang, et L.Li, “Deep Trend: A Deep Hierarchical NeuralNetworkforTraffic FlowPrediction”,p.6,2017.
- [9] P.Ø.Kanestrøm, “Traffic flow forecasting with deep learning”,p. 96.
- [10] Hongsuk Yi, HeeJin Jung, et SanghoonBae, “Deep Neural Networks for traffic flow prediction”, in 2017 IEEE International Conference on Big Data and Smart Computing (BigComp), Jeju Island, SouthKorea, févr.2017,p.328-331.doi:10.1109/BIGCOMP.2017.7881687.
- [11] Balachandran Vijayalakshmi, Kadarkarayandi Raman, NZ. Jhanjhi, Sahil Verma, Madasamy Kaliappan, Kandasamy Vijayalakshmi, ShanmuganathanVimal, Kavita and Uttam Ghosh. “An attention-based deep learning model for traffic flow prediction using spatiotemporal features towards sustainablesmart city”. Int J Commun Syst. 2020;e4609. wileyonlinelibrary.com/journal/dac © 2020 John Wiley & SonsLtd.1of14<https://doi.org/10.1002/dac.4609>.
- [12] C. A. H. Suarez et O. J. S. Parra, “An ARIMA model for forecasting Wi-Fi data network trafficvalues”,p.5.
- [13] X.Luo, L. Niu,etS.Zhang,“AnAlgorithm forTraffic Flow Prediction Based on ImprovedSARIMA and GA”, KSCE J. Civ. Eng., vol. 22, no 10, p. 4107-4115, oct. 2018, doi: 10.1007/s12205-018-0429-4.
- [14] N. Slimani, I. Slimani, N. Sbiti, et M. Amghar, “Machine Learning and statistic predictive modelingforroadtrafficflow”,Int.J.TrafficTransp.Manag.,vol.03,no01,p.17-24,mars2021,doi:10.5383/JTTM.03.01.003.
- [15] N. Slimani, I. Slimani, M. Amghar, et N. Sbiti, “Comparative analysis of traffic prediction modelsbasedona realcasestudy”. INTIS2019.



- [16] S. Chai, "Hong Kong Air Traffic: Explanation and Prediction based on Sparse Seasonal ARIMA Model", ArXiv210805817Stat, août2021, Consulté le: août 29, 2021. [En ligne]. Disponible sur: <http://arxiv.org/abs/2108.05817>
- [17] M. Ge, Z. Junfeng, W. Jinfei, H. Huiting, S. Xinghua, et W. Hongye, "ARIMA-FSVR Hybrid Method for High-Speed Railway Passenger Traffic Forecasting", Math. Probl. Eng., vol. 2021, p. e9961324, mai 2021, doi:10.1155/2021/9961324.
- [18] A. Koesdwiady, R. Soua, and F. Karray, "Improving traffic flow prediction with weather information in connected cars: a deep learning approach," IEEE Transactions on Vehicular Technology, vol. 65, no. 12, pp. 9508–9517, 2016.
- [19] W. Jian, S. Xixiong, and H. Xiaowei, "Researching on microscopic traffic flow characteristics of freeway under rainy environment," in Proceedings of the 2014 Fifth International Conference on Intelligent Systems Design and Engineering Applications, pp. 889–892, Hunan, 2014.
- [20] Z. Zheng, Y. Yang, J. Liu, H.-N. Dai, and Y. Zhang, "Deep and embedded learning approach for traffic flow prediction in urban informatics," IEEE Transactions on Intelligent Transportation Systems, vol. 20, no. 10, pp. 3927–3939, 2019.
- [21] Jianhu Zheng, Mingfang Huang, "Traffic Flow Forecast through Time Series Analysis Based on Deep Learning", DOI 10.1109/ACCESS.2020.2990738, IEEE Access VOLUME XX, 2017.
- [22] Luntian MOU, Pengfei ZHAO and Yanyan CHEN. "Short-Term Traffic Flow Prediction: A Long Short-Term Memory Model Enhanced by Temporal Information". Conference Paper: July 2019 DOI:10.1061/9780784482292.210. <https://www.researchgate.net/publication/332708140>.
- [23] Leong Wai Leong, Kelvin Lee, Kumar Swapnil, Xiao Li, Ho Yao Tong Victor, Nikola Mitrovic, Muhammad Tayyab Asif, Justin Dauwels, and Patrick Jaillet. "Improving traffic prediction by including rainfall data". Conference Paper April 2015 DOI:10.13140/RG.2.1.1.3752.8084 <https://www.researchgate.net/publication/277330077>.
- [24] Noor Afza Mat Razali, Nuraini Shamsaimon, Khairul Khalil Ishak, Suzaimah Ramli, Mohd Fahmi, Mohamad Amran and Sazali Sukardi. "Gap, techniques and evaluation: traffic flow prediction using machine learning and deep learning", Razali et al. Journal of Big Data (2021) 8:152 <https://doi.org/10.1186/s40537-021-00542-7>.
- [25] Shah, I.; Muhammad, I.; Ali, S.; Ahmed, S.; Almazah, M.M.A.; Al-Rezami, A.Y. "Forecasting Day-Ahead Traffic Flow Using Functional Time Series Approach". Mathematics 2022, 10, 4279. <https://doi.org/10.3390/math10224279>. ©2022 by the authors. Licensee MDPI, Basel, Switzerland.
- [26] Clayton Rooke, Jonathan Smith, Kin Kwan Leung, Maksims Volkov, Saba Zuberi. "Temporal Dependencies in Feature Importance for Time Series Predictions". International Conference on Machine Learning 2021 Time Series Workshop, Virtual Workshop, 29 Jul 2021. arXiv:2107.14317v1
- [27] Attila M. Nagy, Vilmos Simon. "Improving traffic prediction using congestion propagation patterns in smart cities" Advanced Engineering Informatics 50(2021)101343 [journal homepage: www.elsevier.com/locate/aei](http://www.elsevier.com/locate/aei) <https://doi.org/10.1016/j.aei.2021.101343> Received 17 December 2020; Received in revised form 13 April 2021; Accepted 14 June 2021
- [28] WeiWanga, Hanyu Zhanga, TongLib, JianhuaGuoc, WeiHuanga, YunWeid, JindeCaoe. "Interpretable model for short term traffic flow prediction". <https://doi.org/10.1016/j.matcom.2019.12.013>. Published by Elsevier B.V. All rights reserved. W. Wang, H. Zhang, T. Li et al. / Mathematics and Computers in Simulation 171 (2020) 264–2789.
- [29] FangMa; "Traffic Optimization Model Based on Regional Road Network Traffic Diversion Technology and Internet of Things". Hindawi Mobile Information Systems Volume 2022, Article ID 3582899, 11 pages <https://doi.org/10.1155/2022/3582899>.
- [30] Xianglong Luo, Liyao Niu, and Shengrui Zhang. "An Algorithm for Traffic Flow Prediction Based on Improved SARIMA and GA". Published Online May 31, 2018 KSCE Journal of Civil Engineering (0000) 00(0):1-9 Copyright ©2018 Korean Society of Civil Engineers DOI 10.1007/s12205-018-0429-4 pISSN 1226-7988, eISSN 1976-3808 [www.springer.com/12205](http://www.springer.com/12205).
- [31] Ibai Laña, Javier Del, Ser Manuel, Vélez Eleni, Vlahogianni. "An Algorithm for Traffic Flow Prediction Based on Improved SARIMA and GA". 10.1109/MITS.2018.2806634 Date of publication: 20 April 2018 1939-1390/18 ©2018 IEEE.
- [32] Chen Chen, Liu, Shaohua Wan, Peng Qiao, and Qingqi Pei, "An Edge Traffic Flow Detection Scheme Based on Deep Learning in an Intelligent Transportation System". IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS. 1524-9050 ©2020 IEEE.
- [33] Anirudh Ameya Kashyap, Shravan Raviraj, Ananya Devarakonda, Shamanth RNayak K, Santhosh K V & Soumya J Bhat | (2022) "Traffic flow prediction models – A review of deep learning techniques", Cogent Engineering, 9:1, 2010510, DOI:10.1080/23311916.2021.2010510.
- [34] S. Kumar et N. Kumar, "Testing normality in the time series of EMP indices: an application and power-comparison of alternative tests", Commun. Stat.- Theory Methods, p.1-14, mai 2021, doi:10.1080/03610926.2021.1914097.
- [35] D. Lakens, "The Practical Alternative to the p Value Is the Correctly Used p Value", Perspect. Psychol. Sci., vol. 16(3) 639–648, p. 10, févr. 2021, doi:10.1177/1745691620958012.
- [36] S. Roy, G. S. Bhunia, et P. K. Shit, "Spatial prediction of COVID-19 epidemic using ARIMA techniques in India", Model. Earth Syst. Environ., vol. 7, no. 2, p. 1385–1391, juin 2021, doi:10.1007/s40808-020-00890-y.
- [37] A. Khalid et A. I. Sarwat, "Unified Univariate-Neural Network Models for Lithium-Ion Battery State-of-Charge Forecasting Using Minimized Akaike Information Criterion Algorithm", IEEE Access, vol. 9, p. 39154–39170, 2021, doi:10.1109/ACCESS.2021.3061478.
- [38] M. S. Yadnya, A. Mauludiyanto, et G. Hendranto, "Akaike Information Criteria Application to Stationary and Nonstationary Rainfalls for Wireless Communication Channel in Surabaya", p. 8, 2008.

