applied sciences

MDPI

Article

Comparative Study Analysis of ANFIS and ANFIS-GA Models on Flow of Vehicles at Road Intersections

Isaac Oyeyemi Olayode [1,2,*], Lagouge Kwanda Tartibu [1] and Frimpong Justice Alex [3]

1 Department of Mechanical and Industrial Engineering, University of Johannesburg, Johannesburg P.O. Box 2028, South Africa
2 SMART Infrastructure Facility, University of Wollongong, Wollongong, NSW 2522, Australia
3 School of Automotive Engineering, Wuhan University of Technology, Wuhan 430070, China
* Correspondence: olayode89@gmail.com or iolayode@uow.edu.au; Tel.: +27-710200947

Abstract: In the last two decades the efficient traffic-flow prediction of vehicles has been significant in curbing traffic congestions at freeways and road intersections and it is among the many advantages of applying intelligent transportation systems in road intersections. However, transportation researchers have not focused on prediction of vehicular traffic flow at road intersections using hybrid algorithms such as adaptive neuro-fuzzy inference systems optimized by genetic algorithms. In this research, we propose two models, namely the adaptive neuro-fuzzy inference system (ANFIS) and the adaptive neuro-fuzzy inference system optimized by genetic algorithm (ANFIS-GA), to model and predict vehicles at signalized road intersections using the South African public road transportation system. The traffic data used for this research were obtained via up-to-date traffic data equipment. Eight hundred fifty traffic datasets were used for the ANFIS and ANFIS-GA modelling. The traffic data comprised traffic volume (output), speed of vehicles, and time (inputs). We used 70% of the traffic data for training and 30% for testing. The ANFIS and ANFIS-GA results showed training performance of (R^2) 0.9709 and 0.8979 and testing performance of (R^2) 0.9790 and 0.9980. The results show that ANFIS-GA is more appropriate for modelling and prediction of traffic flow of vehicles at signalized road intersections. This research adds further to our knowledge of the application of hybrid genetic algorithms in traffic-flow prediction of vehicles at signalized road intersections.

Keywords: traffic flow; road intersections; ANFIS; ANFIS-GA; signalized road intersection; machine learning; fuzzy networks

Citation: Olayode, I.O.; Tartibu, L.K.; Alex, F.J. Comparative Study Analysis of ANFIS and ANFIS-GA Models on Flow of Vehicles at Road Intersections. *Appl. Sci.* **2023**, *13*, 744. https://doi.org/10.3390/app13020744

Academic Editor: Luis Picado Santos

Received: 24 November 2022
Revised: 28 December 2022
Accepted: 30 December 2022
Published: 5 January 2023

1. Introduction

The rapid increase in the growth of vehicle ownership has bought significant strain to the traffic flow of vehicles in developed and developing countries, which has led to tremendous changes in the daily lives of pedestrians and motorists [1–4]. However, it is important to determine the solution to traffic congestion through efficient traffic management and the application of intelligent transportation systems (ITS), especially when referring to a short-term forecast traffic flow (i.e., 60 min) [5–8]. Traffic-flow forecasting helps traffic signal control, volume management, and travel-route planning. In real-life traffic-flow situations, traffic data are significant. Vehicular traffic flow is often defined by non-linear and highly intricate characteristics, making traffic-flow prediction difficult to determine in transportation engineering.

The ability to predict traffic congestion is critical to the successful deployment or application of smart transportation subsystems which include sophisticated traveller information, improved traffic flow management systems, and state-of-the-art public transportation systems. Therefore, free-flowing traffic is crucial for the development of faster transportation and connectivity systems. The majority of techniques used for predicting traffic flow are heavily dependent on machine-learning models which require minor manual interference and consume a lot of time depending on the machine-learning model and the traffic data obtained [9,10].

In more than 20 years of traffic-flow-prediction research, most studies are either classified as model or data-driven, or they can be both. Model-driven methods are also known as parametric techniques, e.g., time series models. However, using this type of parametric technique needs many parameters and assumptions to be used in the transportation network, which makes the performance of the traffic-flow prediction not successful. In recent years, with the technological development of transportation infrastructures, various types of traffic-data-collection methods, notably traffic-flow monitoring equipment and loop detectors, have offered an enormous amount of traffic data for predicting traffic flow. Traffic-data methods, such as artificial neural networks, recurrent neural networks, fuzzy inference systems, genetic algorithms, and particle swarm optimization can be categorized into machine- or deep-learning depending on the traffic flow of vehicles [9–15]. Well-known machine-learning techniques are inadequate when evaluating high-dimensional traffic data and when depending on comprehensive engineering.

According to what was stated by [4–6], by 2050, there will be more than 7 billion people staying in metropolitan areas, which will comprise between two-fourths and two-thirds of the human population. The concept of smart cities is one of the most well-known researched ideologies for future developed urban cities. Smart cities use the integration of urban resources via the application of artificial-intelligence technology associated with big data, cloud computing, deep neural networks, and the Internet of Things (IoT) [16–19]. The smart city framework can assist in city governance, effective regulations, maintenance of the cities' transportation infrastructures, and reduction of air and noise pollution [20–22].

Furthermore, the intelligent-city framework can offer people a more liveable and intelligent-driven living environment [23]. An intelligent transportation system (ITS) is an integral aspect of an innovative city framework. An ITS has an effective integration of both on-road and off-road resources by applying technologically innovative information and communication technology. ITSs are used in road transportation systems to provide real-time traffic-flow information, effective freeway traffic control, and an efficient vehicular cloud framework [24]. With the assistance of intelligent transportation systems, safety, effective mobility, and sustainable transportation, conducive road transportation systems can be achieved. For ITS to be achieved and implemented, an effective and efficient prediction of vehicular traffic flow is mandatory [25–29].

The usage of an accurate vehicular traffic-flow prediction system in ITS is to offer regular continuous and accurate road-travel information depending on the conditions of the road, such as the impacts of flow of vehicles on road intersections and freeways, which are vital to road traffic control and management and integration of road resources. The significant difficulties in applying traffic-flow prediction in intelligent transportation systems can be divided into accuracy, efficiency, and prediction [30–35].

The foundation of efficiency issues for traffic-flow prediction is the primary characteristic of vehicular traffic-flow patterns. The patterns of vehicular traffic flow are usually rigid and not flexible. The variations in their patterns are impacted by traffic lights from the traffic flow of vehicles, unstable changes in weather, and other impeding factors. Some of these have a long-term impact, causing the traffic-flow variations to exhibit specific trends and irregularities, while there are also short-term impacts. The times series variations can be categorized into (1) *trend variation*, which can be defined as the process that allows change to occur while using time and the gradual rise, reduction, or no rise in a specific direction; (2) *periodic change,* which is also known as seasonal change and is a process that allows cyclical fluctuations at a specific period; (3) *cyclical variation*, which is defined as the processes that are unfixed and subject to periodic fluctuations; and (4) *random variation* which is defined as when accidental constraints impact the processes and show irregular variations. The definition of time series is usually a combination of all these variations. This is due to the impact of natural and human interference on road traffic. It possesses intricate non-linear features, which causes enormous difficulties regarding the efficiency of traffic-flow prediction.

The efficiency of the model used in traffic-flow prediction of vehicles depends on the implementation cost and the prediction results. The implementation cost is primarily caused by the model training and application of this model in an extensive urban road transportation system [36]. An enormous amount of time is used for predictive model training, specifically for the model associated with machine learning, such as a machine-learning-dependent model combined with the structure of deep learning. However, when applying the predictive model to a large urban road network, the distinctiveness of the characteristics of various types of road networks can be understood by using predictive models [37,38]. These predictive models comprise different types of parameters which are associated with each segment of the road.

Traffic-flow prediction has become a study hotspot in many road-transportation systems, especially in road intersections, freeway congestions, and un-signalized road intersections. Even though ANFIS combines the learning power of artificial neural networks with fuzzy logic knowledge representation, some complexities exist when creating membership functions in ANFIS. The foundation of these complexities can be found when tuning the ANFIS function to create the optimal model with high efficiency and optimal performance features. The primary aim of combining genetic algorithm (GA) with ANFIS is to decrease errors by tuning and optimizing processes on the ANFIS membership functions. The contribution of this study to the field of transportation engineering especially sub-fields such as road intersections and traffic flow modelling are threefold:

- This study extends our knowledge of the prediction of traffic flow at signalized road intersections using traffic volume, speed of vehicles, and time as our inputs and output in modelling the traffic flow.
- This is the first study to undertake a comparative analysis of using a genetic algorithm combined with an adaptive neuro-fuzzy inference system to model the traffic flow of vehicles at signalized road intersections.
- This study contributes to the growing area of using metaheuristics algorithms in traffic flow prediction of non-autonomous vehicles at signalized road intersections.

This paper is organized as follows: The Section 2 explains in detail the development of the ANFIS and ANFIS-GA models, the location of the study, and the traffic data analysis. The Section 3 is concerned with the research results and discussions, and the Section 4 comprehensively explains the conclusion, recommendations for future work, and study limitations.

2. Materials and Methods

A flowchart of the methodologies used in this study is shown in Figure 1. This study was populated by traffic data from a developing country in conjunction with a traffic-data company known for its prowess in traffic-congestion-monitoring solutions. The size of the dataset used for this research was limited to 850 traffic datasets. The traffic dataset was obtained before COVID-19 lockdowns and collected for more than ten days, considering the traffic volume of vehicles from each road intersection. In this research, we applied the primary and secondary data collection techniques. The primary technique involved collecting traffic datasets from the signalized road intersections by applying station-wide installed GPS-controlled devices at these road intersections. The secondary data involved voluntary visitation to the South Africa Ministry of Transportation and interviewing transportation and traffic engineers, not excluding the technical staff, on obtaining useful traffic flow information from these road intersections. The method of traffic data collection is illustrated in Figure 2.

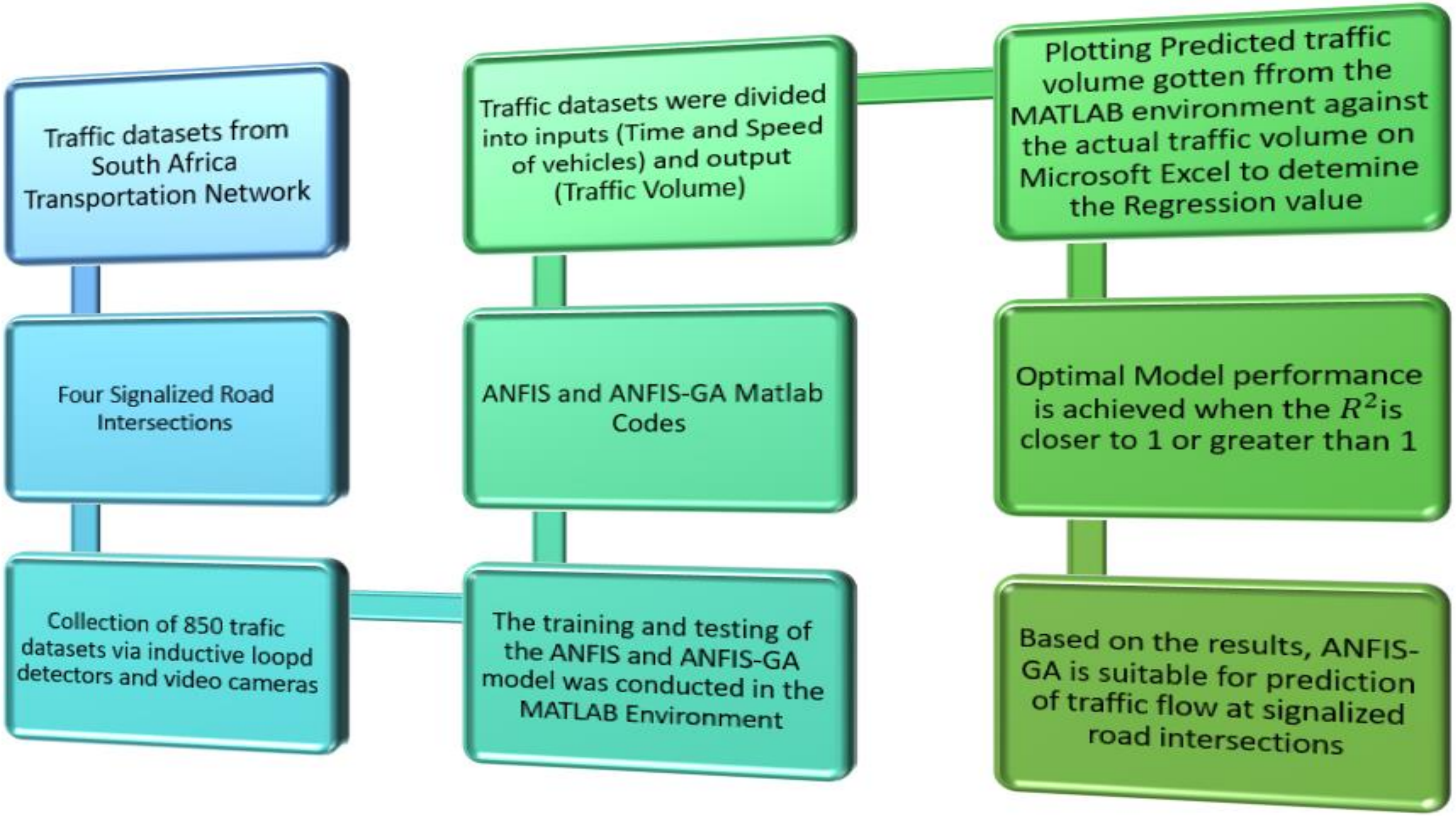

Figure 1. Methodological approaches of the study.

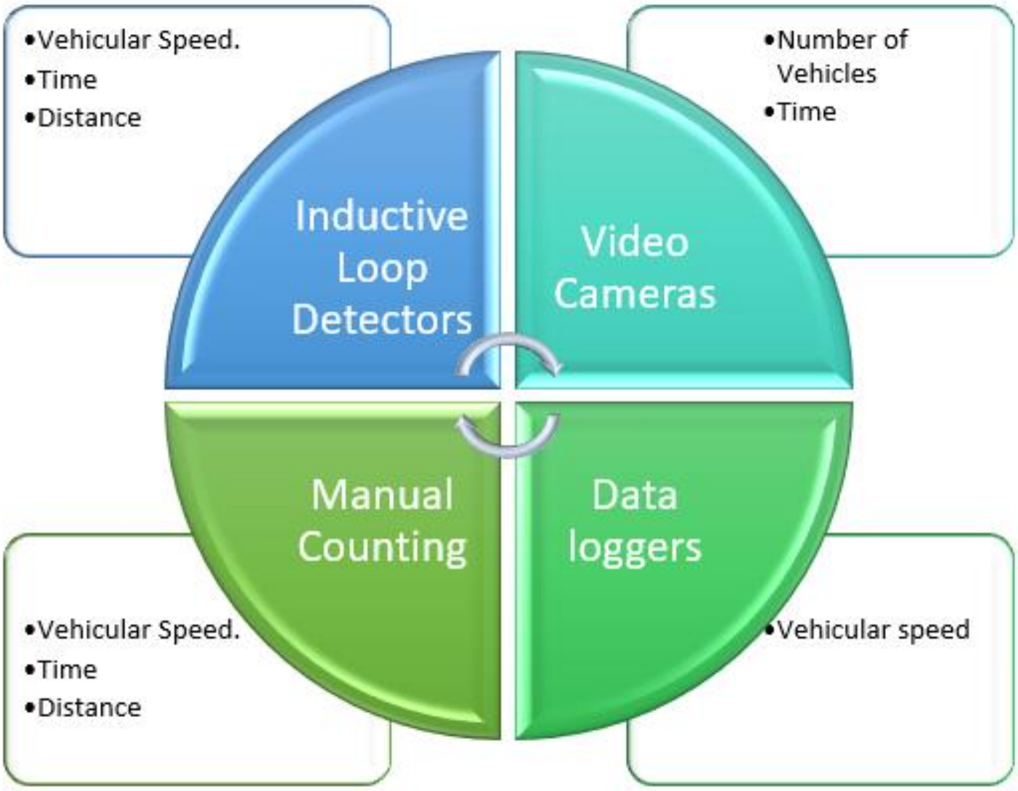

Figure 2. Collection of traffic datasets.

2.1. Models Employed

In this study, we used models such as creative algorithms (artificial neural network and fuzzy logic); hybrid models (adaptive neuro-fuzzy inference system trained by genetic algorithm (ANFIS-GA).

2.2. Study Location

The research study focused on the South African transportation system due to the high traffic volume and density of vehicles in Southern Africa. The traffic data were obtained from signalized road intersections between Johannesburg and Pretoria in Gauteng province. The chosen road intersections used for this study are known to have the highest traffic volume and traffic density of vehicles in South Africa. Based on the information obtained from the Ministry of Transportation, these road intersections sometimes experience more than half a million vehicles depending on the periods of the day [6].

The road intersections consist of several lanes (Figure 3a–d), with the maximum lanes being four and the minimum being two. They are all moving in a northbound direction depending on the vehicle's direction. The location and features of these road intersections are explained in Table 1:

Table 1. Features of the Signalized Road Intersections.

Road Intersections	Date	Distance (m)	Direction	Number of Lanes	Speed Limit (Km/h)	Number of Vehicles
Road Intersection 1	15 July 2019–27 July 2019	13.5	Northbound	4	120	12,067,153
Road Intersection 2	15 July 2019–29 July 2019	10.40	Northbound	3	120	18,340,250
Road Intersection 3	15 July 2019–29 July 2019	8.0	Northbound	4	120	11,444,024
Road Intersection 4	15 July 2019–29 July 2019	4.60	Northbound	2	120	16,151,125

Figure 3. Location of traffic data collection (**a**–**d**).

Traffic flow data were collected for fifteen days between 15 July 2019 and 29 July 2019 using installed video cameras at road intersections and video cameras and roadside detectors. Figure 4 illustrates the traffic-data input and output used for the ANFIS and ANFIS-GA models.

2.3. Model Development

2.3.1. Adaptive Neuro-Fuzzy Inference System (ANFIS)

ANFIS network architecture comprises two main parts, namely premise and consequence. ANFIS training can be defined as evaluating parameters that belong to various parts using different types of optimization algorithms. ANFIS uses the pre-existing data pairs, also known as the input–output pairs, during the model's training. In 1993, [39] devised the fuzzy rules called IF-THEN. These fuzzy rules explain the connectivity of the parts by explaining the structure of the five ANFIS layers. These layers are illustrated in Figure 5, below. Based on the primary aim of this research, the ANFIS structure comprised two inputs and one output, as shown in Figure 5. This ANFIS structure was made up of four membership functions and rules. The different layers are explained based on the ANFIS research done by [20,40].

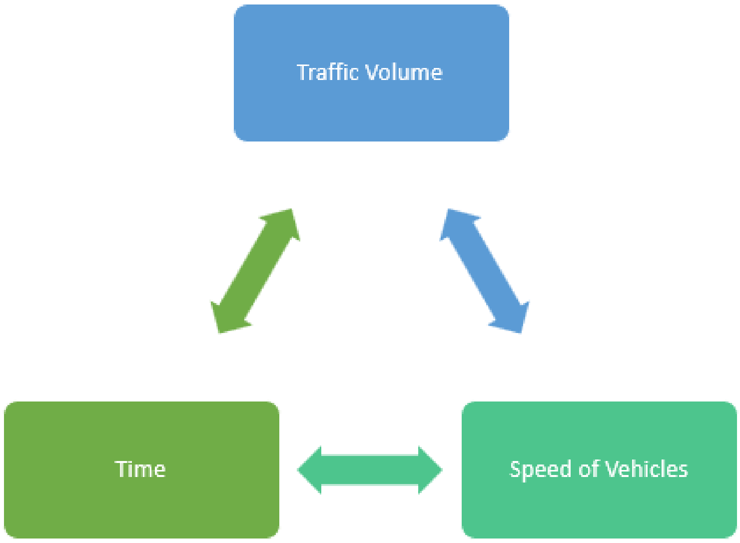

Figure 4. Traffic flow parameters.

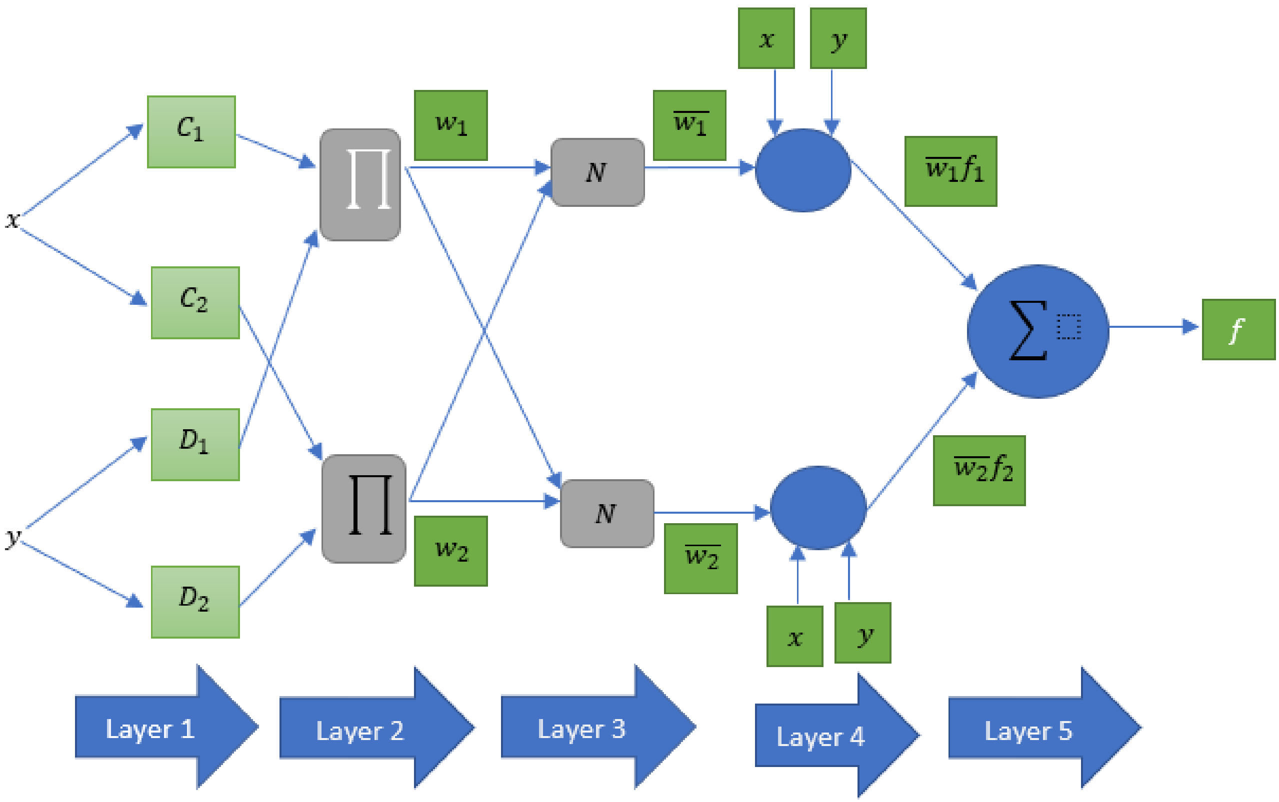

Figure 5. A typical structure of an ANFIS model.

Layer 1

This first layer is identified as the layer of fuzzification and uses membership functions to determine and differentiate fuzzy clusters from data inputs. Variables are usually used to evaluate the type of membership functions, and these variables are also known as premise variables. An example of premise variables is {a, b, c}. The degree of membership of each type of membership function is evaluated by applying these variables as explained in the equations below. The different degrees of membership evaluated using this layer are shown in the equation below.

$$\mu A_{i(x)} = gbellmf(x; a, b, c) = \frac{1}{1+\left|\frac{x-c}{a}\right|^{2b}}$$
$$O_i^1 = \mu A_i(x)$$

Layer 2

In this layer, the firing strengths w_i for layer two, rules are created by applying membership values that can be found in the fuzzification layer. The w_i values can be found by using the multiplication method on the membership values as in the equation below.

$$O_i^2 = w_i = \mu A_i(x).\mu B_i(y)$$
$$i = 1, 2$$

Layer 3

This layer is also known as normalization. It evaluates the strengths of the normalized firing belonging for different rules. The normalized variable is directly proportional to the firing strength of the *i*th rule compared to the overall firing strengths as given in the equation below.

$$O_i^3 = \overline{w_i} = \frac{w_i}{w_1 + w_2 + w_3 + w_4} i \,\epsilon\, \{1, 2, 3, 4\}$$

Layer 4

This layer is different from the first three layers because it is known as the defuzzification layer. The equation used to determine the weighted variables of each of the rules in the nodes is evaluated by applying a mathematical expression called the 1st order polynomial.

$$O_i^4 = \overline{w_i} f_i = \overline{w_i}(p_i x + q_i y + r_i)$$

The w_i is called the normalization layer output; the set of variables used are p_i, q_i, r_i. These variables are also known as consequence variables. The overall number of consequence variables attached to each rule is greater than the number of inputs.

Layer 5

This layer is also called the summation layer. The initial ANFIS output can be achieved by adding the outputs obtained for the defuzzification layer rules.

$$O_i^5 = \sum_i \overline{w_i} f_i = \frac{\sum_i w_i f_i}{\sum_i w_i}$$

2.3.2. Development of the Adaptive Neuro-Fuzzy Inference System

ANFIS is not different from the artificial neural network in terms of heuristic model features. The significant dissimilarity is that the ANFIS model is metaheuristic in nature and comprises the combination of a fuzzy inference system and an artificial neural network to achieve an improved fuzzy inference system. The dataset from the four road intersections was divided into three groups to achieve an optimal traffic-flow predictive performance of the vehicles by using the adaptive neuro-fuzzy inference model. Because the traffic datasets obtained from these road intersections were raw data, to obtain an optimal solution there is a need for continuous iterations. The total number and membership function's structure, not excluding the optimization algorithm, are important in developing the model architecture. In this study, two membership functions were used during the ANFIS development. For the input variables, we used the triangular membership function (trimf) such as speed with traffic density for the input variable, and the output variable was represented by the Gaussian membership function (Gaussmf) (traffic volume). The model was implemented (trained) in the MATLAB 2020a interface. Table 2 tabulates the proposed training algorithms based on an approximate reasoning for evaluating the traffic flow of vehicles at these road intersections. The fuzzy logic controller system applied in this study was dependent on the IF-THEN rules. The fuzzy rules applied for the four signalized road intersections are dependent on the traffic-data and traffic-flow variables as shown in Table 2:

Table 2. Fuzzy rules for the road intersections.

Road Intersections	Intersection_1	Intersection_2	Intersection_3	Intersection_4
Vehicular speed	High	High	High	High
Distance	Low	Low	Very Low	Low
Estimated time	Very Low	Very Low	Very Low	Low

The input variables, comprising speed of vehicle and time, were copied from the Microsoft excel sheet into the fuzzy toolbox which can be found in the "anfised interface" in the MATLAB environment that uses a class function called the Sugeno which is used for the fuzzification. To be able to apply the fuzzy rules on the obtained traffic datasets, we evaluated the membership functions. The fuzzy rule viewers use the continuous discretization triangular inputs to analyze the required alpha cuts needed to develop the fuzzy inference system inputs. Figure 6 depicts the interrelationship between the input and output membership functions.

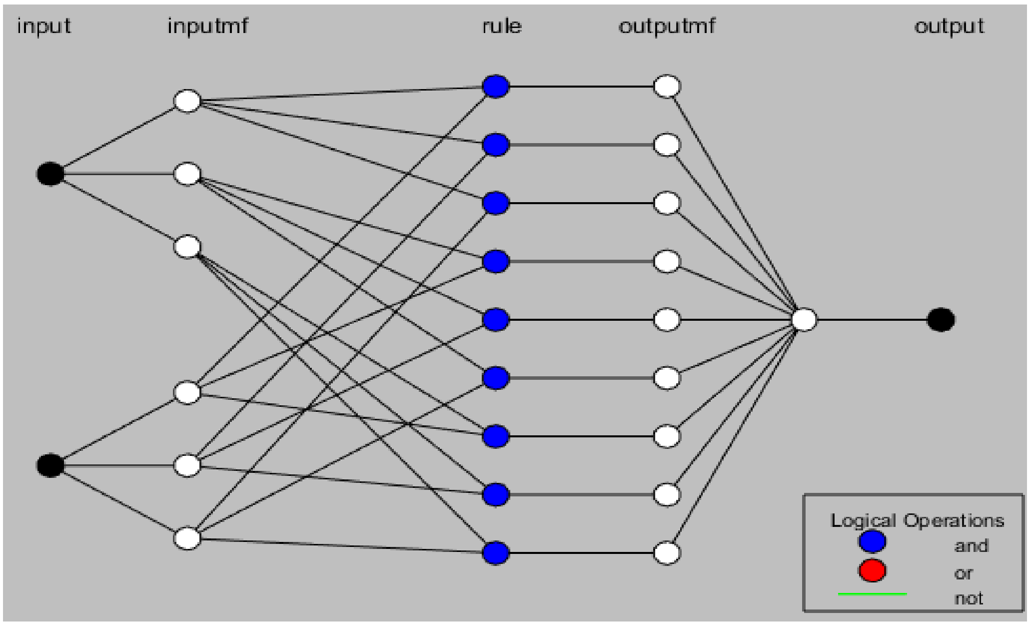

Figure 6. A typical structure of an ANFIS model showing the inputs and output.

These input- and output-variables approach was devised based on the approach used by [7,41,42]. The ANFIS model development, training, and testing were conducted in the MATLAB 2020a user interface tools and command-line functionality. Table 3 shows the parameters used in the training and testing of the model in the MATLAB environment in tabular form. Table 4 shows how the traffic datasets were divided and used to train and test the ANFIS and ANFIS-GA models.

Table 3. The parameters of the ANFIS model.

Parameters	Values
Type	Takagi–Sugeno
Clustering technique	Subtractive clustering
Input MF type	Gaussian
Output MF type	Linear
Number of rules	10
Number of clusters	5
Maximum iterations	100
Defuzzification method	Weighted average
Cluster radius	0.67

Table 4. The breakdown of the traffic datasets.

Signalized Road Intersections	Training	Testing	Total
Intersection 1	144	62	206
Intersection 2	150	65	215
Intersection 3	157	67	224
Intersection 4	144	61	205
	595	**255**	**850**

Note. Traffic data samples were taken from each intersection and divided based on 70% for training and 30% for testing. Furthermore, the data training requires a larger traffic dataset compared to the traffic data needed for testing the model. This is due to the need to train the model with as much data as possible to improve the accuracy and efficiency of the model when it comes to prediction and pattern recognition. When traffic data are entered into the ANFIS and ANFIS-GA models, they learn the patterns from the traffic-flow data and make decisions based on this data.

2.4. Adaptive Neuro-Fuzzy Inference System Optimized by Genetic Algorithm (ANFIS-GA)

A genetic algorithm, also known as a GA, can be defined as a global search heuristics method applied in various engineering applications to evaluate or determine solutions for optimization problems. It is used to analyze and evaluate complex search problems in optimization. A genetic algorithm is a special evolutional algorithm technique that is significant when selecting or choosing a natural selection evolutionary process that uses inheritance, mutation, selection, and recombination [43]. In this research, we combined GA with ANFIS to increase the efficiency of the ANFIS predictive model. GA was combined with ANFIS to enhance the ANFIS model's performance and reduce the rates of error of the regression values by using the tuning and optimization method on the membership functions in the Sugeno-type fuzzy inference system. The development of the hybrid ANFIS-GA model is displayed in Figure 7.

The genetic algorithm model starts with a group of solutions, also known as chromosomes. A new population is created by completing a previous population. The fitness of new solutions formed from selected (offspring) is assigned. This process is carried out repeatedly till a condition, i.e., the improvement of the optimal solution, is met. To accomplish this, the ANFIS algorithm, which is part of the fitness function, plays an important role *f(x)*. The fitness with ANFIS fitness function intervention is represented by:

$$f_1(x) - \frac{1}{m}\sqrt{\sum_i^m o(d_i - a_i)^2}$$

where m is known as the number of characteristics, a_i is output derived from the ANFIS, and d_i is the predicted traffic volume.

The next fitness function is:

$$f_2(x) - \frac{1}{n-m}\sqrt{\sum_i^n m(d_i - a_i)^2}$$

where n is known as the overall number of input characteristics, d_i is regarded as the minimum, a_i is represented as the actual value of the traffic volume, and $n - m$ indicates the left-over undesired attributes.

The final equation is known as minimized *f(x)*, and it is represented as:

$$f(x) - \frac{f_1(x) + f_2(x)}{2}$$

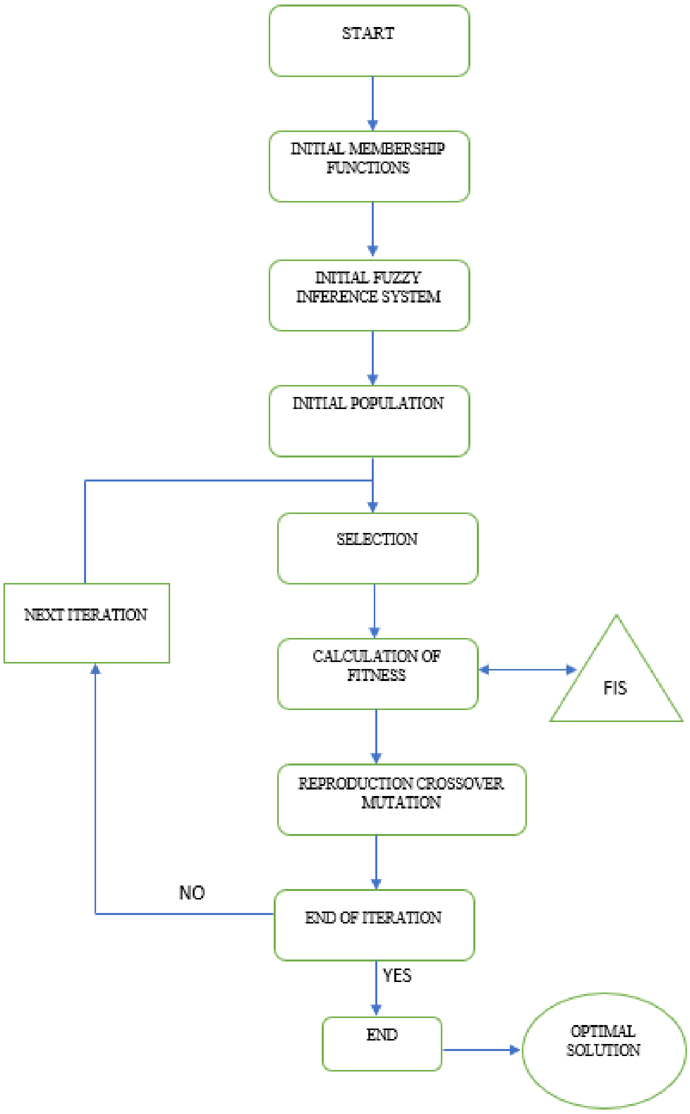

Figure 7. ANFIS-GA model.

For this research, we evaluated the initialization of the parameters for the genetic algorithm. These comprise the overall number of iterations, population size, percentage of mutation, and crossover. It is very important to note that the selection of these parameters is, to a great extent, the determinant when it comes to the capability of the designed controller. The tuning parameter range of the ANFIS-GA is illustrated in Table 5.

Table 5. Tuning parameters used for the ANFIS-GA model.

Population Size	Number of Iterations	Percentage of Crossover	Percentage of Mutation	Rate of Mutation	Selection Pressure	Selection Function
100	500	0.8%	0.3%	0.02	8	Roulette wheel

Note. These tuning parameters can be found in the ANFIS-GA codes; changing these parameters has a significant impact on the ANFIS-GA model training and testing.

In addition, as soon as the fitness $f(x)$ of each chromosome x in the population is investigated and evaluated, a will be created. The steps above are carried out repeatedly until the training and testing of the model are completed. The caveat is that the better the fitness, the bigger the chance of the parent chromosomes being chosen. This leads to crossover for the parents to create brand-new offspring that possess a crossover probability. The next available mutation will create a probability leading to a mutation of new offspring at each available position in the chromosome. The solution will undergo reproduction, crossover, and mutation coupled with parameter settings found in Table 5. The optimal solution in the present population will likely return when the end condition is satisfied. Evaluating the optimal solution will assist the genetic algorithm in searching for an optimal membership function.

2.5. Statistical Indicators Used for Evaluation of the ANFIS and ANFIS-GA Models

The ANFIS and ANFIS-GA models can be validated using different statistical methods. The most acknowledged statistical error indicators are the mean absolute bias error (MABE), mean absolute percentage error (MAPE), root mean square error (RMSE), and coefficient of determination R^2. However, in this research, we only focused on the RMSE and R^2.

a. The mean absolute bias error can be defined as the average overall quantity of all the absolute bias errors determined when comparing the actual and predicted traffic volume. It is mathematically denoted as:

$$\mathrm{MABE} = \frac{1}{N}\sum\nolimits_{i=1}^{N} |(V_{i,P} - V_{i,M})|$$

b. The MAPE is the mean absolute percentage difference that can be determined between predicted and actual traffic volumes. This is mathematically stated as:

$$\mathrm{MAPE} = \frac{1}{N}\sum\nolimits_{i=1}^{N} \left|\left(\frac{V_{i,P} - V_{i,M}}{V_{i,M}}\right)\right| \times 100$$

c. The RMSE is determined by knowing the model's accuracy, which is determined by calculating the comparison between the predicted traffic volume and the actual traffic volume. The value is always positive and not negative. It is mathematically represented as:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(V_{i,P} - V_{i,M})^2}$$

d. The R^2 also known as the coefficient of determination, signifies the optimal relationship between both the predicted traffic volume of vehicles at the four signalized road intersections and the actual traffic volume. This is mathematically represented as:

$$R^2 = 1 - \frac{\frac{1}{N}\sum_{i=1}^{N}(V_{i,P} - V_{i,M})^2}{\frac{1}{N}\sum_{i=1}^{x}(V_{i,P} - V_{M,avg})^2}$$

where $V_{i,P}$ represents the predicted traffic volume of vehicles at the four signalized road intersections and $V_{i,M}$ indicates the actual traffic volume of vehicles on these signalized road intersections. Figure 8 below shows the breakdown of the Off-peak and On-peak hours of traffic flow of vehicles.

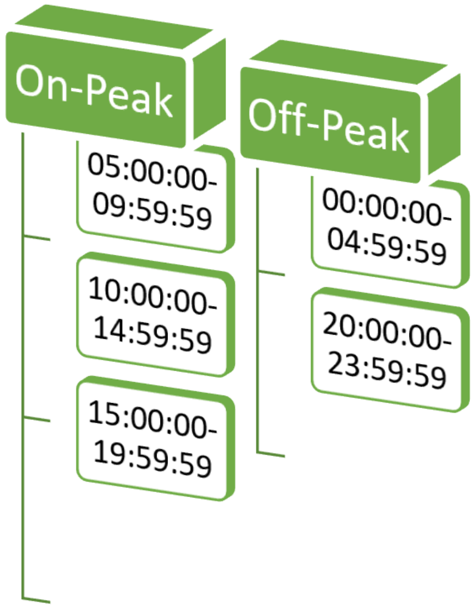

Figure 8. Off-peak and on-peak hours. Note: Figure 8 shows the period of the day based on the on-peak and off-peak hours.

3. Results and Discussions

3.1. Prediction of Traffic Volume at These Four Signalized Road Intersections

- **Signalized Road Intersection 1**

From Figure 9, the single most striking observation is that during period 3 (10:00:00–14:59:59), there are usually many vehicles on the road except on Sundays, which always experiences smaller traffic volumes.

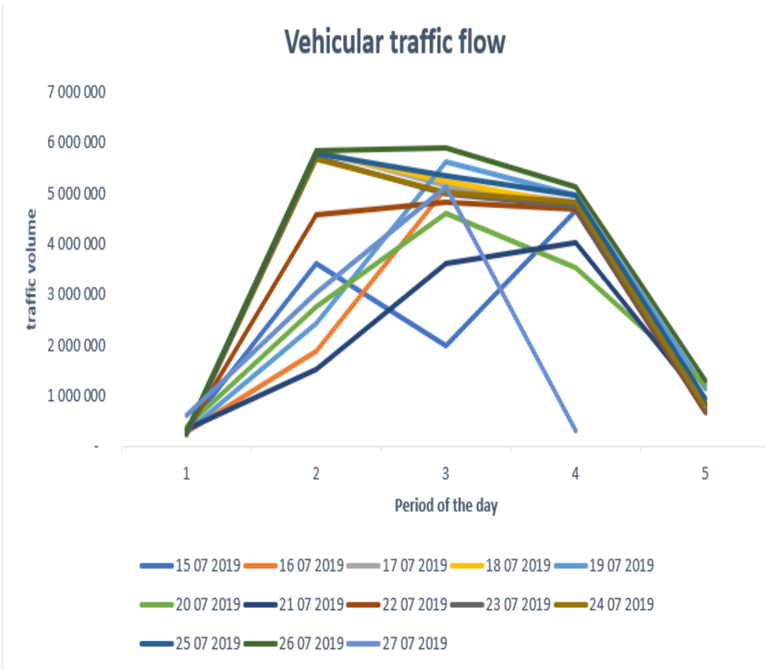

Figure 9. Road intersection 1.

- **Signalized Road Intersection 2**

From Figure 10, the most interesting find in this road intersection is that the on-peak and off-peak periods of the day play an enormous role in determining when there is going to be high or lower traffic volume unless there is non-recurrent traffic congestion (breakdown of vehicle or road accident on the road).

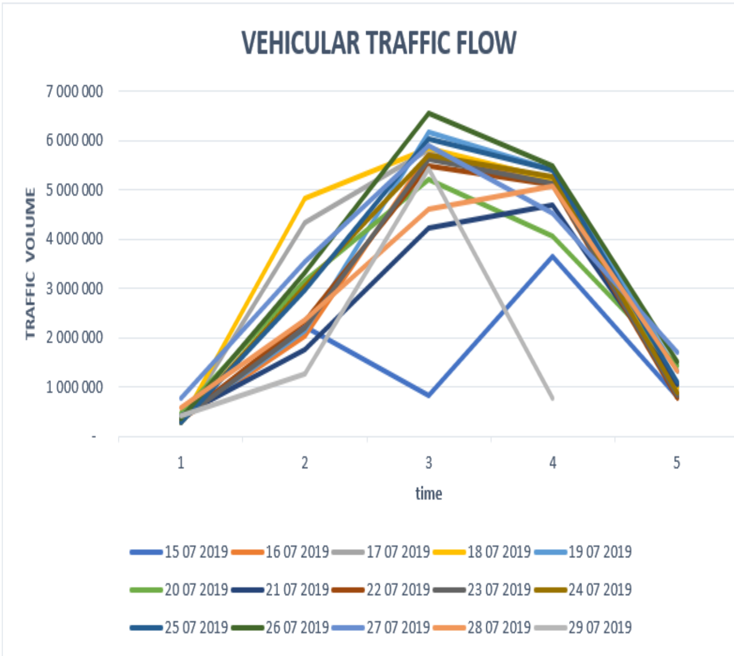

Figure 10. Road intersection 2.

- **Signalized Road Intersection 3**

From Figure 11, It is interesting to note that in road intersection 3, on-peak periods are usually between 10:00:00 and 14:59:59 and 15:00:00 and 19:59:59, and the on-peak periods are mainly on weekdays. Reduced traffic volume is prominent during weekends and public holidays.

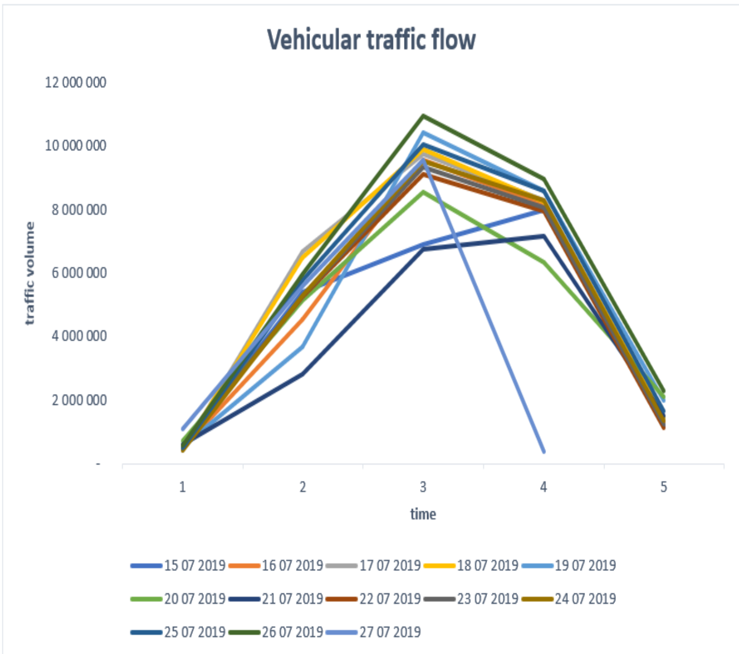

Figure 11. Road intersection 3.

- **Signalized Road Intersection 4**

From Figure 12, another significant result when observing road intersection 4 is that the periods 05:00:00–09:59:59, 10:00:00–14:59:59, and 15:00:00–19:59:59 are all regarded as on-peak periods, which is when there is an increase in traffic volume.

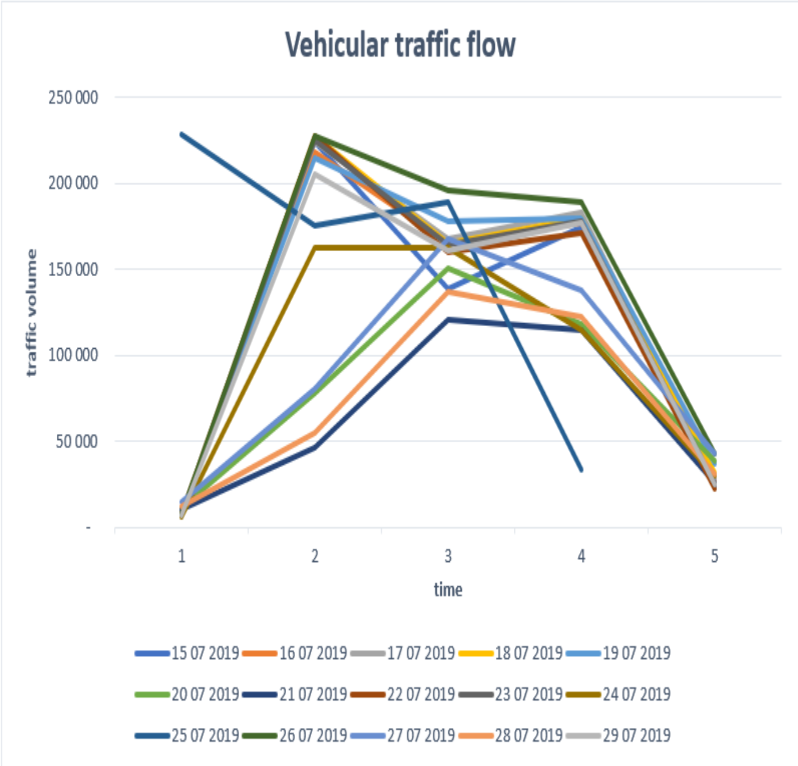

Figure 12. Road intersection 4.

3.2. ANFIS-GA

In this research study, ANFIS-GA and standalone ANFIS models were developed to predict the traffic flow of non-autonomous vehicles at signalized road intersections using the South Africa Road transportation systems as a case study. The above-mentioned models were trained and tested using traffic data from the signalized road intersections. The traffic data were divided according to the data collected from these road intersections.

Based on the results obtained when we compared the ANFIS-GA and ANFIS models results showed that the performance of the predictive models was dependent on the lowest root mean square error (RMSE). According to Figures 13 and 14, the ANFIS-GA results show that the performance of the training and testing was based on the size of population (n = 100), the number of iterations (n = 500), the percentage of crossovers, the rate of mutation, and the selection function. Figures 15 and 16 show the ANFIS-GA model's corresponding performance on the traffic volume of vehicles at the four signalized road intersections. It is significant to note that although the ANFIS model sometimes produced the best R^2 performance, the results of this study have shown that ANFIS-GA provided the best model performance for any of the four signalized road intersections from which the traffic data is collected. This is illustrated by the R^2 values in the testing of the ANFIS-GA model in Figure 16, which shows 0.9980 as compared to the testing of the ANFIS model, which shows R^2 = 0.9790 in Figure 20.

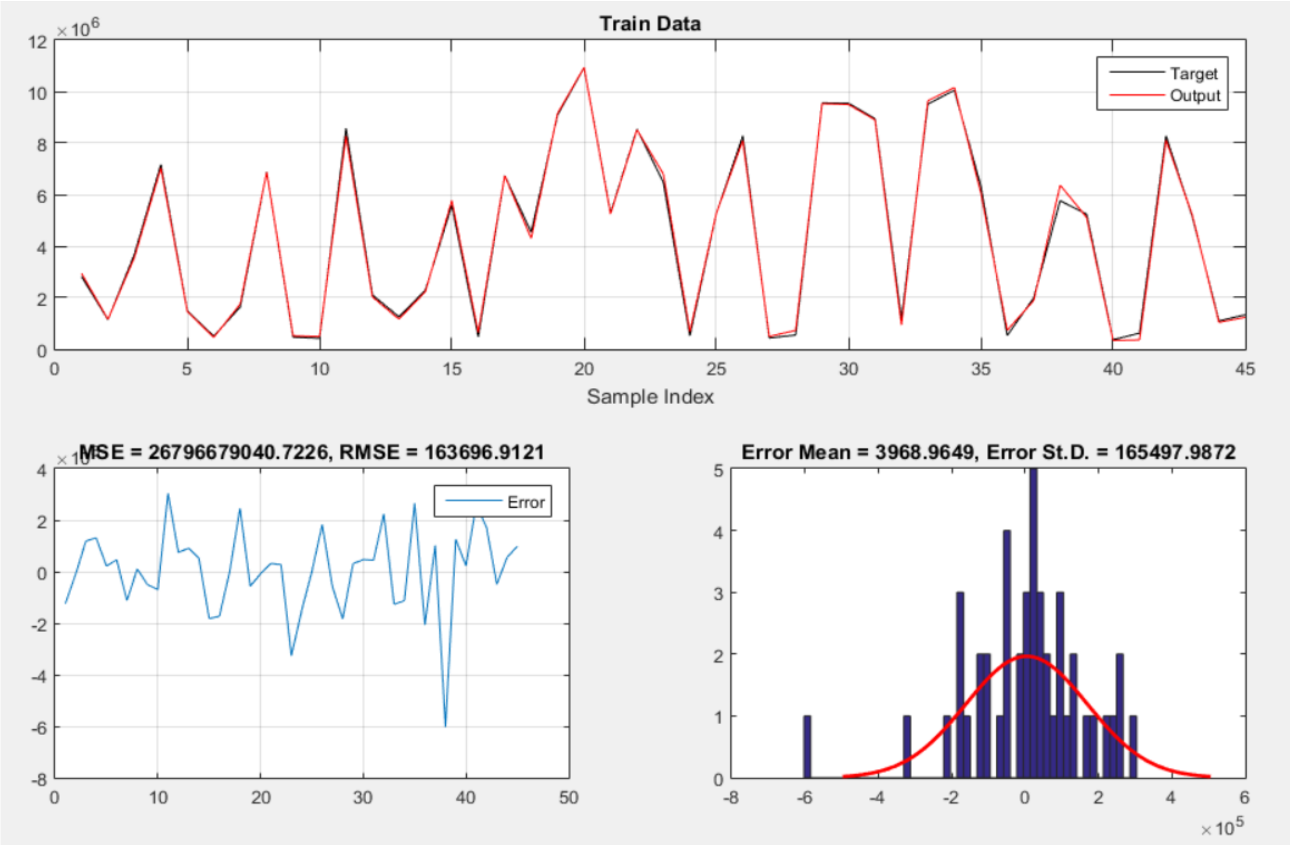

Figure 13. Training of the ANFIS-GA model with the traffic datasets.

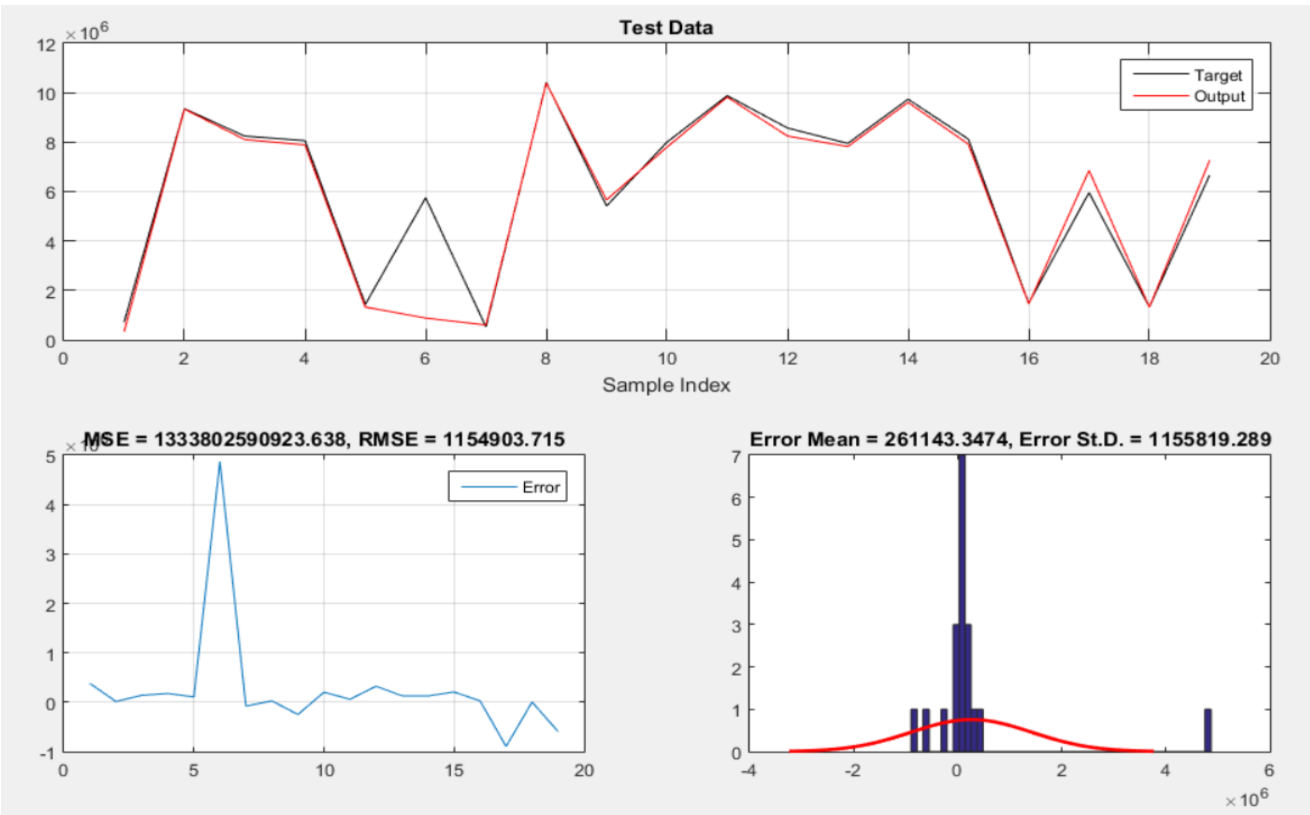

Figure 14. ANFIS-GA model testing with the traffic datasets.

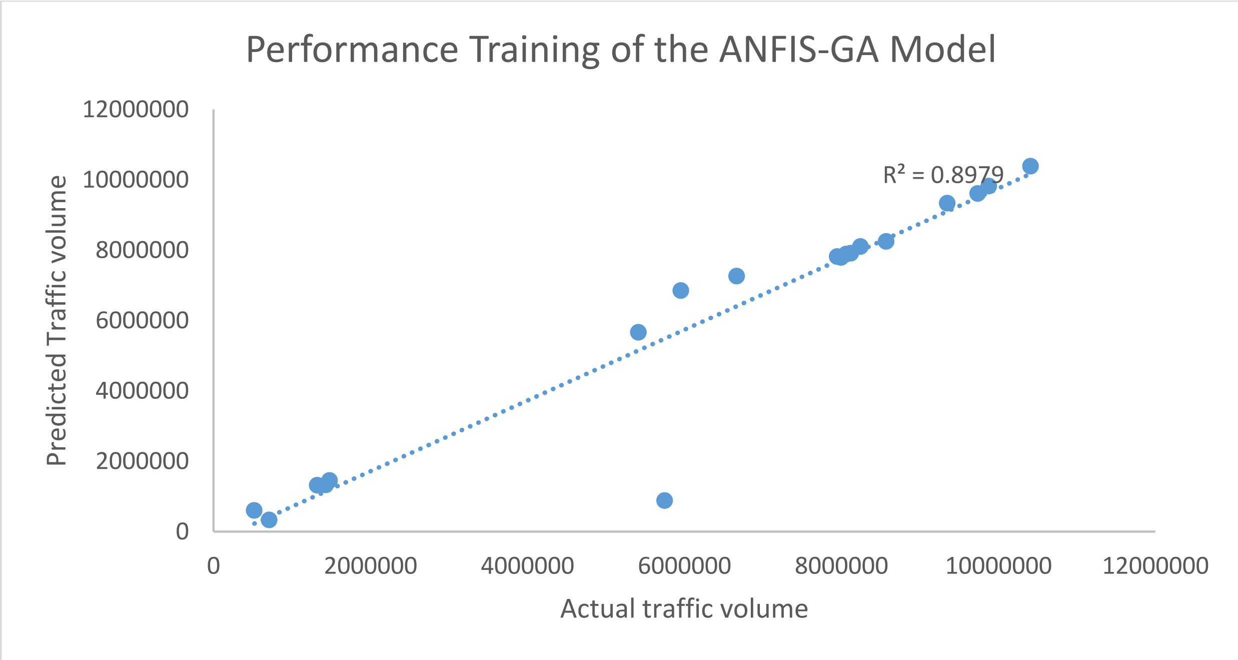

Figure 15. ANFIS-GA model training.

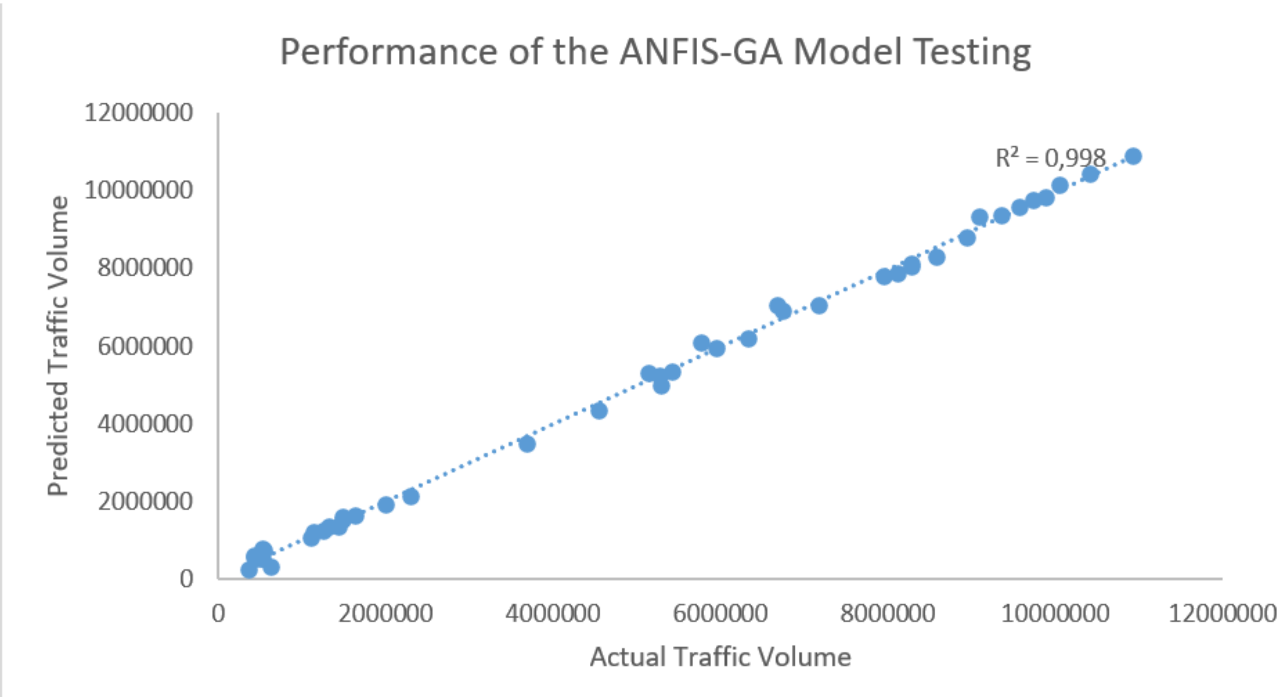

Figure 16. ANFIS-GA model testing.

The 850 traffic datasets obtained from the four congested signalized road intersections on the South African road network were divided into 595 and 255 for the training and testing of the ANFIS-GA and ANFIS models, respectively. The division of the traffic data for each of the four intersections is shown in tabulated form in Table 4. A trial-and-error method was used to achieve the optimum performance of the ANFIS and ANFIS-GA models. This method was used during the ANFIS and ANFIS-GA model training and testing to determine the optimal parameter for the number of populations and iterations. The best optimal parameters for the ANFIS-GA and ANFIS performance training and testing are shown in Figures 15–18. It is interesting to note that in all four signalized road intersections, the ANFIS-GA and ANFIS models' parameters significantly impacted the predictive performance of the models.

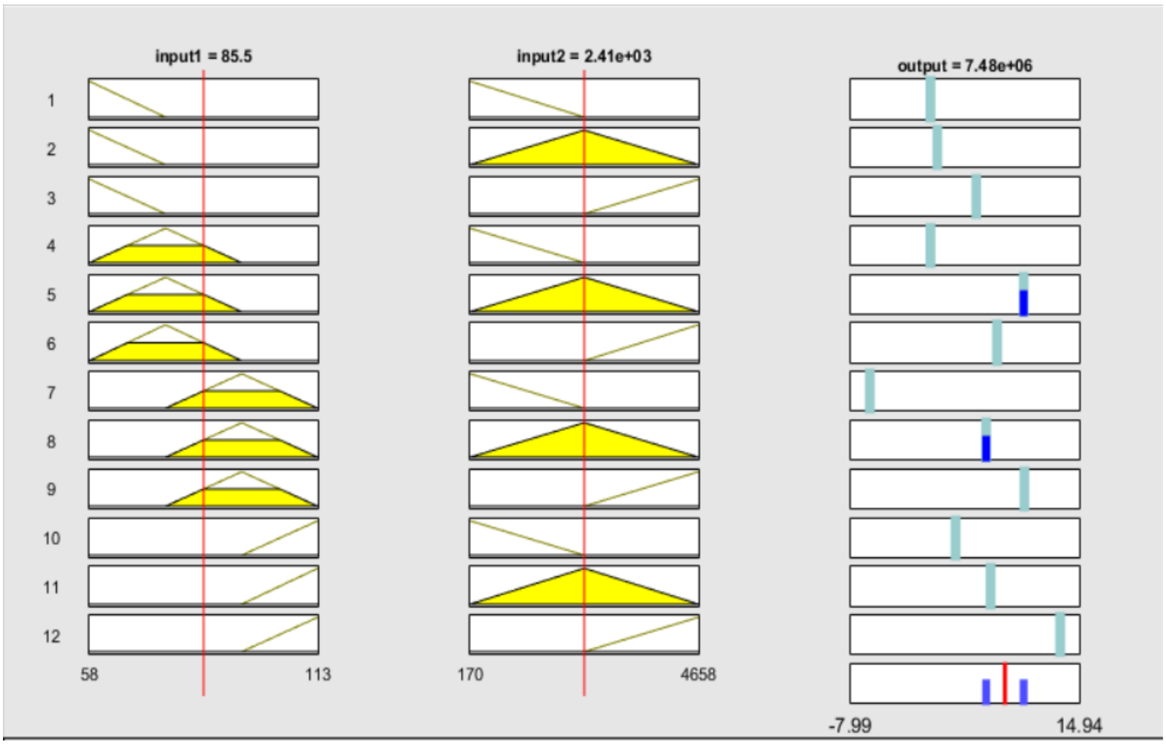

Figure 17. ANFIS model rule viewer.

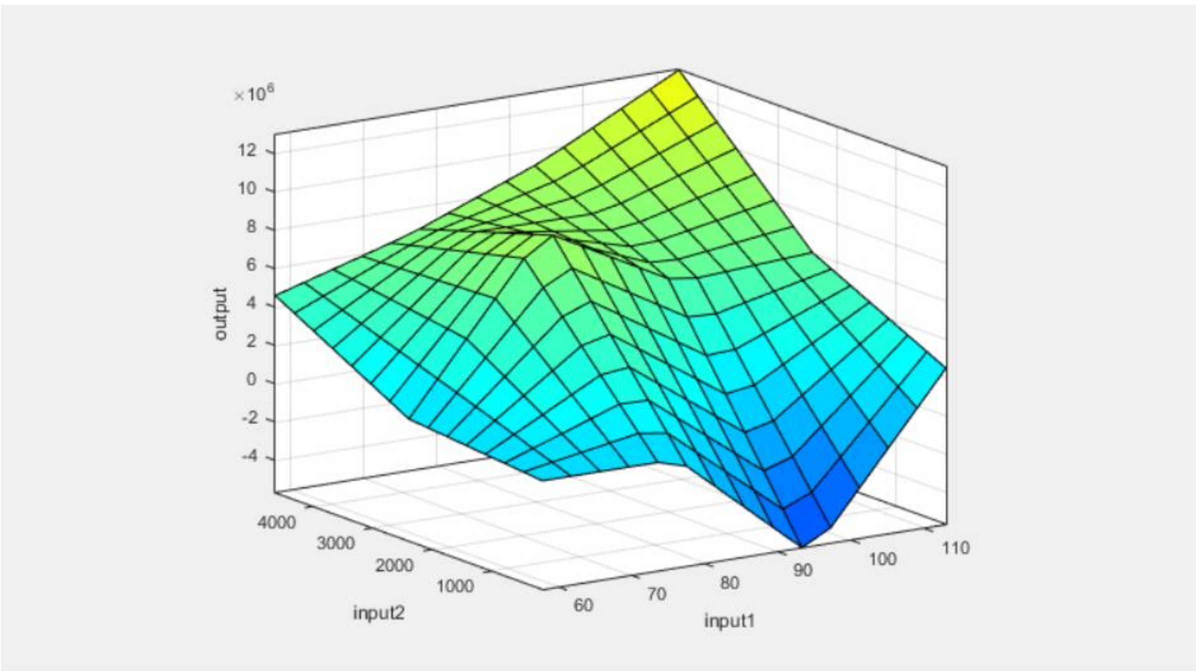

Figure 18. 3D image of the ANFIS model. (Input 1= time, input 2 = speed of vehicles, output = traffic volume).

The fuzzy rule viewer representation used in this research uses a continuous discretization of the inputs triangular mf to evaluate the cuts by alpha technique to create the fuzzy inference inputs (Figure 17).

Figure 18 shows the three-dimensional surface views of the inputs and output using the ANFIS rule viewer.

To determine the ANFIS model training accuracy, we compared the observed and predicted output of the traffic volume of vehicles in Figure 19, with the training of the ANFIS model accuracy being 0.9709.

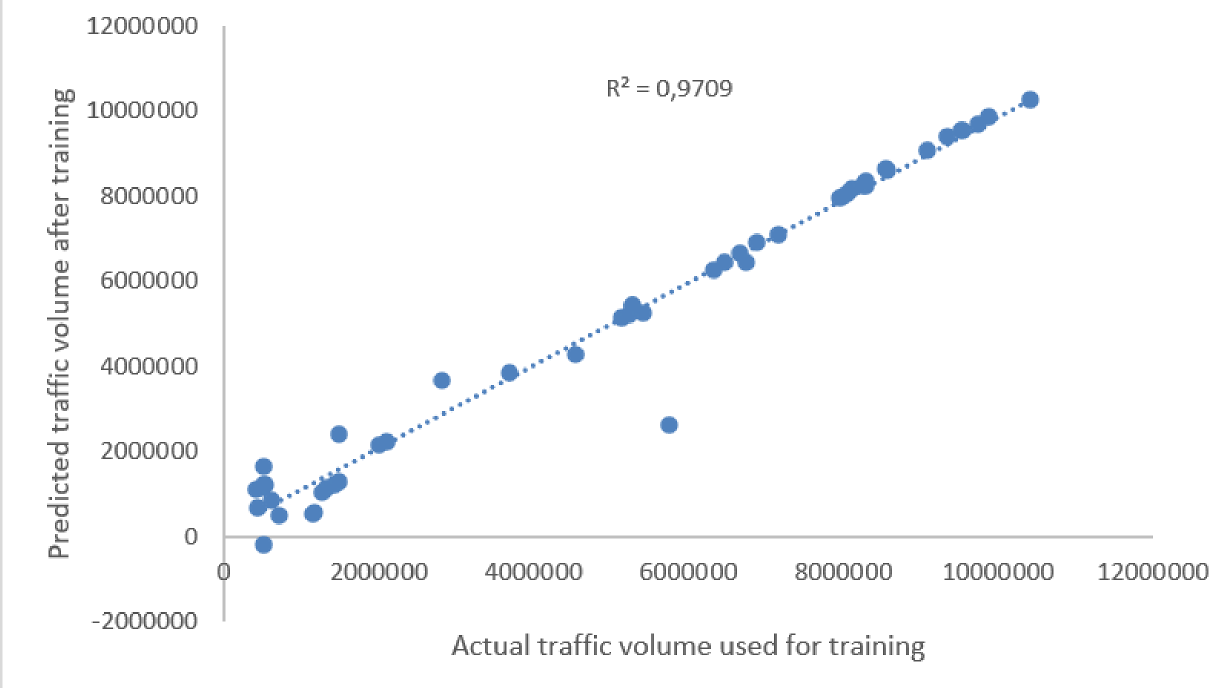

Figure 19. ANFIS model training accuracy.

The observed and predicted output of the traffic volume of vehicles at the road intersections were compared in Figure 20 to determine the accuracy of the testing of the ANFIS model with the testing accuracy being 0.9790.

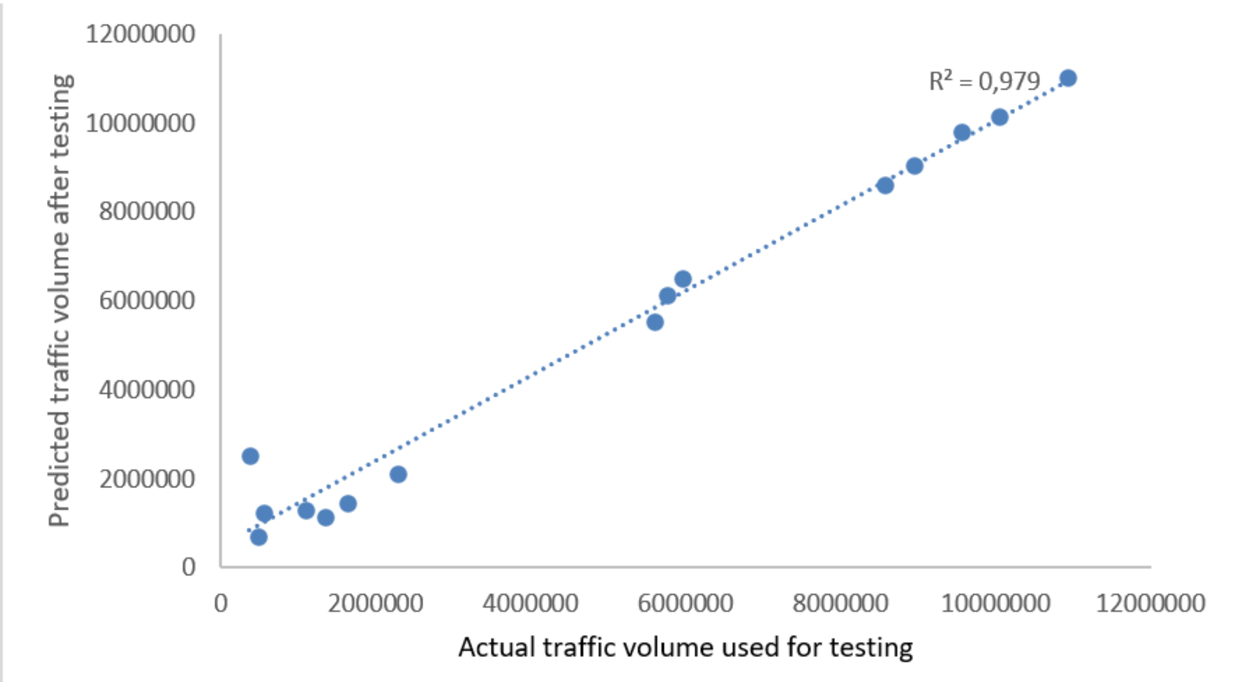

Figure 20. ANFIS model testing accuracy.

4. Conclusions and Recommendations

The main goal of the current study was to determine the prediction and modelling of data-driven novel vehicular traffic flow of non-autonomous vehicles at a signalized road intersection using optimized fuzzy inference systems with genetic algorithms. Based on the results and discussions, the following conclusions can be drawn:

- The relevance of traffic volume in evaluating the vehicular traffic flow at road intersections was supported by the ANFIS and ANFIS-GA models' research findings in predicting and modelling traffic volume of vehicles.
- The obvious finding from this study is that the ANFIS-GA ($R_{Training} = 0.8979$ and $R_{Testing} = 0.9980$) results when compared to the ANFIS ($R_{Training} = 0.9709$ and $R_{Testing} = 0.9790$) is more suitable and appropriate for the modelling and predicting of traffic volume.
- The second significant finding was that this research study has torchlit the importance of off-peak and on-peak periods in evaluating the vehicular traffic flow at signalized road intersections.
- The evidence from the ANFIS and ANFIS-GA model results suggests that the traffic data inputs, and outputs (traffic volume, time, and speed of vehicles) are well correlated to each other.
- The adaptive neuro-fuzzy inference system (ANFIS) and adaptive neuro-fuzzy inference system optimized by genetic algorithm (ANFIS-GA) are reliable traffic-flow predictors of traffic volume at road intersections.
- This research has shown that the important advantage of using a hybrid ANFIS-GA in modelling the traffic flow of vehicles is that it tunes the ANFIS model's membership functions to reduce the error parameters during the training and testing of the models.

The findings of this study make a significant contribution to the current literature on traffic-flow prediction and how important it is to traffic-volume prediction in reducing traffic congestion at signalized road intersections.

Recommendations

- Further research needs to be carried out to improve the efficiency of traffic-volume prediction. A spatial relationship between light-, heavy-, and truck-traffic volume should be investigated.
- Further study can be carried out to evaluate the effects of other traffic-flow parameters (traffic density and speed of vehicles) on the modelling of traffic volume.
- Further study could assess the effects of genetic and particle swarm optimization algorithms on vehicle traffic volume at unsignalized road intersections.
- Finally, further study should focus on the comparison of the results of this research with other machine-learning models.
- A major limitation of this study is that it does not consider the different weather that can influence the collection of traffic datasets and flow of vehicles. Future transportation research should focus on factoring in the interference of different seasons and weather conditions on the flow of vehicles and how this can impede the collection of traffic datasets on freeways and road intersections.

Author Contributions: Conceptualization, I.O.O. and L.K.T.; methodology, I.O.O., L.K.T. and F.J.A.; validation, I.O.O., L.K.T. and F.J.A.; formal analysis, I.O.O. and F.J.A.; investigation, I.O.O. and L.K.T.; writing—original draft preparation I.O.O. and F.J.A.; writing—review and editing, I.O.O. and F.J.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: Not Applicable.

Data Availability Statement: The traffic data used for this research are available on reasonable request from the corresponding author. The ANFIS-GA codes used in this research can be found in https://yarpiz.com/319/ypfz104-evolutionary-anfis-training (accessed on 15 July 2022).

Acknowledgments: The authors acknowledge the University of Johannesburg, South Africa; the University of Wollongong, NSW, Australia; and the Wuhan University of Technology, China.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Barberi, S.; Arena, F.; Termine, F.; Canale, A.; Olayode, I.O.; Zuccalà, Y. BIM applied to intelligent transport systems. *Proc. AIP Conf. Proc.* **2022**, *2611*, 060011.
2. Barberi, S.; Arena, F.; Termine, F.; Canale, A.; Olayode, I.O. Safety aspects of intelligent transport systems applied to road intersections. *Proc. AIP Conf. Proc.* **2022**, *2611*, 060012.
3. Barberi, S.; Arena, F.; Termine, F.; Canale, A.; Zuccalà, Y.; Olayode, I.O. Smart city: Effects on urban mobility and expected developments due to autonomous vehicles. *Proc. AIP Conf. Proc.* **2022**, *2611*, 060014.
4. Ghorbanzadeh, O.; Moslem, S.; Blaschke, T.; Duleba, S. Sustainable urban transport planning considering different stakeholder groups by an interval-AHP decision support model. *Sustainability* **2018**, *11*, 9. [CrossRef]
5. Zhang, J.; Wang, F.-Y.; Wang, K.; Lin, W.-H.; Xu, X.; Chen, C. Data-driven intelligent transportation systems: A survey. *IEEE Trans. Intell. Transp. Syst.* **2011**, *12*, 1624–1639. [CrossRef]
6. Isaac, O.O.; Tartibu, L.K.; Okwu, M.O. Prediction and Modelling of Traffic Flow of Human-driven Vehicles at a Signalized Road Intersection Using Artificial Neural Network Model: A South Africa Road Transportation System Scenario. *Transp. Eng.* **2021**, *6*, 100095.
7. Olayode, I.O.; Severino, A.; Campisi, T.; Tartibu, L.K. Prediction of Vehicular Traffic Flow using Levenberg-Marquardt Artificial Neural Network Model: Italy Road Transportation System. *Commun.-Sci. Lett. Univ. Zilina* **2022**, *24*, E74–E86. [CrossRef]
8. 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; pp. 287–292.
9. Li, Y.; Shahabi, C. A brief overview of machine learning methods for short-term traffic forecasting and future directions. *Sigspatial Spec.* **2018**, *10*, 3–9. [CrossRef]
10. Devi, S.; Neetha, T. Machine Learning based traffic congestion prediction in a IoT based Smart City. *Int. Res. J. Eng. Technol.* **2017**, *4*, 3442–3445.
11. Peng, X.; Zhu, H.; Feng, J.; Shen, C.; Zhang, H.; Zhou, J.T. Deep clustering with sample-assignment invariance prior. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *31*, 4857–4868. [CrossRef] [PubMed]
12. Hu, P.; Peng, D.; Wang, X.; Xiang, Y. Multimodal adversarial network for cross-modal retrieval. *Knowl.-Based Syst.* **2019**, *180*, 38–50. [CrossRef]
13. Bratsas, C.; Koupidis, K.; Salanova, J.-M.; Giannakopoulos, K.; Kaloudis, A.; Aifadopoulou, G. A comparison of machine learning methods for the prediction of traffic speed in urban places. *Sustainability* **2020**, *12*, 142. [CrossRef]
14. Boukerche, A.; Wang, J. Machine Learning-based traffic prediction models for Intelligent Transportation Systems. *Comput. Netw.* **2020**, *181*, 107530. [CrossRef]
15. Xie, P.; Li, T.; Liu, J.; Du, S.; Yang, X.; Zhang, J. Urban flow prediction from spatiotemporal data using machine learning: A survey. *Inf. Fusion* **2020**, *59*, 1–12. [CrossRef]
16. Bao, W.; Yuan, D.; Yang, Z.; Wang, S.; Zhou, B.; Adams, S.; Zomaya, A. sFOG: Seamless fog computing environment for mobile IoT applications. In Proceedings of the 21st ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems, Montreal, QC, Canada, 28 October–2 November 2018; pp. 127–136.
17. Hu, W.; Feng, Z.; Chen, Z.; Harkes, J.; Pillai, P.; Satyanarayanan, M. Live synthesis of vehicle-sourced data over 4G LTE. In Proceedings of the 20th ACM International Conference on Modelling, Analysis and Simulation of Wireless and Mobile Systems, Miami Beach, FL, USA, 21–25 November 2017; pp. 161–170.
18. Severino, A.; Pappalardo, G.; Olayode, I.O.; Canale, A.; Campisi, T.J.T.E. Evaluation of the Environmental Impacts of Bus Rapid Transit System on Turbo Roundabout. *Transp. Eng.* **2022**, *9*, 100130. [CrossRef]
19. Kwon, H. Friend-guard textfooler attack on text classification system. *IEEE Access* **2021**, 1. [CrossRef]
20. Olayode, I.O.; Severino, A.; Tartibu, L.K.; Arena, F.; Cakici, Z.J.I. Performance Evaluation of a Hybrid PSO Enhanced ANFIS Model in Prediction of Traffic Flow of Vehicles on Freeways: Traffic Data Evidence from South Africa. *Infrastructures* **2022**, *7*, 2. [CrossRef]
21. Severino, A.; Pappalardo, G.; Curto, S.; Trubia, S.; Olayode, I.O. Safety Evaluation of Flower Roundabout Considering Autonomous Vehicles Operation. *Sustainability* **2021**, *13*, 10120. [CrossRef]
22. Tollazzi, T.; Tesoriere, G.; Guerrieri, M.; Campisi, T. Environmental, functional and economic criteria for comparing "target roundabouts" with one-or two-level roundabout intersections. *Transp. Res. Part D Transp. Environ.* **2015**, *34*, 330–344. [CrossRef]

23. Domingues, A.C.; Silva, F.A.; Loureiro, A.A. On the Analysis of Users' Behavior Based on Mobile Phone Apps. In Proceedings of the 17th ACM International Symposium on Mobility Management and Wireless Access, Miami Beach, FL, USA, 25–29 November 2019; pp. 25–32.
24. Olayode, I.; Tartibu, L.; Okwu, M.; Uchechi, U. Intelligent transportation systems, un-signalized road intersections and traffic congestion in Johannesburg: A systematic review. *Procedia CIRP* **2020**, *91*, 844–850. [CrossRef]
25. Alipour, B.; Tonetto, L.; Ketabi, R.; Yi Ding, A.; Ott, J.; Helmy, A. Where are you going next? A practical multi-dimensional look at mobility prediction. In Proceedings of the 22nd International ACM Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems, Miami Beach, FL, USA, 25–29 November 2019; pp. 5–12.
26. Meneguette, R.I.; Nakamura, L.H. A flow control policy based on the class of applications of the vehicular networks. In Proceedings of the 15th ACM International Symposium on Mobility Management and Wireless Access, Miami Beach, FL, USA, 21–25 November 2017; pp. 137–144.
27. Sun, S.; Zhang, C.; Yu, G. A Bayesian network approach to traffic flow forecasting. *IEEE Trans. Intell. Transp. Syst.* **2006**, *7*, 124–132. [CrossRef]
28. Du, B.; Zhang, C.; Shen, J.; Zheng, Z. A Dynamic Sensitivity Model for Unidirectional Pedestrian Flow With Overtaking Behaviour and Its Application on Social Distancing's Impact During COVID-19. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 10404–10417. [CrossRef]
29. Mackenzie, J.; Roddick, J.F.; Zito, R. An evaluation of HTM and LSTM for short-term arterial traffic flow prediction. *IEEE Trans. Intell. Transp. Syst.* **2018**, *20*, 1847–1857. [CrossRef]
30. Hou, Q.; Leng, J.; Ma, G.; Liu, W.; Cheng, Y. An adaptive hybrid model for short-term urban traffic flow prediction. *Phys. A Stat. Mech. Its Appl.* **2019**, *527*, 121065. [CrossRef]
31. Jiang, J.; Dellaert, N.; Van Woensel, T.; Wu, L. Modelling traffic flows and estimating road travel times in transportation network under dynamic disturbances. *Transp. Trans. GIS* **2019**, *47*, 2951–2980. [CrossRef]
32. Chen, D. Research on traffic flow prediction in the big data environment based on the improved RBF neural network. *IEEE Trans. Ind. Inform.* **2017**, *13*, 2000–2008. [CrossRef]
33. Chen, Q.; Song, Y.; Zhao, J. Short-term traffic flow prediction based on improved wavelet neural network. *Neural Comput. Appl.* **2020**, *33*, 1–10. [CrossRef]
34. Olayode, I.O.; Tartibu, L.K.; Campisi, T. Stability Analysis and Prediction of Traffic Flow of Trucks at Road Intersections Based on Heterogenous Optimal Velocity and Artificial Neural Network Model. *Smart Cities* **2022**, *5*, 1092–1114. [CrossRef]
35. Dell'Acqua, P.; Bellotti, F.; Berta, R.; De Gloria, A. Time-aware multivariate nearest neighbor regression methods for traffic flow prediction. *IEEE Trans. Intell. Transp. Syst.* **2015**, *16*, 3393–3402. [CrossRef]
36. Olayode, I.O.; Tartibu, L.K.; Okwu, M.O.; Ukaegbu, U.F. Development of a Hybrid Artificial Neural Network-Particle Swarm Optimization Model for the Modelling of Traffic Flow of Vehicles at Signalized Road Intersections. *Appl. Sci.* **2021**, *11*, 8387. [CrossRef]
37. Zheng, F.; Liu, C.; Liu, X.; Jabari, S.E.; Lu, L. Analyzing the impact of automated vehicles on uncertainty and stability of the mixed traffic flow. *Transp. Res. Part C Emerg. Technol.* **2020**, *112*, 203–219. [CrossRef]
38. Polson, N.; Sokolov, V. Bayesian analysis of traffic flow on interstate I-55: The LWR model. *Ann. Appl. Stat.* **2015**, *9*, 1864–1888. [CrossRef]
39. Jang, J.-S. ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Trans. Syst. Man Cybern.* **1993**, *23*, 665–685. [CrossRef]
40. Ghorbanzadeh, O.; Rostamzadeh, H.; Blaschke, T.; Gholaminia, K.; Aryal, J. A new GIS-based data mining technique using an adaptive neuro-fuzzy inference system (ANFIS) and k-fold cross-validation approach for land subsidence susceptibility mapping. *Nat. Hazards* **2018**, *94*, 497–517. [CrossRef]
41. Zaki, J.; Ali-Eldin, A.; Hussein, S.E.; Saraya, S.F.; Areed, F.F. Framework for Traffic Congestion Prediction. *Int. J. Sci. Eng. Res.* **2016**, *7*, 1205–1210.
42. Kukadapwar, S.R.; Parbat, D. Modeling of traffic congestion on urban road network using fuzzy inference system. *J. Am. J. Eng. Res.* **2015**, *4*, 143–148.
43. Mitchell, M. *An Introduction to Genetic Algorithms*; MIT Press: Cambridge, MA, USA, 1998.