

## Article

# Comparative Analysis of Parametric and Neural Network Models for Rural Highway Traffic Volume Prediction

Mohammed Al-Turki 

Department of Civil Engineering, Jubail Industrial College (JIC), Royal Commission of Jubail and Yanbu (RCJY), P.O. Box 10099, Jubail Industrial City 31961, Saudi Arabia; moe\_kt@hotmail.com

## Abstract

The information and communication technology revolution has provided researchers with new opportunities to enhance traffic prediction methods. Accurate long-term traffic forecasts are essential for sustainable infrastructure planning, supporting proactive maintenance and efficient resource allocation. They also enable environmental impact assessments and help reduce carbon footprints through optimized traffic flow, minimized idling, and better planning for low-emission infrastructure. Most traffic prediction studies focus on short-term urban traffic, but there remains a gap in methods for long-term planning of rural highways, which pose significant challenges for intelligent transportation systems. This paper assesses and compares six prediction models for long-term daily traffic volume prediction, including two traditional time series methods (ARIMA and SARIMA) and four artificial neural networks (ANNs): three feedforward networks trained with Bayesian Regularization (BR), Scaled Conjugate Gradient (SCG), and Levenberg–Marquardt (LM), along with a nonlinear autoregressive (NAR) network. Applying mean absolute percentage error (MAPE) as the performance metric, the results showed that all models effectively captured the data's nonlinearity, though their accuracy varied significantly. The NAR model proved to be the most accurate, with a minimum average MAPE of 2%. The Bayesian Regularization (BR) algorithm achieved superior performance (average MAPE: 4.50%) among the feedforward ANNs. Notably, the ARIMA, SARIMA, and ANN-LM models exhibited similar performance. Accordingly, the NAR model is recommended as the optimal choice for long-term traffic prediction. Implementing these models with optimal design will enhance long-term traffic volume forecasting, supporting sustainable transportation and improving intelligent highway operation systems.

**Keywords:** ARIMA; Artificial Neural Networks; long term prediction; nonlinear autoregressive; rural highway; traffic volume



Academic Editor: Giulio Erberto Cantarella

Received: 24 October 2025

Revised: 14 November 2025

Accepted: 19 November 2025

Published: 24 November 2025

**Citation:** Al-Turki, M. Comparative Analysis of Parametric and Neural Network Models for Rural Highway Traffic Volume Prediction.

*Sustainability* **2025**, *17*, 10526.

<https://doi.org/10.3390/su172310526>

**Copyright:** © 2025 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Transportation infrastructure availability and efficiency are essential for sustainable development and crucial to a country's economic growth. The success of a transportation system, such as highway networks, depends on its ability to safely, efficiently, and cost-effectively move individuals and goods between locations [1]. Various approaches, including models and theories, have been developed to evaluate and forecast trends in highway traffic flow. The primary objectives of traffic prediction are to support transportation planning and optimize traffic management concerning environmental preservation, economic efficiency, accessibility, reliability and mobility.

Traffic volume prediction involves estimating the number of vehicles expected to use a specific highway at future times [2]. For implementing advanced control strategies, traffic management is improved with the help of accurate traffic volume prediction by providing early information to the engineers [3]. In order to tackle traffic challenges, traffic volume prediction helps traffic planners and engineers to design and develop efficient highways. For safe and efficient roads, congestion can be reduced by managing real-time traffic data. Urbanization is expected to reach 68% by 2050. The rate of urbanization has increased from 2% in 1800 to over 50% in 2007 with over 20 million people moving from rural to urban zone each year. Cities make up to 3% of total land area contributes to 80% of global Gross Domestic Product (GDP) [4]. Poor traffic management has caused death to approximately 1.35 million people globally. This high death toll is serious economic and environmental issue [5]. The maintenance cost is increase for road surface issues due to increase in the road network [6].

Predicted traffic volumes supply critical data for a variety of aspects of managing highways; including planning, design, and decision-making for operations. Advanced highway management strategies rely upon Intelligent Transportation Systems (ITS) that are typically regarded as being more sustainable and therefore preferred over expensive construction alternatives. ITS is dependent on sophisticated communications technology within both vehicles and roadway infrastructure [7]. Cheng et al. [8] used data from 99 urban areas in the United States collected between 1994 and 2014 and determined that implementation of 511 intelligent transportation systems resulted in significant reductions in traffic congestion with a total of 175 million hours of travel saved, an estimated annual savings of over \$4.7 billion, the saving of approximately 53 million gallons of fossil fuel and reduction of more than 10 billion pounds of carbon dioxide emissions. In addition to economic impacts, the social impact of traffic relates to public health and safety. Vehicle emissions have dramatically increased because of continued urbanization and increasing traffic densities in major metropolitan areas [9]. There has been extensive research linking vehicle traffic, roadway geometrics, and accident frequency, demonstrating that improvement in highway geometries results in decreased crash occurrences [10]. Predictions of long-term traffic volume will allow for more accurate budgeting and scheduling of roadway maintenance, resulting in more economically sustainable use of maintenance funds, longer life expectancy of infrastructure and reduced costs associated with reactive maintenance [11].

However, traffic forecasting needs to be reliable for ITS to be implemented effectively. The prediction of traffic volume provides a basis for highway delay and service level assessments which can assist in reducing traffic congestion by implementing specific traffic control strategies or optimal routing [12]. For highway users, traffic prediction provides valuable information, such as current traffic conditions, expected trip durations, and highway status. Real-time and forecasted traffic data are relied upon by the majority of route planning and navigation tools [13]. Additionally, traffic forecast contributes to intersection management by enabling adaptive traffic signal operations.

An extensive literature review reveals that many researchers have contributed to developing various approaches for short-term traffic flow prediction. However, few investigations have addressed long-term traffic volume forecasting, primarily due to the limited availability of reliable long-term traffic data. Additionally, majority of current research in both short-term and long-term prediction rely on small datasets or simulation data [14]. Some of these studies have used only a single dataset for development, testing, and comparison of their proposed methods, which raises concerns about the models' generalizability and accuracy on other or new data [15]. The majority of research to date has employed parametric methods, which have certain limitations. Long-term traffic volume

prediction using nonparametric techniques, such as deep learning and artificial neural networks (ANNs), is still in the early stages and requires further investigation.

This study aims to evaluate various prediction models for long-term forecasting of daily traffic volume on rural highways in Saudi Arabia. While many studies focus on short-term urban traffic prediction, this research addresses a significant gap by benchmarking models for long-term rural highway traffic forecasting, an area underrepresented in existing literature. The main contributions are:

- The identification of NAR network as a highly accurate and data-efficient model for long-term rural traffic prediction, achieving an average MAPE of 2%, demonstrating that accurate forecasts can be achieved primarily using past traffic data with minimal reliance on extensive exogenous inputs.
- The finding that traditional parametric models (ARIMA, SARIMA) can compete with, and in some cases outperform, complex neural network architectures (specifically ANN-SCG and ANN-LM in this study), challenging the common assumption that nonparametric methods are universally superior for this task.
- A practical benchmark and guidance for transportation planners, comparing six diverse models on real-world data and highlighting the trade-offs between accuracy, model complexity, and computational cost for rural highway applications.
- This study serves as the first comprehensive evaluation for long-term rural traffic forecasting in Saudi Arabia and establishes a foundational reference for the Saudi Ministry of Transport and future researchers in the region, guiding model selection based on accuracy, complexity and cost.

The results from this research can aid highway network management by applying and expanding intelligent transportation systems.

In this paper, Section 2 reviews traffic prediction modeling, key factors, and relevant studies, Section 3 details the materials, methods, data collection, models, and evaluation framework, Section 4 discusses the results and compares model performances and Section 5 focuses on main findings, recommendations and future research.

## 2. Literature Review

### 2.1. Traffic Prediction Modeling

Traffic management benefits from traffic prediction models that use historical data to forecast future conditions [16]. These models are generally classified into short-term (e.g., less than 30 min) and long-term (e.g., a day or a week) forecast [17–20]. Recent advancements in methods have enabled more accurate collection of traffic big data [18]. Examples of data collection techniques include various highway sensors such as loop detectors, radars, cameras, GPS-enabled mobile applications, and probe vehicles.

### 2.2. Key Factors of Traffic Volume Prediction

Several factors can influence future traffic volume on any highway. These factors are generally related to highway characteristics such as distance, capacity, and quality; trip characteristics including purpose and time; and highway users' attributes like age and other economic factors such as fuel prices, tolls, taxes, and the availability of alternative transportation modes [21]. The traffic volume is significantly affected by economic considerations such as toll charges, fuel costs and taxes [22,23].

### 2.3. Prediction Methods Groups

Parametric and nonparametric methods, based on theories such as artificial intelligence and statistics, are utilized by researchers to forecast short-term and long-term traffic

data [24]. These methods are also classified as machine learning methods and classical statistical methods [25].

Parametric methods explicitly specify and examine how inputs relate to outputs through a parameterized and interpretable model, enabling precise quantitative analysis of their relationship. Typically, parametric methods rely on strong assumptions such as the normality of residuals, a defined model structure, and the stationarity of the time series. In nonlinear and disordered traffic flow data, these assumptions may not hold and often result in poor performance. Conversely, the advantages of parametric methods include their ease of implementation and interpretability. They offer quick learning from limited training data and can perform well even with imperfect data fitting [26]. The statistical time series analysis techniques form the foundation of the most popular parametric methods [14,15,27], including historical averaging [14], smoothing techniques, growth trends [28,29], ARIMA model and SARIMA, Kalman filtering models and various regression models [30–32].

Using data-driven approaches, nonparametric methods analyze how inputs relate to outputs [33]. In contrast to parametric methods, which depend on pre-defined relationships, these approaches learn directly from training data, which allows for flexible modeling of different functional forms without prior assumptions. Moreover, intelligent learning algorithms are capable of recognizing nonlinear relationships in data. Poor interpretability is a common challenge with these methods, which require intensive training procedures and large amounts of data. Nonparametric regressions (e.g., k-nearest neighbor (k-NN) methods) [14,34,35], machine learning techniques (e.g., ANN [36] and SVM [14,18,28,29,37,38]) and deep learning neural networks [20,39,40] are frequently used in traffic prediction studies.

Recently, among various nonparametric approaches, ANNs have been prominently considered for prediction tasks. Capacity to categorize patterns, predict outcomes and make decisions based on past experience forms the basis for ANN optimization algorithms, which are developed based on concepts derived from scientific research focusing on the nature of the brain [37]. Given large datasets and limited knowledge of the underlying processes, an ANN consists of a network of interconnected neurons that can receive, process, and transmit signals [14]. Predicted and actual outputs are fundamental to supervised learning, a subset of ANN training that minimizes the error between them. Conversely, some ANNs utilize unsupervised learning, such as self-organization, where outputs are unavailable, and the network organizes input cases into a useful configuration. Several studies have applied ANNs in traffic prediction [41,42], employing techniques such as backpropagation, RBANN, FNM, SVR, GAs, SAE, SVM, CNN and MLP.

#### 2.4. Traffic Prediction Methods

Recent research has focused on improving traffic prediction and management. Xu et al. [43] proposed MetaSTC framework, which employs a spatio-temporal clustering strategy and a meta-learner to address the challenges of large-scale, heterogeneous traffic data. Yang et al. [44] introduce the parallel self-learned and predefined joint spatial-temporal GCN (PSPJSTGCN), a traffic flow forecasting model that combines self-learned and predefined graph structures using a gated mechanism to enhance the extraction of spatial-temporal dependencies. Xiao et al. [45] introduce the adaptive spatiotemporal dynamic graph convolutional network (AST-DGCN), that utilizes dynamic graphs, self-attention, and residual correction. Zhou et al. [46] introduce a large-scale spatio-temporal multimodal fusion framework that uses Convolutional Neural Networks (CNNs) for spatial feature extraction and Recurrent Neural Networks (RNNs) for accurate traffic prediction. This framework can perform location-based predictions and has been successfully applied to real-world large-scale maps, demonstrating its practical utility in urban traffic management.

Lee [47] introduces the HI-XGB model that uses Shapley values and XGBoost to achieve highly accurate urban traffic speed predictions. Min et al. [48] propose a deep multimodal model that fuses data from different sensors to accurately estimate traffic speed. Shen [49] introduces an interpretable XGBoost model for accurate long-term traffic speed prediction. Nesa and Yoon [50] employs regression models and XAI techniques to predict traffic speeds. Wiczorek et al. [51] presents an adaptive vehicle recognition system using fused energy and deep optical flow features with graph mining and an ANN for high performance in dense traffic. Satyanarayana et al. [52] achieves high-accuracy, real-time vehicle detection on low-power devices using a CNN trained on road marks. Singh et al. [53] combines YOLOv8 with Transformers and a modified pyramid pooling model for robust real-time vehicle detection and tracking. Bakirci [54] explores YOLOv8 for vehicle detection in aerial imagery. These innovative approaches reflect a continued focus to address the challenges of transportation systems. The following sections discuss the latest studies on short-term and long-term traffic volume prediction.

#### 2.4.1. Short-Term Traffic Prediction

Among parametric methods, time series models are frequently employed for short-term traffic prediction [55,56]. Due to their low complexity, ARIMA models, originally established by Box and Jenkins, are broadly applied in traffic prediction and several approaches for learning and predicting time series data, which ranges from basic time series analysis to newer methods utilizing ANNs [57]. Forecasting future values is attained by applying a mathematical function that represents historical behavior at regular intervals to the data [58]. Reflecting the effect of historical data while managing uncertainty, the predicted values are modeled as a linear combination of error terms and past data. Smith and Williams [59] applied ARIMA approach to predict single-point traffic flow. Lee et al. [17] found that a specific subset of ARIMA models provided the best prediction accuracy. Box and Jenkins developed Seasonal ARIMA (SARIMA), which is effective for data with strong seasonal patterns. Li et al. [20] observed that SARIMA models deliver accurate predictions. Hong et al. [60] also applied a SARIMA model, which incorporates seasonal differencing to account for peak and non-peak flow periods. However, they highlighted that these models require significant time for outlier detection and parameter estimation. Based on the ARIMA framework, many other time series models have been emerged in the past few years. For example, Chen et al. [12] developed a Logistic Smooth Transition AutoRegressive (LSTAR) model for short-term traffic prediction and proposed a new parameter estimation method (LSTARIMA), which demonstrated higher accuracy than other time series models. Their research also discovered that an optimized ARIMA model can perform effectively compared to the conventional ARIMA by fine-tuning through sequential traffic data analysis.

The rapid advancement in data acquisition, storage, and management technologies has enhanced the potential of nonparametric methods for future traffic prediction [3,58]. Additionally, with the emergence of artificial intelligence applications, new approaches have been introduced into traffic prediction, primarily involving machine learning techniques such as deep learning neural networks and ANN [61]. Recent studies have applied various nonparametric methods for short-term traffic forecasting and compared their performance with standard parametric models [62–64].

Among nonparametric and artificial intelligence approaches, ANNs have been widely applied for traffic prediction. Chen [37] discussed decomposition models, Winters' models, ARIMA, and various neural network models for traffic volume forecasting. His results show that the radial basis function neural network (RBFNN) outperforms other models in volume prediction. Smith et al. [14] applied backpropagation ANN for traffic forecasting,

finding it superior to traditional time series models due to its ability to capture high variability in traffic data [20]. For the short-term traffic prediction, Park et al. [65] compared standard ANN with an RBF-enhanced ANN, concluding that the RBF model is more accurate and requires less training time. Kumar et al. [66] used a multilayer perceptron (MLP) for predicting accurate short-term traffic using time, density, speed, and day of the week. Siddiquee and Hoque [29] utilizes backpropagation ANN for predicting accurate daily traffic volume. Park [65] created a hybrid neural-fuzzy model, which integrates fuzzy C-means with RBF neural networks for categorizing traffic patterns and improving prediction accuracy. For traffic flow prediction, Alsehaimi et al. [67] recommended the Adaptive Spatio-Temporal Attention-Based Multi-Model (ASTAM), which improves the MAE, RMSE, and MAPE metrics.

#### 2.4.2. Long-Term Traffic Prediction

Limited literature is available on the use of parametric and non-parametric models for long-term traffic prediction. Yin et al. [68] introduce the Dual-module Adaptive Transformer and Spatio-Temporal Attention Network (DAT-STAN) to improve traffic flow prediction by capturing traffic patterns and integrating spatio-temporal features. Ali et al. [69] proposed the Multi-scale Attention-Based Spatio-Temporal Graph Convolution Recurrent Network (MASTGCNet) for enhancing traffic prediction accuracy in smart city networks. This technique combines multi-scale feature extraction, dual attention mechanisms and a blend of GRUs and GCNs to capture complex spatio-temporal relationships. It addresses the limitations of traditional fixed-graph methods by dynamically modeling correlations and long-term dependencies, while also incorporating resource allocation strategies to reduce energy consumption in edge computing. Experimental reveal that MASTGCNet significantly outperforms existing approaches, representing a valuable advancement in intelligent traffic management. Table 1 summarizes studies related to long-term traffic forecasting with various parametric and nonparametric methods.

**Table 1.** Summary of the published long-term traffic prediction studies.

| Authors                       | Prediction Methods                                                                                         | Main Findings                                                                                                         |
|-------------------------------|------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------|
| Yasdi [70]                    | Recurrent neural networks with backpropagation                                                             | The model showed good results for traffic prediction                                                                  |
| Stutz and Runkler [71]        | Fuzzy clustering method                                                                                    | The method showed promising results for traffic prediction                                                            |
| Yin [72]                      | Fuzzy neural model (FNM); ANN with backpropagation                                                         | FNM provided more accurate predictions compared to Backpropagation networks                                           |
| Ahmed and Gazder [28]         | ANNs with Multi-layer perceptron (MLP); Linear regression                                                  | The MLP has better accuracy than the linear regression technique                                                      |
| Al-MASAEID AND Al-Omoush [59] | Aggregate regression; Disaggregate trend; Empirical Bayesian                                               | Aggregate regression and empirical Bayesian analysis provided similar results. Trend method was not efficient.        |
| Ratrout and Gazder [73]       | ANN with (MLP); ANN with radial basis function neural network (RBANN); Linear regression analysis          | ANNs have better accuracy than linear regression technique                                                            |
| Su et al. [19]                | Nonparametric kernel regression; SARIMA; Back propagation NN; Least Squares Support Vector Machine (LSSVM) | Nonparametric kernel regression is more accurate and effective than all of the compared models                        |
| Khalifa et al. [74]           | Holt-Winters; ARIMA; Random Forest; MLP; AdaBoost; Long Short-Term Memory networks (LSTM); Extra Trees     | Holt-Winters and ARIMA showed unsatisfactory performance. The MLP and LSTM gave a better performance than the others. |

Table 1. *Cont.*

| Authors          | Prediction Methods                                                                  | Main Findings                                                                                                              |
|------------------|-------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------|
| Zang et al. [75] | Residual net and Deconvolutional Neural Network                                     | The model performed better than any other existing model for traffic long-term flow prediction.                            |
| Zang et al. [76] | Multiscale Spatiotemporal Feature Learning Network (MSTFLN)                         | The model could effectively predict the long-term traffic Information.                                                     |
| Wu et al. [77]   | Denoising schemes and support vector machine                                        | Ensemble Empirical Mode Detection (EEMD) outperforms other denoising algorithms in prediction accuracy.                    |
| Zhao et al. [78] | Deep belief networks (DBNs)                                                         | Parallel DBN learning reduces pre-training and fine-tuning times, enhancing efficiency and effectiveness.                  |
| Li et al. [79]   | Gaussian interval type-2 fuzzy set                                                  | Forecasted traffic range fully encompassed the actual traffic volume within upper and lower bounds.                        |
| Tang et al. [80] | DNN based traffic flow model                                                        | The DNN-BTF model, using CNN and RNN to extract spatial-temporal features, outperformed all models.                        |
| He et al. [81]   | Spatio-Temporal Convolutional Neural Network (STCNN) model                          | STCNN model showed significantly better performance than any other predictive model                                        |
| Li et al. [82]   | T2F-LSTM neural network model                                                       | The introduction of interval sets of T2F provided a better LSTM model performance for Long-term traffic volume prediction. |
| Li et al. [83]   | Wavelet-Decomposed Convolutional Neural Network-Long Short-Term Memory (W-CNN-LSTM) | W-CNN-LSTM combines wavelet, CNN, and LSTM for improved long-term traffic flow prediction.                                 |
| Park et al. [84] | Graph Convolutional Network                                                         | This model predicts traffic better, incorporating weather data and outperforming traditional methods.                      |
| Toba et al. [85] | Combination of K-means clustering, LSTM and Fourier transform                       | This method accurately predicts long-term traffic trends, capturing periodicity and variations effectively.                |

### 2.5. Findings of the Literature Review

Studies generally show that nonparametric methods are more suitable and accurate than parametric approaches. However, a few studies report conflicting results, particularly with certain advanced parametric models. Statistical time-series models are easy to use and have low computational complexity, but their performance in long-term traffic prediction suffers due to the high variability and nonlinearity of traffic data, which these models struggle to capture. For example, ARIMA models focus on mean values and, as linear statistical algorithms, face challenges in modeling traffic flow fluctuations and high variability. Conversely, the literature indicates that nonparametric methods are more successful for traffic prediction. They are flexible, data-driven, and do not rely on distribution assumptions [86].

## 3. Materials and Methods

### 3.1. Data Collection

Daily traffic volume in this study represents the vehicle count on a highway in one direction over a 24 h period. Using loop detector pairs in every highway lane for both directions, the Saudi Ministry of Transport (MOT) gathered data from 2013 to 2019. However, due to maintenance or other issues, detector shutdowns occurred, resulting in missing data.

### 3.2. Selection of the Highways Sample

Saudi Arabia has over 60 rural highways connecting remote regions, each differing in characteristics due to their unique economic and social roles [29]. This research aims to predict the daily traffic volume on selected rural highways to evaluate the performance of various prediction models. A representative sample of three traffic datasets from different highways was used, selected based on data availability, highway type, and average monthly traffic volume. The primary criterion was sufficient traffic data over long periods. While expanding the dataset to include more highway types and regions would improve generalizability, it poses significant data collection challenges. This study provides a crucial baseline using the most comprehensive available data. Table 2 presents the selected highways, including their names and features.

**Table 2.** Selected highways and their characteristics.

| Highway Name  | Highway Code | No of Lanes | Highway Distance (km) | Data Period                          | No of Available Days |
|---------------|--------------|-------------|-----------------------|--------------------------------------|----------------------|
| Riyadh—Dammam | RTD          | 3           | 383                   | (13 January 2014)–(10 July 2019)     | 1685                 |
| Riyadh—Qassim | RTQ          | 3           | 317                   | (25 February 2013)–(22 May 2019)     | 1748                 |
| Qassim—Riyadh | QTR          | 3           | 317                   | (26 February 2013)–(02 October 2019) | 1905                 |

### 3.3. Inputs Processing

As discussed in the literature, various factors can influence highway traffic volumes over different time periods. In this study, key factors such as time characteristics, specifically, the date of the month, day of the week and fuel prices were selected as input variables for some of the prediction models. Table 3 illustrates a typical dataset with examples. The date was divided into four components:

- ×1: the day of the month (1–31),
- ×2: the month (1 January to 12 December),
- ×3: the year (1 for 2013 up to 7 for 2019),
- ×4: the day of the week (1–Sunday to 7–Saturday).

Additionally, ×5 represents the fuel price for each day, with four possible values: 0.45, 0.75, 1.37, and 1.44 Saudi Riyals, reflecting fuel prices in Saudi Arabia over the past seven years. The total daily traffic volume ( $y$ ) is represented in the last column.

Based on literature review findings and the availability of specific data, this research selected these inputs. To reduce model complexity and improve predictive accuracy, some outliers in the original traffic volume dataset were adjusted.

**Table 3.** Example of the Typical Dataset Structure.

| Date<br>(×1) | Month<br>(×2) | Year<br>(×3) | Day of the Week<br>(×4) | Fuel Price (×5)<br>(Saudi Riyals) | Total Daily Volume ( $y$ )<br>(No. of Vehicles) |
|--------------|---------------|--------------|-------------------------|-----------------------------------|-------------------------------------------------|
| 13           | 1             | 1            | 2                       | 0.45                              | 13,378                                          |
| 14           | 1             | 1            | 3                       | 0.45                              | 14,322                                          |
| 15           | 1             | 1            | 4                       | 0.45                              | 15,704                                          |

### 3.4. Data Exploration

Before developing prediction models, traffic volume data from three highways were analyzed to identify general patterns and relationships with input variables. Seasonal, monthly, and daily patterns were revealed in the analysis of daily traffic volumes over the entire year. The data show nonlinear and disordered patterns. As expected, traffic volume is strongly correlated with the day of the week (×4) and significantly related to daily

time characteristics, including the date ( $\times 1$ ) and month ( $\times 2$ ). Additionally, each highway exhibits different traffic patterns due to variations in their economic and social significance.

### 3.5. Selection of the Prediction Models

Suitable prediction models were selected and applied using the three datasets described earlier to assess their effectiveness for long-term traffic volume forecasting. Six models were chosen: two parametric models (ARIMA and SARIMA) for comparison, and four different types of ANN. Classical statistical methods, fully connected ANNs, and a recurrent NAR network are included in this selection, which provides a broader benchmark. The chosen set, representing parametric, nonparametric, and neural network approaches, is broadly used, despite the existence of other models, e.g., Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), and LSTM. Without the complexity of hybrid or highly advanced architectures, this enables a clear evaluation of their suitability for long-term rural highway traffic prediction and provides a solid foundation for future comparisons. For linear and seasonal time-series problems, parametric models acted as established baseline methods due to their proven performance [87]. In traffic volume forecasting, feedforward ANNs are a key class of non-parametric models that can learn complex, nonlinear relationships without strong assumptions [88]. Sharing the ability to model temporal dependencies with more complex models like LSTM, the NAR network was specifically selected for its simpler architecture. All models were implemented using MATLAB R2018b. Following sections discuss Statistical Time Series models, Artificial Neural Network and Neural Network Time Series.

### 3.6. Statistical Time Series Models

ARIMA is a non-seasonal linear time series model characterized by three key parameters:  $p$  (autoregressive order),  $D$  (degree of non-seasonal differencing), and  $q$  (moving average order). Equation (1) shows that the predicted value in the ARIMA model is a linear combination of past observations and errors.

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q} \quad (1)$$

where  $p$  and  $q$  are autoregressive and moving average,  $Y_t$  is the actual observation and  $\epsilon_t$  is the random error at  $t$ , and  $\theta_1, \theta_2, \dots, \theta_q$  and  $\phi_0, \phi_1, \dots, \phi_p$  are coefficients, respectively [89].

As a time series model, ARIMA uses past traffic volumes to forecast future traffic. Based on the literature, ARIMA parameters (constant, AR, and MA) can be estimated using the Box–Jenkins method or other approaches. In this study, the optimal values of  $p$ ,  $D$ , and  $q$  were selected through trial and error, guided by the minimum AIC. These values were then used to estimate ARIMA parameters for the prediction model. The SARIMA model, which captures seasonality, employed the same parameter estimation approach [90].

### 3.7. Artificial Neural Networks (ANNs)

The primary goal of ANNs is to model the relationship between inputs and outputs by training on data to produce accurate target outputs [91]. In this study, MATLAB's neural network fitting app was used. The ANN consists of three layers: input, hidden, and output, and employs the sigmoid function as the transfer function. Several training algorithms are available; however, this research selected BR, LM and SCG to compare their performance and determine which works best for long-term traffic volume prediction on rural highways. During training, the network memorizes input–output relationships, and the error decreases to a small value. However, when tested on new data, the error may become large, a phenomenon known as overfitting. This can be mitigated through

generalization techniques that restrict the network size, although determining the optimal size is challenging and typically approached with different methods.

The ANNs used in this research are feedforward neural networks (FFNNs) that model the relationship between the input variables  $x = [x_1, x_2, \dots, x_n]$  and the output  $y$  through a series of nonlinear transformations. The network consists of an input layer, one hidden layer, and an output layer. The output of each neuron  $jj$  in the hidden layer is computed as:

$$h_j = f^{(1)}(\sum_{i=1}^n w_{ji}^1 x_i + b_j^i) \quad (2)$$

where  $w_{ji}^1$  are the weights connecting input  $ii$  to hidden neuron  $j$ ,  $b_j^i$  is the bias term for neuron  $j$  and  $f^{(1)}$  is the activation function. The final network output is then:

$$\hat{y} = f^{(2)}(\sum_{j=1}^m w_j^2 h_j + b^2) \quad (3)$$

where  $m$  is the number of hidden neurons,  $w_j^2$  are the weights from the hidden layer to the output,  $b^2$  is the output bias, and  $f^{(2)}$  is a linear activation function for regression tasks. The goal of training is to find the parameters  $\theta = \{\mathbf{W}^{(1)}, \mathbf{b}^{(1)}, \mathbf{W}^{(2)}, \mathbf{b}^{(2)}\}$  that minimize a loss function, such as the MSE, between the predicted outputs  $\hat{y}$  and the actual targets  $y$ .

The first two algorithms, SCG and LM, use early stopping to address this issue. Early stopping divides data into three sets: training, validation, and test. The model's weights and biases are trained on the training set, while the validation set monitors generalization and determines when to stop training. The test set evaluates the model's performance on unseen data.

### 3.7.1. Scaled Conjugate Gradient (SCG)

The SCG algorithm adjusts weights by searching along the conjugate gradient, offering faster convergence than first-order methods by avoiding costly line searches. Developed by Möller [92], it remains popular among researchers with large datasets who prioritize convergence speed over memory. In SCG algorithm the Hessian matrix can be approximated by

$$s_k = \frac{\dot{E}(w_k + \sigma_k p_k)}{\sigma_k} + \lambda_k p_k \quad (4)$$

$$\beta_k = \frac{(\|g_{k+1}\|^2 - g_{k+1}^T g_k)}{g_k^T g_k} \quad (5)$$

$$p_{k+1} = -g_{k+1} + \beta_k p_k \quad (6)$$

where  $E$  is the total error function and  $\dot{E}$  is the gradient of  $E$ ,  $\lambda_k$ ,  $\beta_k$  and  $\sigma_k$  are scaling factors [93]. The algorithm itself is complex and completes its search in multiple steps [94].

### 3.7.2. Levenberg–Marquardt (LM)

The LM is a fast optimization technique based on numerical methods [95], closely related to quasi-Newton methods. It achieves near second-order training speed by approximating Hessian matrices ( $H$ ), which involve second derivatives of the performance function at current weights and biases. By avoiding these complex calculations, LM converges faster than conjugate gradient methods. Instead, it approximates  $H$  with  $J^T J$ , where  $J$  contains first-order derivatives of network errors, making the computation more efficient. This approximation is effective when the performance index is a sum of squares, common in feedforward networks:

$$H = J^T J \quad (7)$$

The update rule then becomes:

$$x_{n+1} = x_n - \frac{J^T e}{[J^T J + kI]} \quad (8)$$

where  $e$  represents network errors,  $I$  is the identity matrix, and  $k$  is a scalar that interpolates between gradient descent and Newton's method based on performance criteria [96].

### 3.7.3. Bayesian Regularization (BR)

Regularization is used by BR to reduce overfitting [97]. To promote smoother responses and improve generalization, this approach uses a modified performance index (PI) that penalizes large weights and biases. The modified PI adds an MSWB term to the standard mean squared error (MSE):

$$PI = \alpha MSE + (1 - \alpha)MSWB \quad (9)$$

where MSE is the mean squared error, MSWB is the mean sum of squares of weights and biases, and  $\alpha$  is the performance ratio parameter. Selecting the optimal  $\alpha$  is crucial; a small value reduces performance, while a large value risks overfitting. In BR,  $\alpha$  is automatically estimated by treating weights and biases as random variables with known distributions, relating  $\alpha$  to their variances.

All ANNs in this study had a single hidden layer with five input nodes date ( $\times 1$ ), month ( $\times 2$ ), year ( $\times 3$ ) day of the week ( $\times 4$ ) and fuel price ( $\times 5$ ) and one output node (total daily volume  $y$ ). Data were split into 70% for training, 15% for validation and 15% for testing. To identify the optimal number of hidden neurons, models were tested with 5 to 600 neurons; 125 neurons yielded the lowest MAPE on the validation set. This number was used for all models to ensure consistency and optimal performance.

### 3.8. Neural Network Time Series

This is another neural network type used for fitting problems, employing a two-layer feedforward architecture. It utilizes dynamic filtering with previous time series values to predict future data, using tapped delay lines for nonlinear filtering and prediction. In this study, the NAR model was chosen to represent this group. The NAR model can be trained with any training function and is mathematically expressed as:

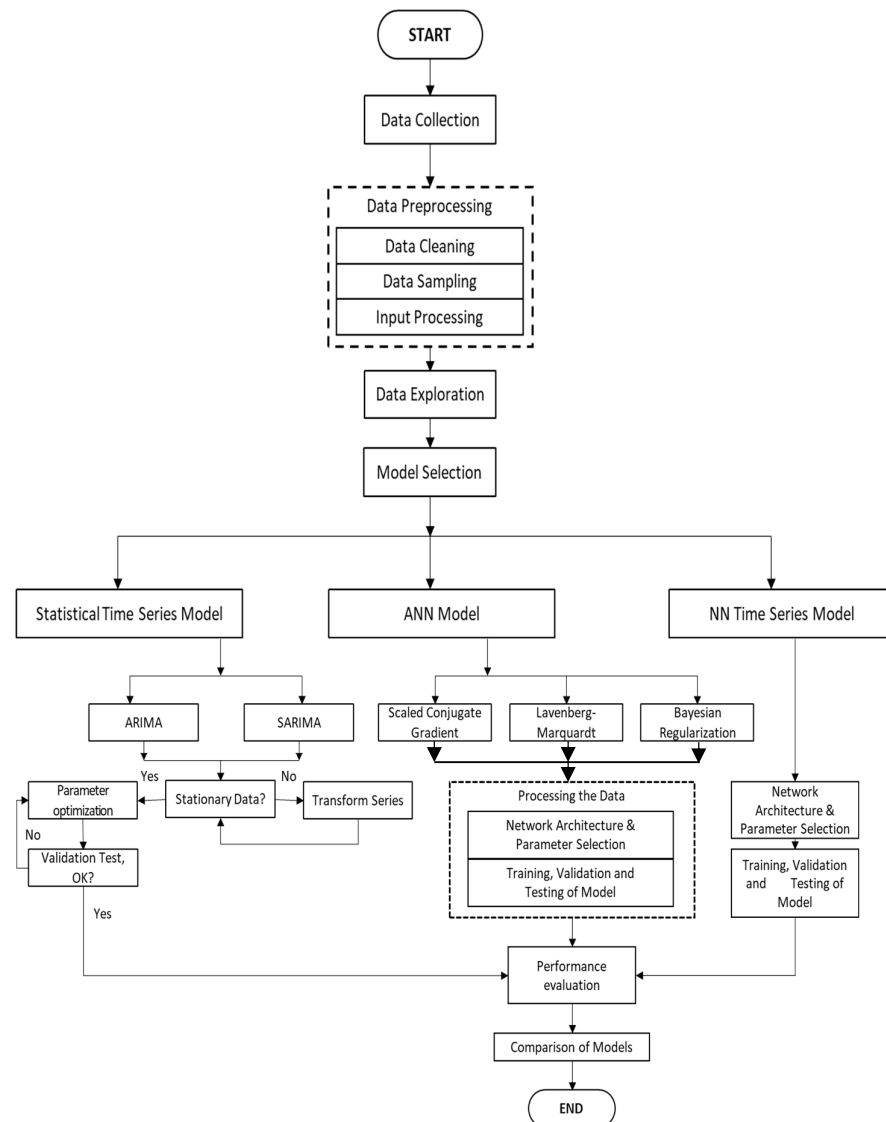
$$\hat{Y} = f(y(t-1) + y(t-2) + \dots + y(t-d)) \quad (10)$$

where  $f$  is a nonlinear function and predicted values depend on the earlier values of the regressed  $d$  [98]. As a time series model, it has a single input (past traffic volumes) and predicts the next day's volume. The NAR model was implemented with a single hidden layer and trained using Bayesian regularization. Based on literature, 50 hidden neurons were selected. The key hyperparameter is the number of feedback delays; performance was tested for delays ranging from 1 to 200, with 200 delays yielding the lowest MAPE. Data were split into 70% training, 15% validation, and 15% testing, consistent with the other models.

### 3.9. Model Evaluation

Various techniques can be employed to evaluate the effectiveness of prediction models, including regression analysis, mean absolute error (MAE), Mean Absolute Percentage Error (MAPE), Mean Square Error (MSE), Variance of Absolute Percentage Error (VAPE) and the probability of percentage errors. MAPE is suitable for transportation planning, is scale-independent and unit-free, and adjusts for differences in error magnitude. Due

to its ease of interpretation as an average percentage error, MAPE was used to evaluate all prediction models in this study, which helps in stakeholder communication. It enables consistent comparison across datasets that have different traffic volumes. RMSE and  $R^2$  highlight various aspects of error, but the clarity and consistency of MAPE's trends provide a more reliable foundation for model comparison [99–101]. The prediction methodology is illustrated with a flowchart diagram (Figure 1).



**Figure 1.** Methodology flow diagram for Long-term traffic prediction [93,102].

#### 4. Results and Discussion

A critical step in this study was the systematic optimization of hyperparameters for neural network models to ensure fair comparison and maximize predictive accuracy. For feedforward ANNs, the number of hidden neurons was optimized, performance was evaluated across 5 to 600 neurons, with the lowest MAPE on the validation set achieved at 125 neurons. This configuration was used for all ANN variants (BR, SCG, LM). For the NAR model, the key hyperparameter was the number of feedback delays, ranging from 1 to 200. Results in Table 7 indicated that accuracy improved with increasing delays, plateauing around 25 delays. The minimum MAPE was consistently obtained with 200 delays, which was selected for the final NAR model despite higher computational costs. For parametric models (ARIMA and SARIMA), the Box-Jenkins methodology was employed; the optimal

parameters ( $p, D, q$ ) and seasonal parameters ( $ps, Ds, qs$ ) were determined by minimizing the Akaike Information Criterion (AIC) for the Riyadh-Dammam (RTD) dataset. The optimal structures were  $(7, 1, 2)$  for ARIMA and  $(7, 1, 2) \times (7, 1, 2)$  for SARIMA, which were then applied across all datasets for consistency.

#### 4.1. Results of the Statistical Time Series Models

Initially, the RTD highway dataset was used to determine the optimal hyperparameters for ARIMA ( $p, D, q$ ) and SARIMA ( $p, D, q$ )  $\times$  ( $ps, Ds, qs$ ) models. These optimal parameters were then applied to develop models for three datasets (RTD, RTQ, and QTR). Table 4 presents the model characteristics and their prediction accuracy, assessed via MAPE. Results indicated that both ARIMA and SARIMA models achieved promising prediction accuracies across all datasets, with average MAPE values of 7.42% and 7.00%, respectively. The SARIMA model demonstrated marginally better performance. These findings suggest that, despite their assumptions, both parametric models effectively predicted nonlinear traffic volume data.

**Table 4.** ARIMA and SARIMA models characteristics and MAPE (%) results of RTD, RTQ and QTR highways.

| Highway Code | Time Series Model                                                         |                                                                                                                      |
|--------------|---------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------|
|              | ARIMA                                                                     | SARIMA                                                                                                               |
| RTD          | MAPE = 7.098%<br>( $p, D, q$ ) = $(7, 1, 2)$<br>$AIC = 2.983 \times 10^4$ | MAPE = 6.599%<br>$AIC = 3.114 \times 10^4$<br>( $p, D, q$ ) $\times$ ( $ps, Ds, qs$ ) = $(7, 1, 2) \times (7, 1, 2)$ |
| RTQ          | MAPE = 7.218%<br>( $p, D, q$ ) = $(7, 1, 2)$<br>$AIC = 2.814 \times 10^4$ | MAPE = 6.836%<br>$AIC = 2.858 \times 10^4$<br>( $p, D, q$ ) $\times$ ( $ps, Ds, qs$ ) = $(7, 1, 2) \times (7, 1, 3)$ |
| QTR          | MAPE = 7.926%<br>( $p, D, q$ ) = $(7, 1, 2)$<br>$AIC = 3.324 \times 10^4$ | MAPE = 7.574%<br>$AIC = 3.429 \times 10^4$<br>( $p, D, q$ ) $\times$ ( $ps, Ds, qs$ ) = $(7, 1, 2) \times (7, 1, 2)$ |
| Average MAPE | 7.400                                                                     | 7.000                                                                                                                |

#### 4.2. Results of the Artificial Neural Networks

The three selected training methods were applied on each dataset to evaluate and compare their prediction capabilities. Table 5 presents the ANN-MAPE results against different hidden neurons.

**Table 5.** ANN-MAPE (%) results of different number of hidden neurons.

| No. of Hidden Neurons | 5    | 25   | 50   | 75   | 100  | 125  | 150  | 200  | 500  | 600  |
|-----------------------|------|------|------|------|------|------|------|------|------|------|
| MAPE (%)              | 13.2 | 6.37 | 4.92 | 3.95 | 3.56 | 3.22 | 3.25 | 3.48 | 3.82 | 3.70 |

Table 6 presents the prediction accuracies of the three neural networks (with 125 neurons), evaluated using MAPE for RTD, RTQ, and QTR highways. Overall, all three ANNs demonstrated acceptable predictive performance, though their accuracies varied. The Bayesian Regularization (BR) training method yielded the lowest MAPE values across all datasets, with an average MAPE of 4.50%. The LM method produced higher MAPE values, averaging 8.00%, while the SCG method resulted in the highest errors, with an average MAPE of 13.80%. These results highlight the importance of selecting an appropriate training method in addition to optimizing the number of hidden neurons.

**Table 6.** ANNs MAPE (%) results of RTD, RTQ and QTR highways.

| ANN Training Method/Highway Code | RTD    | RTQ    | QTR    | Average |
|----------------------------------|--------|--------|--------|---------|
| Bayesian Regularization (BR)     | 3.222  | 4.432  | 5.869  | 4.466   |
| Scaled Conjugate Gradient (SCG)  | 14.618 | 12.409 | 14.459 | 13.800  |
| Levenberg–Marquardt (LM)         | 7.484  | 6.936  | 9.598  | 8.000   |

#### 4.3. Results of the Neural Network Time Series

Table 7 presents the prediction accuracy of the NAR model, evaluated using MAPE for RTD, RTQ, and QTR highways. Overall, the results indicate that the NAR model achieved high prediction accuracy, with MAPE values of 1.65%, 1.70% and 2.57% for RTD, RTQ, and QTR, respectively, at a delay of 200. Increasing the delay number from 1 to 200 significantly reduced MAPE values; however, minimal differences were observed beyond a delay of 25. MAPE values for delays of 50, 100 and 200 were nearly similar across all datasets.

**Table 7.** NAR model MAPE (%) results of RTD, RTQ and QTR highways.

| Delay Numbers | RTD    | RTQ    | QTR    | Average |
|---------------|--------|--------|--------|---------|
| 1             | 15.221 | 10.371 | 13.970 | 13.130  |
| 2             | 9.623  | 9.281  | 10.936 | 9.990   |
| 3             | 7.605  | 8.710  | 8.036  | 8.100   |
| 4             | 6.309  | 8.416  | 7.614  | 7.430   |
| 5             | 5.602  | 7.768  | 6.886  | 6.700   |
| 6             | 4.509  | 7.038  | 6.577  | 6.000   |
| 7             | 3.801  | 5.558  | 5.384  | 4.866   |
| 8             | 3.704  | 5.531  | 5.012  | 4.733   |
| 9             | 3.607  | 5.265  | 4.795  | 4.500   |
| 15            | 3.408  | 5.649  | 5.040  | 4.666   |
| 25            | 2.801  | 4.330  | 4.233  | 3.766   |
| 50            | 1.952  | 3.778  | 3.558  | 3.300   |
| 100           | 1.702  | 4.243  | 2.428  | 2.766   |
| 200           | 1.648  | 1.695  | 2.566  | 2.000   |

#### 4.4. Comparisons of the Applied Prediction Models

Table 8 shows minor variations in prediction performance across highways, attributable to differences in their traffic volume patterns. While all models provided usable forecasts, their accuracy varied substantially, with the NAR model being the most accurate, achieving a minimum average MAPE of 2%. Overall, all models performed well for traffic volume prediction across the three datasets; however, their prediction accuracies differed. The NAR model consistently outperformed others, with the lowest average MAPE (2%). In contrast, the ANN with the SCG training method was the least accurate (average MAPE: 13.83%), performing worse than traditional ARIMA and SARIMA models. This contradicts the common assumption that nonparametric methods are superior to classical time-series models. The ARIMA, SARIMA, and ANN-LM models showed similar performance (average MAPE 7.5%), indicating that well-specified parametric models can be competitive with neural networks for this task. Notably, the strong performance of the NAR model, which relies solely on past traffic volumes, suggests that additional exogenous factors (e.g., date, month) used as inputs in feedforward ANNs may be unnecessary for high accuracy, thereby simplifying data collection. The ANN models, particularly the impact of key variables like the day of the week, month, and date, appear crucial only when used with neural networks; in contrast, ARIMA, SARIMA and NAR models using only past traffic volumes, delivered more accurate results than some ANNs with additional factors. Regarding computational efficiency, ANNs, ARIMA, and SARIMA models ran

significantly faster than the NAR model, which, with 50 hidden neurons and 200 delays, required nearly 19 h for training. The ANN with the Bayesian Regularization (BR) method took longer than other training approaches.

**Table 8.** Comparison of the MAPE (%) of all prediction models.

| Prediction Model Name                                    | RTD    | RTQ    | QTR    | Average |
|----------------------------------------------------------|--------|--------|--------|---------|
| ARIMA                                                    | 7.098  | 7.218  | 7.926  | 7.420   |
| SARIMA                                                   | 6.599  | 6.836  | 7.574  | 7.000   |
| ANN with Bayesian Regularization (BR)                    | 3.222  | 4.432  | 5.869  | 4.500   |
| ANN with Scaled Conjugate Gradient (SCG)                 | 14.618 | 12.409 | 14.459 | 13.830  |
| ANN with Levenberg–Marquardt (LM)                        | 7.484  | 6.936  | 9.598  | 8.100   |
| NN time series with Nonlinear Autoregressive model (NAR) | 1.647  | 1.695  | 2.566  | 2.000   |

Figure 2 illustrates the predicted traffic volumes from all models alongside the actual traffic volumes for RTD, RTQ, and QTR. The shown portion, selected randomly, includes 42 data points representing every fifth day near the end of each dataset. These figures compare the predicted and actual volumes, highlighting how well each model captures traffic patterns. Across all datasets, the NAR model's predictions closely follow the actual data, validating the high accuracy indicated by the earlier MAPE analysis.

The superior performance of the NAR model can be attributed to its architecture, which is fundamentally designed for sequential data. The optimization of the feedback delays to 200 provided the model with an extended memory, enabling it to capture long-range temporal dependencies and complex, nonlinear patterns within the traffic volume time series that other models could not. The ability to model such extended temporal contexts is a key advantage of recurrent architectures over feedforward networks for traffic prediction [45]. The NAR model's architecture is well-suited to capture the intense intermittency and stochasticity of traffic flow [67]. Furthermore, BR incorporates a regularization term that penalizes overly complex models, effectively controlling overfitting and promoting smoother, more generalizable predictions [97]. The use of the Bayesian Regularization training algorithm enhanced learning stability by effectively preventing overfitting, leading to a more robust and generalizable model. In contrast, the linear assumptions of the ARIMA and SARIMA models limited their ability to capture the high stochasticity of the data, while the feedforward ANNs, despite their nonlinear capabilities, lacked inherent temporal memory and were more sensitive to the choice of training algorithms.

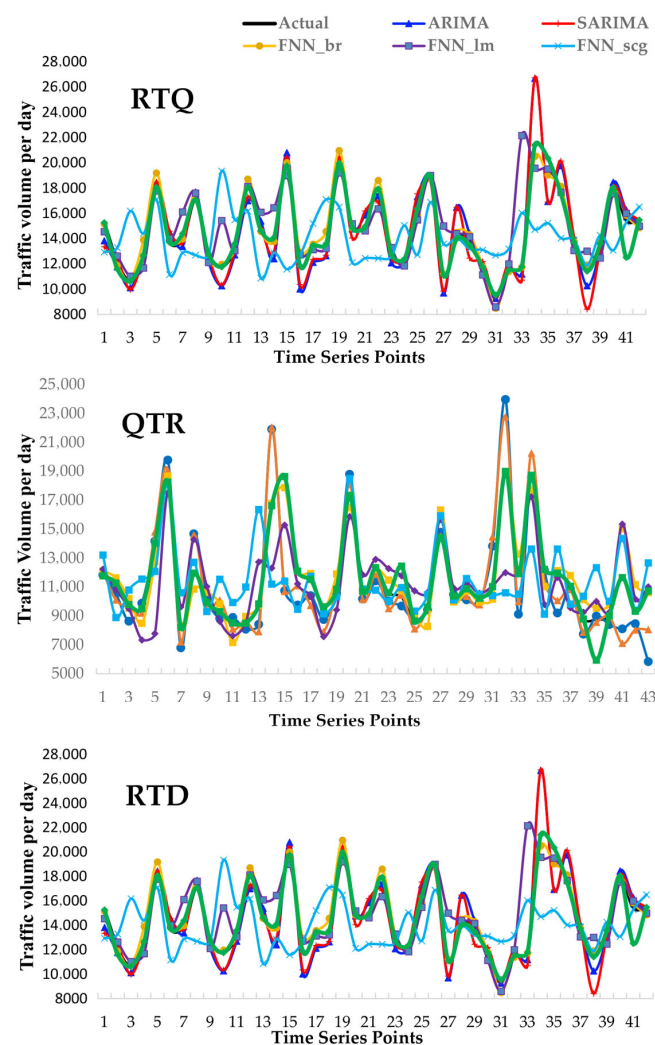
This comprehensive evaluation of six prediction models on three distinct rural highway datasets provides clear evidence for their applicability in long-term traffic volume forecasting. The detailed findings offer several key conclusions with practical benefits for transportation planners and engineers.

1. The NAR model is the most accurate predictor. It consistently achieved the lowest prediction error across all datasets, with an average MAPE of 2%. Its design, which relies solely on past traffic volumes to forecast future values, proved exceptionally effective at capturing the underlying temporal dynamics of rural highway traffic, making it a highly recommended model for applications where forecasting precision is paramount.
2. The superiority of neural networks is not universal. While the ANN trained with Bayesian Regularization (ANN-BR) was a strong performer (Average MAPE: 4.5%), the results challenge the assumption that nonparametric methods always outperform classical ones. The ANN-SCG model performed poorly, and the ANN-LM model showed accuracy comparable to the much simpler and more interpretable ARIMA

and SARIMA models. This indicates that well-specified parametric models remain competitive for this specific task.

3. Model design and configuration are critical to performance. The significant performance gap between ANN-BR and the other ANN training algorithms underscores the importance of the training method and inherent regularization. Furthermore, the systematic search for optimal hyperparameters, such as the number of hidden neurons in ANNs and the feedback delays in the NAR model, was proven to be a necessary step, as these choices profoundly influenced predictive accuracy.
4. A clear trade-off exists between accuracy and computational cost. The NAR model's superior accuracy came at the cost of significantly longer training times (up to 19 h for 200 delays) compared to the near-instantaneous training of ARIMA/SARIMA and the relatively fast training of other ANNs. This highlights a critical practical consideration: the choice of model often involves balancing the need for high accuracy against available computational resources and time constraints for model development.

In summary, this benchmark study demonstrates that effective long-term traffic prediction for rural highways is achievable with multiple modeling approaches. The choice of an optimal model depends on the specific priorities of the project, whether they favor maximum accuracy (NAR), a balance of performance and simplicity (ANN-BR), or computational efficiency and interpretability (ARIMA/SARIMA).



**Figure 2.** Portion of predicted traffic volumes of all the applied models with actual traffic volumes of RTQ, QTR and RTD.

#### 4.5. Implications for Sustainable Transportation

Traffic congestion and suboptimal traffic flow are major sources of excess fuel consumption and pollutant emissions. The relationship between traffic flow and emissions is well-established in international methodologies [9]. Accurate long-term traffic volume forecasts are crucial for modeling vehicle emissions and implementing effective mitigation strategies, as errors can significantly impact emission estimates. The NAR model, with a low MAPE of 2%, offers superior accuracy, reducing uncertainty and enabling transportation agencies to better plan for environmental impacts, development projects, and electrification efforts, ultimately aiding in reducing the carbon footprint of rural highways and supporting national climate goals.

Long term traffic volume forecasts are vital for the economic sustainability of highway infrastructure, as they optimize maintenance costs and resource allocation. The highly accurate NAR model enhances life-cycle cost management by enabling better planning, reducing waste, and extending infrastructure lifespan, ultimately ensuring cost-effective investments and maximizing public funds' returns through a proactive, data-driven approach [11].

### 5. Conclusions and Recommendations

Accurate long-term traffic volume forecasts are crucial for infrastructure investment decisions, feasibility studies, and sustainable transportation planning. This study addressed a critical gap by benchmarking forecasting models specifically for rural highways, an area underrepresented in the literature. In addition to that, in the case of the highway charging concept, the traffic volume prediction can help to manage the toll booth system in an efficient way. The predicted volumes analysis will help determine which level of maintenance is required, when it is required, and when is the best time to perform the maintenance works without causing heavy traffic congestion. In addition to all of the above applications, the predicted traffic volume is also useful to predict and analyze some related events such as the number of accidents or the environmental impacts such as air and noise pollution. This research has evaluated six different prediction models for the application of long-term prediction of the daily traffic volume using three datasets of long duration counts of daily traffic volumes. The main conclusions of this research are:

1. Highly accurate long-term traffic prediction for rural highways is achievable. The evaluated models effectively captured the complex, nonlinear patterns in the data.
2. Model choice significantly impacts accuracy. The NAR model was the most accurate (avg. MAPE: 2%), followed by the ANN-BR (avg. MAPE: 4.5%). The ARIMA, SARIMA, and ANN-LM models formed a third tier of performance (avg. MAPE: ~7.5%).
3. The superiority of neural networks is not universal. The poor performance of ANN-SCG and the comparable performance of ANN-LM to traditional ARIMA/SARIMA models demonstrate that well-specified parametric models remain competitive and can outperform certain neural network approaches.

This study underscores that model configuration is paramount. The significant performance differences between ANN training algorithms (BR vs. SCG/LM) highlight that the choice of training method and hyperparameter tuning is as critical as the choice of the model class itself. For accurate traffic prediction using ANNs, it is vital to optimize the number of hidden neurons and select a suitable training method. For the NAR model, the results have concluded that a higher delay number will lead to better prediction accuracy. However, the delay number should have important consideration in designing the NAR model, especially when long-term data of past days are limited.

The prediction results of the NAR model have shown that it is a more preferred and dependable model since the model is totally developed based on data and it is free of the

assumptions of data distribution, highly flexible and easy to use. Using this model for long-term traffic prediction with a smaller number of delays may be more cost-effective since it needs a smaller number of past traffic volumes and required less running time compared to higher delay numbers. The NAR model transforms rural highway planning with its high accuracy achieved through minimal data, improving efficiency and practicality.

For practical implementation, the NAR model is recommended where high accuracy is critical, despite its computational cost. For scenarios requiring a balance of speed and performance, the ANN-BR or even the simpler SARIMA model are viable alternatives. Implementing these models can significantly enhance highway network management and Intelligent Transportation Systems applications.

The models demonstrated strong performance, but their applicability to other regions and cultural contexts needs further validation, particularly for the NAR model. Future research should focus on testing these models across different geographical settings to assess their generalizability. Incorporating additional features such as holiday indicators, school vacations, and weather conditions could improve model accuracy. Exploring these variables, especially in ANN models that handle multiple inputs, may also help explain residual variance during atypical days. Furthermore, while this study systematically optimized key hyperparameters like the number of hidden neurons and NAR delays, a more exhaustive hyperparameter search (e.g., using automated grid or random search routines across a broader parameter space) could be undertaken in future work to ascertain if further marginal gains in prediction accuracy are achievable.

This research supports sustainable transportation by improving long-term traffic forecasts that enhance economic viability, environmental impact mitigation, and traffic safety. It enables proactive planning and resource optimization, ultimately promoting safer, greener, and more efficient transportation systems. Furthermore, long-term traffic volume forecasts are a critical enabler of sustainable transportation systems, directly supporting environmental goals through accurate emission planning and economic goals through optimized resource allocation.

This study establishes a robust benchmark for long-term rural highway traffic prediction. Future work should focus on three key directions to build upon these findings. First, the benchmark should be expanded to include advanced deep learning architectures, particularly LSTM and Transformer networks, to rigorously test their performance against the superior NAR model identified here. Second, a formal sensitivity analysis is crucial to quantify the importance of the various input variables (e.g., day of week, fuel price) in the FFNN models, providing deeper insight into the factors driving predictions. Finally, applying this extended benchmark to a wider variety of highway datasets from different regions is essential to validate the generalizability and robustness of the models' performance across diverse traffic patterns.

**Funding:** This research received no external funding.

**Data Availability Statement:** No new data were created or analyzed in this study. Data sharing is not applicable to this article.

**Conflicts of Interest:** The author declares no conflicts of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

|       |                                          |
|-------|------------------------------------------|
| AIC   | Akaike Information Criterion             |
| ANN   | Artificial Neural Network                |
| ARIMA | Autoregressive Integrated Moving Average |

|            |                                                                                         |
|------------|-----------------------------------------------------------------------------------------|
| ASTAM      | Adaptive Spatio-Temporal Attention-Based Multi-Model                                    |
| AST-DGCN   | Adaptive Spatiotemporal Dynamic Graph Convolutional Network                             |
| BR         | Bayesian Regularization                                                                 |
| BS         | Background Subtraction                                                                  |
| CB         | CatBoost                                                                                |
| CNN        | Convolutional Neural Network                                                            |
| DAT-STAN   | Dual-module Adaptive Transformer and Spatio-Temporal Attention Network                  |
| DBN        | Deep Belief Network                                                                     |
| DNN        | Deep Neural Network                                                                     |
| DOF        | Deep Optical Flow                                                                       |
| EEMD       | Ensemble Empirical Mode Decomposition                                                   |
| EN         | Energy                                                                                  |
| FCM        | Fuzzy C-Means                                                                           |
| FNM        | Fuzzy Neural Model                                                                      |
| GA         | Genetic Algorithm                                                                       |
| GDP        | Gross Domestic Product                                                                  |
| GM         | Graph Mining                                                                            |
| GPS        | Global Positioning System                                                               |
| ITS        | Intelligent Transportation System                                                       |
| k-NN       | k-Nearest Neighbor                                                                      |
| LASSO      | Least Absolute Shrinkage and Selection Operator                                         |
| LM         | Levenberg–Marquardt                                                                     |
| LSSVM      | Least Squares Support Vector Machine                                                    |
| LSTM       | Long Short-Term Memory                                                                  |
| LSTAR      | Localized Space-Time Autoregressive                                                     |
| MAE        | Mean Absolute Error                                                                     |
| MAPE       | Mean Absolute Percentage Error                                                          |
| MASTGCNNet | Multi-scale Attention-Based Spatio-Temporal Graph Convolution Recurrent Network         |
| MetaSTC    | Metapopulation Spatio-Temporal Clustering                                               |
| MLP        | Multilayer Perceptron                                                                   |
| MOT        | Ministry of Transport                                                                   |
| MSE        | Mean Square Error                                                                       |
| MSTFLN     | Multiscale Spatiotemporal Feature Learning Network                                      |
| NAR        | Nonlinear Autoregressive                                                                |
| PI         | Performance Index                                                                       |
| PSPJSTGCN  | Parallel Self-learned and Predefined Joint Spatial–Temporal Graph Convolutional Network |
| QTR        | Qassim—Riyadh (Highway code)                                                            |
| RBF        | Radial Basis Function                                                                   |
| RBANN      | Radial Basis Function Artificial Neural Network                                         |
| RF         | Random Forest                                                                           |
| RFE        | Recursive Feature Elimination                                                           |
| RNN        | Recurrent Neural Network                                                                |
| RTD        | Riyadh–Dammam (Highway code)                                                            |
| RTQ        | Riyadh—Qassim (Highway code)                                                            |
| SAR        | Saudi Riyal                                                                             |
| SARIMA     | Seasonal Autoregressive Integrated Moving Average                                       |
| SAE        | Stacked Auto-Encoder                                                                    |
| SCG        | Scaled Conjugate Gradient                                                               |
| SHAP       | Shapley Additive Explanations                                                           |
| STCNN      | Spatio-Temporal Convolutional Neural Network                                            |
| SVM        | Support Vector Machine                                                                  |
| SVR        | Support Vector Regression                                                               |
| T2F-LSTM   | Type-2 Fuzzy Long Short-Term Memory                                                     |

|      |                                       |
|------|---------------------------------------|
| UAV  | Unmanned Aerial Vehicle               |
| VAPE | Variance of Absolute Percentage Error |
| XAI  | Explainable Artificial Intelligence   |
| XGB  | Extreme Gradient Boosting             |

## References

1. Kashifi, M.T.; Al-Turki, M.; Sharify, A.W. Deep Hybrid Learning Framework for Spatiotemporal Crash Prediction Using Big Traffic Data. *Int. J. Transp. Sci. Technol.* **2023**, *12*, 793–808. [\[CrossRef\]](#)
2. Taheri, A.; Moghadasi, R.; Rasaizadi, A. The Short-Term Prediction of Traffic Parameters: A Review of Parametric and Nonparametric Approaches. *AUT J. Civ. Eng.* **2022**, *6*, 369–386.
3. Le, L.-V.; Sinh, D.; Tung, L.-P.; Lin, B.-S.P. A Practical Model for Traffic Forecasting Based on Big Data, Machine-Learning, and Network KPIs. In *15th IEEE Annual Consumer Communications & Networking Conference (CCNC)*; IEEE: Las Vegas, NV, USA, 2018; pp. 1–4.
4. Gao, Y.; Zhu, J. Characteristics, Impacts and Trends of Urban Transportation. *Encyclopedia* **2022**, *2*, 1168–1182. [\[CrossRef\]](#)
5. Sayed, S.A.; Abdel-Hamid, Y.; Hefny, H.A. Intelligent Traffic Flow Prediction Using Deep Learning Techniques: A Comparative Study. *SN Comput. Sci.* **2025**, *6*, 60. [\[CrossRef\]](#)
6. Doğan, G.; Ergen, B. A New Mobile Convolutional Neural Network-Based Approach for Pixel-Wise Road Surface Crack Detection. *Measurement* **2022**, *195*, 111119. [\[CrossRef\]](#)
7. Zhang, Y.; Zhang, Y. A Comparative Study of Three Multivariate Short-Term Freeway Flow Forecasting Methods with Missing Data. *J. Intell. Transp. Syst.* **2016**, *20*, 205–218. [\[CrossRef\]](#)
8. Cheng, Z.; Pang, M.-S.; Pavlou, P.A. Mitigating Traffic Congestion: The Role of Intelligent Transportation Systems. *Inf. Syst. Res.* **2020**, *31*, 653–674. [\[CrossRef\]](#)
9. Shepelev, V.; Slobodin, I.; Gritsenko, A.; Fadina, O. Forecasting the Amount of Traffic-Related Pollutant Emissions by Neural Networks. *Front. Built Env.* **2022**, *8*, 945615. [\[CrossRef\]](#)
10. Karlaftis, M.G.; Golias, I. Effects of Road Geometry and Traffic Volumes on Rural Roadway Accident Rates. *Accid. Anal. Prev.* **2002**, *34*, 357–365. [\[CrossRef\]](#)
11. Do, V.M.; Tran, Q.H.; Le, K.G.; Vuong, X.C.; Vu, V.T. Enhanced Deep Neural Networks for Traffic Speed Forecasting Regarding Sustainable Traffic Management Using Probe Data from Registered Transport Vehicles on Multilane Roads. *Sustainability* **2024**, *16*, 2453. [\[CrossRef\]](#)
12. Chen, Y.; Guizani, M.; Zhang, Y.; Wang, L.; Crespi, N.; Lee, G.M. When Traffic Flow Prediction Meets Wireless Big Data Analytics meet. *IEEE Network* **2017**, *33*, 161–167. [\[CrossRef\]](#)
13. Chen, J.; Li, D.; Zhang, G.; Zhang, X. Localized Space-Time Autoregressive Parameters Estimation for traffic Flow Prediction in Urban Road Networks. *Appl. Sci.* **2018**, *8*, 277. [\[CrossRef\]](#)
14. Smith, B.L. Forecasting Freeway Traffic Flow for Intelligent Transportation Application. *Transp. Res. Part. A* **1997**, *123*, 261–266.
15. Lin, L.; Wang, Q.; Sadek, A. Short-Term Forecasting of Traffic Volume: Evaluating Models Based on Multiple Data Sets and Data Diagnosis Measures. *Transp. Res. Rec. J. Transp. Res. Board.* **2013**, *2392*, 40–47. [\[CrossRef\]](#)
16. Zhao, L.; Song, Y.; Zhang, C.; Liu, Y.; Wang, P.; Lin, T.; Deng, M.; Li, H. T-GCN: A Temporal Graph Convolutional Network for Traffic Prediction. *IEEE Trans. Intell. Transp. Syst.* **2020**, *21*, 3848–3858. [\[CrossRef\]](#)
17. Alajali, W.; Zhou, W.; Wen, S. Traffic Flow Prediction for Road Intersection Safety. In *Proceedings of the 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, Guangzhou, China, 8–12 October 2018; IEEE: New York, NY, USA, 2018; pp. 812–820.
18. Hou, Z.; Li, X. Repeatability and Similarity of Freeway Traffic Flow and Long-Term under Big Data. *IEEE Trans. Intell. Transp. Syst.* **2016**, *17*, 1786–1796. [\[CrossRef\]](#)
19. Su, F.; Dong, H.; Jia, L.; Qin, Y.; Tian, Z. Long-Term Forecasting Oriented to Urban Expressway Traffic Situation. *Adv. Mech. Eng.* **2016**, *8*, 1687814016628397. [\[CrossRef\]](#)
20. Li, L.; He, S.; Zhang, J.; Ran, B. Short-Term Highway Traffic Flow Prediction Based on a Hybrid Considering Temporal-Spatial Information. *J. Adv. Transp.* **2016**, *50*, 2029–2040. [\[CrossRef\]](#)
21. Narayanan, S.; Emmanouil, C.; Constantinos, A. Factors Affecting Traffic Flow Efficiency Implications of Connected and Autonomous Vehicles: A Review and Policy Recommendations. *Adv. Transp. Policy Plan.* **2020**, *5*, 1–50.
22. Topić, J.; Škugor, B.; Deur, J. Neural Network-Based Prediction of Vehicle Fuel Consumption Based on Driving Cycle Data. *Sustainability* **2022**, *14*, 744. [\[CrossRef\]](#)
23. Rakhmanto, G.B. Short-Term Traffic Flow Prediction with Support Vector Machine to Reduce Congestion on Toll Roads in Jakarta. *Eduvest J. Univers. Stud.* **2025**, *5*, 6433–6444. [\[CrossRef\]](#)

24. Chen, M.; Chen, R.; Cai, F.; Li, W.; Guo, N.; Li, G. Short-Term Traffic Flow Prediction with Recurrent Mixture Density Network. *Math. Probl. Eng.* **2021**, 2021, 1–9. [\[CrossRef\]](#)
25. Qu, L.; Ma, D.; Wang, Y. Daily Long-Term Traffic Flow Forecasting Based on a Deep Neural. *Expert. Syst. Appl.* **2019**, 121, 304–312. [\[CrossRef\]](#)
26. Brownlee, J. *Master Machine Learning Algorithms: Discover How They Work and implement Them from Scratch*. Machine Learning Mastery; Guiding Tech Media: San Juan, Puerto Rico, 2016.
27. Lee, S.; Fambro, D.B. Application of Subset Autoregressive Integrated Moving Average for Short-Term Freeway Traffic Volume Forecasting. *Transp. Res. Rec.* **1999**, 1678, 179–188. [\[CrossRef\]](#)
28. Ahmed, A.; Gazder, U. Traffic Forecasting for King Fahd Causeway: Comparison of Parametric Technique with Artificial Neural Networks. *Int. J. Adv. Civ. Struct. Environ. Eng. IJACSE* **2013**, 1, 75–79.
29. Siddiquee, M.S.A.; Hoque, S. Predicting the Daily Traffic Volume from Hourly Traffic Data Using Neural Network. *Neural Netw. World* **2017**, 27, 283–294. [\[CrossRef\]](#)
30. Yin, K.; Wu, J.; Wang, W.; Lee, D.-H.; Wei, Y. An Integrated Resilience Assessment Model of Urban Transportation Network: A Case Study of 40 Cities in China. *Transp. Res. Part. A Policy Pr.* **2023**, 173, 103687. [\[CrossRef\]](#)
31. Sporkmann, J.; Liu, Y.; Spinler, S. Carbon Emissions from European Land Transportation: A Comprehensive Analysis. *Transp. Res. D Transp. Env.* **2023**, 121, 103851. [\[CrossRef\]](#)
32. Momin, K.A.; Barua, S.; Jamil, M.S.; Hamim, O.F. Short Duration Traffic Flow Prediction Using Kalman Filtering. In Proceedings of the 6th International Conference on Civil Engineering for Sustainable Development (Iccesd 2022), Khulna, Bangladesh, 10–12 February 2022; AIP Publishing LLC: Melville, NY, USA, 2023; p. 040011.
33. Zheng, Z.; Su, D. Short-Term Traffic Volume Forecasting: A k-Nearest Neighbor Enhanced by Constrained Linearly Sewing Principle Algorithm. *Transp. Res. Part. C Emerg. Technol.* **2014**, 43, 143–157. [\[CrossRef\]](#)
34. Wang, J.; Boukerche, A. Non-Parametric Models with Optimized Training Strategy for Vehicles Traffic Flow Prediction. *Comput. Netw.* **2021**, 187, 107791. [\[CrossRef\]](#)
35. Mladenović, D.; Janković, S.; Zdravković, S.; Mladenović, S.; Uzelac, A. Night Traffic Flow Prediction Using K-Nearest Neighbors Algorithm. *Oper. Res. Eng. Sci. Theory Appl.* **2022**, 5, 152–168. [\[CrossRef\]](#)
36. Alam, M.S.; Deb, J.B.; Al Amin, A.; Chowdhury, S. An Artificial Neural Network for Predicting Air Traffic Demand Based on Socio-Economic Parameters. *Decis. Anal. J.* **2024**, 10, 100382. [\[CrossRef\]](#)
37. Chen, J. *Characterization and Implementation of Neural Network Time Series Models for Traffic Volume Forecasting*; The University of Toledo: Toledo, OH, USA, 1997.
38. Zhang, Q.; Lu, J.; Kuang, W.; Wu, L.; Wang, Z. Short-Term Charging Load Prediction of Electric Vehicles with Dynamic Traffic Information Based on a Support Vector Machine. *World Electr. Veh. J.* **2024**, 15, 189. [\[CrossRef\]](#)
39. Kumar, S. V Traffic Flow Prediction Using Kalman Filtering Technique. *Procedia Eng.* **2017**, 187, 582–587. [\[CrossRef\]](#)
40. Wang, C.; Huang, S.; Zhang, C. Short-Term Traffic Flow Prediction Considering Weather Factors Based on Optimized Deep Learning Neural Networks: Bo-GRA-CNN-BiLSTM. *Sustainability* **2025**, 17, 2576. [\[CrossRef\]](#)
41. Olayode, I.O.; Tartibu, L.K.; Okwu, M.O.; Severino, A. Comparative Traffic Flow Prediction of a Heuristic ANN Model and a Hybrid ANN-PSO Model in the Traffic Flow Modelling of Vehicles at a Four-Way Signalized Road Intersection. *Sustainability* **2021**, 13, 10704. [\[CrossRef\]](#)
42. Olayode, I.O.; Tartibu, L.K.; Okwu, M.O. Traffic Flow Prediction at Signalized Road Intersections: A Case of Markov Chain and Artificial Neural Network Model. In Proceedings of the 2021 IEEE 12th International Conference on Mechanical and Intelligent Manufacturing Technologies (ICMIMT), Cape Town, South Africa, 13–15 May 2021; IEEE: New York, NY, USA, 2021; pp. 287–292.
43. Xu, K.; Yu, Z.; Gao, Y.; Zhang, S.; Fang, J.; Gao, X.; Chen, G. MetaSTC: A Backbone Agnostic Spatio-Temporal Framework for Traffic Forecasting. In Proceedings of the 2024 IEEE International Conference on Data Mining (ICDM), Abu Dhabi, United Arab Emirates, 9–12 December 2024; IEEE: New York, NY, USA, 2024; pp. 899–904.
44. Yang, X.; Li, Q.; Xu, P.; He, D.; Tan, H. Parallel Self-Learned and Predefined Joint Spatial–Temporal Graph Convolutional Networks for Traffic Flow Prediction. *IEEE Internet Things J.* **2025**, 12, 11698–11707. [\[CrossRef\]](#)
45. Xiao, Z.; Shen, Q.; Li, C.; Li, D.; Liu, Q. An Adaptive Spatiotemporal Dynamic Graph Convolutional Network for Traffic Prediction. *Sci. Rep.* **2025**, 15, 27098. [\[CrossRef\]](#)
46. Zhou, B.; Liu, J.; Cui, S.; Zhao, Y. A Large-Scale Spatio-Temporal Multimodal Fusion Framework for Traffic Prediction. *Big Data Min. Anal.* **2024**, 7, 621–636. [\[CrossRef\]](#)
47. Lee, E.H. Traffic Speed Prediction of Urban Road Network Based on High Importance Links Using XGB and SHAP. *IEEE Access* **2023**, 11, 113217–113226. [\[CrossRef\]](#)
48. Min, J.H.; Ham, S.W.; Kim, D.-K.; Lee, E.H. Deep Multimodal Learning for Traffic Speed Estimation Combining Dedicated Short-Range Communication and Vehicle Detection System Data. *Transp. Res. Rec. J. Transp. Res. Board.* **2023**, 2677, 247–259. [\[CrossRef\]](#)
49. Shen, Z. Long-Time Traffic Speed Prediction Model Based on XGBoost 2023. *preprint* **2023**, 1–11. [\[CrossRef\]](#)

50. Nesa, M.; Yoon, Y. Speed Prediction and Nearby Road Impact Analysis Using Machine Learning and Ensemble of Explainable AI Techniques. *Sci. Rep.* **2024**, *14*, 25208. [[CrossRef](#)] [[PubMed](#)]
51. Wiecek, G.; un din Tahir, S.B.; Akhter, I.; Kurek, J. Vehicle Detection and Recognition Approach in Multi-Scale Traffic Monitoring System via Graph-Based Data Optimization. *Sensors* **2023**, *23*, 1731. [[CrossRef](#)] [[PubMed](#)]
52. Satyanarayana, G.S.R.; Deshmukh, P.; Das, S.K. Vehicle Detection and Classification with Spatio-Temporal Information Obtained from CNN. *Displays* **2022**, *75*, 102294. [[CrossRef](#)]
53. Singh, A.; Rahma, M.Z.U.; Rani, P.; Agrawal, N.K.; Sharma, R.; Kariri, E.; Aray, D.G. Smart Traffic Monitoring Through Real-Time Moving Vehicle Detection Using Deep Learning via Aerial Images for Consumer Application. *IEEE Trans. Consum. Electron.* **2024**, *70*, 7302–7309. [[CrossRef](#)]
54. Bakirci, M. Enhancing Vehicle Detection in Intelligent Transportation Systems via Autonomous UAV Platform and YOLOv8 Integration. *Appl. Soft Comput.* **2024**, *164*, 112015. [[CrossRef](#)]
55. Guo, F.; Polak, J.W.; Krishnan, R. Comparison of Modelling Approaches for Short Term Traffic Prediction under Normal and Abnormal Conditions. In Proceedings of the 13th International IEEE Conference on Intelligent Transportation Systems, Funchal, Portugal, 19–22 September 2010; IEEE: New York, NY, USA, 2010; pp. 1209–1214.
56. Hsueh, Y.-L.; Yang, Y.-R. A Short-Term Traffic Speed Prediction Model Based on LSTM Networks. *Int. J. Intell. Transp. Syst. Res.* **2021**, *19*, 510–524. [[CrossRef](#)]
57. Wisconsin Department of Transportation. *Transportation Planning Manual*; Wisconsin Department of Transportation: Madison, WI, USA, 2008.
58. Lv, Y.; Duan, Y.; Kang, W.; Li, Z.; Wang, F.-Y. Traffic Flow Prediction with Big Data: A Deep Learning Approach. *IEEE Trans. Intell. Transp. Syst.* **2014**, *16*, 865–873. [[CrossRef](#)]
59. Al-Masaeid, H.R.; Al-Omoush, N.J. Traffic Volume Forecasting for Rural Roads in Jordan. *Jordan J. Civ. Eng.* **2014**, *159*, 1–13.
60. Hong, W.-C.; Dong, Y.; Zheng, F.; Lai, C.-Y. Forecasting Urban Traffic Flow by SVR with Continuous ACO. *Appl. Math. Model.* **2011**, *35*, 1282–1291. [[CrossRef](#)]
61. Altikat, S. Prediction of CO<sub>2</sub> Emission from Greenhouse to Atmosphere with Artificial Neural Networks and Deep Learning Neural Networks. *Int. J. Environ. Sci. Technol.* **2021**, *18*, 3169–3178. [[CrossRef](#)]
62. Xu, Z.; Yang, Q.; Chen, D. Nonparametric Method for Optimizing the Start Time of Short-Term Road Maintenance Operations Using Real-Time Traffic Map Data. *Adv. Transp. Stud.* **2023**, *59*, 51–70.
63. Mor, A.; Kumar, M. Multivariate Short-Term Traffic Flow Prediction Based on Real-Time Expressway Toll Plaza Data Using Non-Parametric Techniques. *Int. J. Veh. Inf. Commun. Syst.* **2022**, *7*, 32. [[CrossRef](#)]
64. Alrumaidhi, M.; Farag, M.M.G.; Rakha, H.A. Comparative Analysis of Parametric and Non-Parametric Data-Driven Models to Predict Road Crash Severity among Elderly Drivers Using Synthetic Resampling Techniques. *Sustainability* **2023**, *15*, 9878. [[CrossRef](#)]
65. Park, B. Hybrid Neuro-Fuzzy Application in Short-Term Freeway Traffic Forecasting. *Transp. Res. Rec.* **2002**, *1802*, 190–196. [[CrossRef](#)]
66. Kumar, K.; Parida, M.; Katiyar, V.K. Short Term Traffic Flow Prediction for a Non Urban Highway Using Neural Network. *Procedia Soc. Behav. Sci.* **2013**, *104*, 755–764. [[CrossRef](#)]
67. Alsehaime, B.; Alzamzami, O.; Alowidi, N.; Ali, M. An Adaptive Spatio-Temporal Traffic Flow Prediction Using Self-Attention and Multi-Graph Networks. *Sensors* **2025**, *25*, 282. [[CrossRef](#)]
68. Yin, K.; Niu, X.; Zhu, J.; Jiang, Y.; Min, F. DAT-STAN: Dual-Module Adaptive Transformer and Spatio-Temporal Attention Network for Large-Scale Traffic Flow Prediction. *Tsinghua Sci. Technol.* **2025**. [[CrossRef](#)]
69. Ali, A.; Ullah, I.; Singh, S.K.; Sharafian, A.; Jiang, W.; Sherazi, H.I.; Bai, X. Energy-Efficient Resource Allocation for Urban Traffic Flow Prediction in Edge-Cloud Computing. *Int. J. Intell. Syst.* **2025**, *2025*, 1863025. [[CrossRef](#)]
70. Yasdi, R. Prediction of Road Traffic Using a Neural Network Approach. *Neural Comput. Appl.* **1999**, *8*, 135–142. [[CrossRef](#)]
71. Stutz, C.; Runkler, T.A. Classification and Prediction of Road Traffic Using-Specific Fuzzy Clustering. *IEEE Trans. Fuzzy Syst.* **2002**, *10*, 297–308. [[CrossRef](#)]
72. Yin, H.; Wong, S.; Xu, J.; Wong, C. Urban Traffic Flow Prediction Using a Fuzzy-Neural Approach. *Transp. Res. Part. C Emerg. Technol.* **2002**, *10*, 85–98. [[CrossRef](#)]
73. Ratrou, N.T.; Gazder, U. Factors Affecting Performance of Parametric and Non-Parametric for Daily Traffic Forecasting. *Procedia Comput. Sci.* **2014**, *32*, 285–292. [[CrossRef](#)]
74. Khalifa, A.; Idsougou, Y.; Benabbou, L.; Zirari, M. Case Study of the Moroccan Highway Network. *arXiv* **2017**, arXiv:1711.06779. [[CrossRef](#)]

75. Zang, D.; Fang, Y.; Wang, D.; Wei, Z.; Tang, K.; Li, X. Long Term Traffic Flow Prediction Using Residual Net and deconvolutional Neural Network. In *Chinese Conference on Pattern Recognition and Computer Vision (PRCV)*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 62–74.
76. Zang, D.; Ling, J.; Wei, Z.; Tang, K.; Cheng, J. Long-Term Traffic Speed Prediction Based on Multiscale-Temporal Feature Learning Network. *IEEE Trans. Intell. Transp. Syst.* **2018**, *20*, 3700–3709. [\[CrossRef\]](#)
77. Wu, Y.; Tan, H.; Qin, L.; Ran, B.; Jiang, Z. A Hybrid Deep Learning Based Traffic Flow Prediction Method and its Understanding. *Transp. Res. Part. C Emerg. Technol.* **2018**, *90*, 166–180. [\[CrossRef\]](#)
78. Zhao, L.; Zhou, Y.; Lu, H.; Fujita, H. Parallel Computing Method of Deep Belief Networks and Its to Traffic Flow Prediction. *Knowl. Based Syst.* **2019**, *163*, 972–987. [\[CrossRef\]](#)
79. Li, R.; Huang, Y.; Wang, J. Long-Term Traffic Volume Prediction Based on K-Means Gaussian Type-2 Fuzzy Sets. *IEEE/CAA J. Autom. Sin.* **2019**, *6*, 1344–1351. [\[CrossRef\]](#)
80. Tang, J.; Chen, X.; Hu, Z.; Zong, F.; Han, C.; Li, L. Traffic Flow Prediction Based on Combination of Support Vector and Data Denoising Schemes. *Phys. A Stat. Mech. Its Appl.* **2019**, *534*, 120642. [\[CrossRef\]](#)
81. He, Z.; Chow, C.-Y.; Zhang, J.-D. STCNN: A Spatio-Temporal Convolutional Neural Network for long-Term Traffic Prediction. In *20th IEEE International Conference on Mobile Data Management (MDM)*; IEEE: Hong Kong, China, 2019; pp. 226–233.
82. Li, R.; Hu, Y.; Liang, Q. T2F-LSTM Method for Long-Term Traffic Volume Prediction. *IEEE Trans. Fuzzy Syst.* **2020**, *28*, 3256–3264. [\[CrossRef\]](#)
83. Li, Y.; Chai, S.; Ma, Z.; Wang, G. A Hybrid Deep Learning Framework for Long-Term Traffic Flow Prediction. *IEEE Access* **2021**, *9*, 11264–11271. [\[CrossRef\]](#)
84. Park, S.; Kim, M.; Kim, J. Hourly Long-Term Traffic Volume Prediction with Meteorological Information Using Graph Convolutional Networks. *Appl. Sci.* **2024**, *14*, 2285. [\[CrossRef\]](#)
85. Toba, A.-L.; Kulkarni, S.; Khallouli, W.; Pennington, T. Long-Term Traffic Prediction Using Deep Learning Long Short-Term Memory. *Smart Cities* **2025**, *8*, 126. [\[CrossRef\]](#)
86. Smeeton, N.; Neil, H.S.; Peter, S. *Applied Nonparametric Statistical Methods*, 5th ed.; CRC Press: Boca Raton, FL, USA, 2025.
87. Sayed, S.A.; Abdel-Hamid, Y.; Hefny, H.A. Artificial Intelligence-Based Traffic Flow Prediction: A Comprehensive Review. *J. Electr. Syst. Inf. Technol.* **2023**, *10*, 13. [\[CrossRef\]](#)
88. Duraku, R.K.; Ramadan, R. Development of Traffic Volume Forecasting Using Multiple Regression Analysis and Artificial Neural Network. *Civ. Eng. J.* **2019**, *5*, 1698–1713. [\[CrossRef\]](#)
89. Ariyo, A.A.; Adewumi, A.O.; Ayo, C.K. Stock Price Prediction Using the ARIMA Model. In *UKSim-AMSS 16th International Conference on Computer Modelling and Simulation*; IEEE: Cambridge, UK, 2014; pp. 106–112.
90. Tantekin, A. Heating Degree Days (HDD) Forecasting and Heating Degree Days Factor (HDDF) Assessment in Selected European Countries. *Energy* **2025**, *332*, 137277. [\[CrossRef\]](#)
91. Farinango, P.; Estrada, R.; Vera, N. Evaluation of Indoor Positioning System Design Based on Empirical Propagation Models versus Optimizable Regression Models. *Procedia Comput. Sci.* **2025**, *265*, 301–309. [\[CrossRef\]](#)
92. Møller, M.F. A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning. *Neural Netw.* **1993**, *6*, 525–533. [\[CrossRef\]](#)
93. Baghirli, O. Comparison of Lavenberg-Marquardt, Scaled Conjugate Gradient and Bayesian Regularization Backpropagation Algorithms for Multistep Wind Speed Forecasting Using Multilayer Perceptron Neural Network. Master's Thesis, Uppsala University, Disciplinary Domain of Science and Technology, Earth Sciences, Department of Earth Sciences, Uppsala, Sweden, 2015.
94. Mohamad, N.; Zaini, F.; Johari, A.; Yassin, I.; Zabidi, A. Comparison between Levenberg-Marquardt and Scaled Conjugate Training Algorithms for Breast Cancer Diagnosis Using. In *6th International Colloquium on Signal Processing & Its Applications*; IEEE: Malacca City, Malaysia, 2010; pp. 1–7.
95. Fischer, A.; Izmailov, A.F.; Solodov, M.V. The Levenberg–Marquardt Method: An Overview of Modern Convergence Theories and More. *Comput. Optim. Appl.* **2024**, *89*, 33–67. [\[CrossRef\]](#)
96. Demuth, H.; Beale, M.; Hagan, M. *MATLAB User's Guide, Version 4.0: Neural Network Toolbox*; MathWorks Inc: Natick, MA, USA, 2005.
97. Fiorentini, N.; Pellegrini, D.; Losa, M. Overfitting Prevention in Accident Prediction Models: Bayesian Regularization of Artificial Neural Networks. *Transp. Res. Rec. J. Transp. Res. Board.* **2023**, *2677*, 1455–1470. [\[CrossRef\]](#)
98. Wang, K.W.; Deng, C.; Li, J.P.; Zhang, Y.Y.; Li, X.Y.; Wu, M.C. Hybrid Methodology for Tuberculosis Incidence Time-Series Based on ARIMA and a NAR Neural Network. *Epidemiol. Infect.* **2017**, *145*, 1118–1129. [\[CrossRef\]](#) [\[PubMed\]](#)
99. Kim, S.; Kim, H. A New Metric of Absolute Percentage Error for Intermittent Demand Forecasts. *Int. J. Forecast.* **2016**, *32*, 669–679. [\[CrossRef\]](#)
100. Khoshvaght, H.; Permala, R.R.; Razmjou, A.; Khiadani, M. A Critical Review on Selecting Performance Evaluation Metrics for Supervised Machine Learning Models in Wastewater Quality Prediction. *J. Env. Chem. Eng.* **2025**, *13*, 119675. [\[CrossRef\]](#)

101. Alali, Y.; Harrou, F.; Sun, Y. Unlocking the Potential of Wastewater Treatment: Machine Learning Based Energy Consumption Prediction. *Water* **2023**, *15*, 2349. [[CrossRef](#)]
102. Khan, F.M.; Gupta, R. Arima and Nar Based Prediction Model for Time Series Analysis Of-19 Cases in India. *J. Saf. Sci. Resil.* **2020**, *1*, 12–18. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.