PREDICTION OF VEHICULAR TRAFFIC FLOW USING LEVENBERG-MARQUARDT ARTIFICIAL NEURAL NETWORK MODEL: ITALY ROAD TRANSPORTATION SYSTEM

Isaac Oyeyemi Olayode[1,*], Alessandro Severino[2], Tiziana Campisi[3], Lagouge Kwanda Tartibu[1]

[1]Mechanical and Industrial Engineering Technology, University of Johannesburg, Johannesburg, South Africa
[2]Faculty of Civil Engineering and Architecture, University of Catania, Catania, Italy
[3]Faculty of Engineering and Architecture, Kore University of Enna, Italy

*E-mail of corresponding author: olayode89@gmail.com

Resume

In the last decades, the Italian road transport system has been characterized by severe and consistent traffic congestion and in particular Rome is one of the Italian cities most affected by this problem. In this study, a Levenberg-Marquardt (LM) artificial neural network heuristic model was used to predict the traffic flow of non-autonomous vehicles. Traffic datasets were collected using both inductive loop detectors and video cameras as acquisition systems and selecting some parameters including vehicle speed, time of day, traffic volume and number of vehicles. The model showed a training, test and regression value (R^2) of 0.99892, 0.99615 and 0.99714 respectively. The results of this research add to the growing body of literature on traffic flow modelling and help urban planners and traffic managers in terms of the traffic control and the provision of convenient travel routes for pedestrians and motorists.

Article info

Received 5 August 2021
Accepted 7 October 2021
Online 22 December 2021

Keywords:

traffic flow,
traffic congestion,
Levenberg-Marquardt artificial neural network model,
artificial intelligence,
Italy transportation system

Available online: https://doi.org/10.26552/com.C.2022.2.E74-E86

ISSN 1335-4205 (print version)
ISSN 2585-7878 (online version)

1 Introduction

The evolution of private vehicle ownership over the years has put tremendous pressure on road transportation systems worldwide, which has led to an epidemic called the traffic congestion. In addition, road accidents and other road transportation-related deaths are becoming a perennial problem. The significant causes of traffic congestions in urban areas are illustrated in Figure 1 below.

Intelligent transportation systems were introduced to find solutions to this issue, such as traffic flow information management of vehicles and pedestrians. The objective of these solutions was so that the future traffic flow conditions can be adequately predicted and appropriate measures would be put in place to tackle traffic flow problems on freeways and road intersections [1]. Prediction of traffic volume is the fundamental and primary component of intelligent transportation systems; this is essential for urban planners and transportation researchers to better manage traffic control of vehicles in road transportation systems. This will enable the efficient operation of road transportation systems [2]. Traffic volume over the years has been identified as a primary traffic flow variable that is a significant determinant in knowing if traffic congestion will occur or not. It is a complex system related to traffic density, speed of vehicles and length of the road. Its features are internal periodicity and correlation and these features determine how a traffic volume can be measured. The techniques for predicting the traffic volume are categorized in two): conventional mathematical models and computation intelligence. Conventional models focus on mathematical statistics and calculus. Mathematical techniques comprise of exponential smoothing [3], Kalman filtering [4-5] and autoregressive integrated moving average (ARIMA) [6]. Statistical models are well-known by researchers for their effective mathematical theories and innovative insights [7].

A typical example is linked to a time series predictive model, which has an accuracy that depends on the quality of the training datasets. However, the traffic volume has been known to be intricate and non-straightforward when it comes to vehicles on the road, which means it cannot present an orderly state of functions of vehicles on the road, which makes achieving high efficacy difficult. In the last decades, marked by the fourth industrial revolution, machine learning has become widely used in transportation due to its artificial intelligence capabilities. An important branch of machine learning is artificial neural networks. This type of machine learning has unique predictive capabilities because of its capacity to adapt to learning,

Figure 1 *Primary causes of traffic congestion*

associative and memory features and large-scale and distributed processing features [8]. Artificial neural networks consist of backpropagation neural networks [9-14], radial basis function neural networks [15] and fuzzy-neural network [16-18], all these neural networks can approximate non-linear functions and perform excellently depending on the type of datasets for training and testing and the field of applications of the neural network. However, the neural networks comprise randomness when choosing weights and thresholds needed for speed convergence and the neural network results. Some transportation researchers have applied the combination of predictive techniques to apply genetic algorithms for optimization of weights and neural network thresholds. The caveat about this model is that it has an appropriate generalization. However due to the interference from weather and other climatic factors during the collection of traffic data, the data collected is not 100% accurate.

In recent times, the number of vehicles on the road are unknown due to lack of adequate traffic data collection equipment and most urban roads are not constructed adequately because of lack of the proper road design and planning. This has led to severe traffic congestions. Traffic congestions cause road accidents and hidden dangers to pedestrians‘ travel and transportation planning [19]. Other leading causes of traffic congestions are increased travel times, waste of road transportation resources, more greenhouse gas emissions (GHG) [10, 20] and causes severe harm to people's mental and physical health [21]. However, one can apply an in-vehicle communication method or use mobile applications to collect large unfiltered traffic datasets to assist in the traffic volume prediction. The efficiency in predicting the traffic volume will assist drivers in travel routes recommendation, peak-hours avoidance and driver experience optimization [22]. Another advantage of traffic volume prediction is that it offers reference recommendations to urban planners regarding the construction of road infrastructure and road transportation planning. Therefore, there has been growing research interest in predicting the traffic volume in urban areas using historical data. There has been no shortage of literature review on the issue of prediction of traffic data not excluding greenhouse gas emissions from a taxi. Transportation researchers have used Kalman Filter, Hidden Markov and ARIMA models to predict the traffic volume using conventional methods using geographical positional sensors. But, because of the excessive costs of installing roadside sensors, there are limitations to their usage.

With the increase in innovative development of mobile internet, artificial intelligence mobile devices are widely applied and there is the ease in accessing user information [23-26]. Convolutional Neural Networks (CNN) can be applied to capture spatial characteristics and Recursive Neural Networks (RNN) to capture temporal characteristics. Neural networks consist of variants and hybrid models, which are used in the field of traffic flow predictions. There are three types of challenges pre-existing techniques used for predicting traffic volume are experiencing: (1) The problem of single traffic data source from taxi and bicycle. The traffic datasets can only be taken from a minutely small part of an area with different vehicles, causing limited data and inefficient prediction. (2) They usually take into consideration the spatial characteristics such as the length of the freeway or the road intersections and the number of lanes on the road. Application of artificial intelligence over the last few years have been successfully used in various fields of transportation, such as service computing techniques [27], edge computing methods [28] and social networks [29].

This research focuses on modelling the vehicle traffic flow in the Italian transport system using vehicle speed, time, traffic volume and number of vehicles on the road as input parameters. The aim of the research was to demonstrate whether the artificial intelligence techniques, such as the Levenberg-Marquardt-Artificial neural network model, can be used to model the traffic flow of light vehicles (which are commonly used by road users due to their convenience and ease of use). In addition, understanding of the traffic flow patterns using traffic volume was analyzed. This parameter plays an important role in determining whether the traffic congestion will occur on the motorway or highway. The results obtained improve the research in the area of traffic flow modelling using a soft computing technique (Levenberg-Marquardt artificial neural network model (LM-ANNM)). The article was divided into five parts defined respectively by the introduction with an outline of the main theoretical-scientific bases, the second part focused on the study of the Italian traffic flow model. The third part dealt with the methodology, while the fourth part underlined the results obtained and, finally, the conclusions included some considerations for future developments.

2 Methodology

In this section, the location of the study is explained in detail, how the traffic datasets were extracted and different traffic flow variables that were considered.

2.1 Traffic datasets and location of the study

According to the Global traffic scorecard, Rome was ranked in 2019 as one of the top ten most congested cities globally, in terms of vehicular traffic flow. The traffic scorecard even reported that motorists in the most congested city in Italy, identified as Rome, lose over 200 hours stuck in traffic jams. The traffic datasets sources for this study were obtained before COVID-19 in early 2019. The traffic data covered up to 400 million km taking into consideration Highways and non-Highways. The traffic datasets showed that highways in Rome in 2019 experienced 29% in traffic congestion and non-highways (inner roads: unsignalized road intersections) experienced 42% in traffic congestion. It was deduced from the traffic datasets that when motorists travel around Rome highways regularly before 8 am on Tuesday, they will save up to 7 hours per year (i.e. for a 30-minute journey due to low traffic volume during this time period). This is the average daily congestion levels shown in Figure 2, these congestion levels are calculated using the hourly traffic datasets. Note that

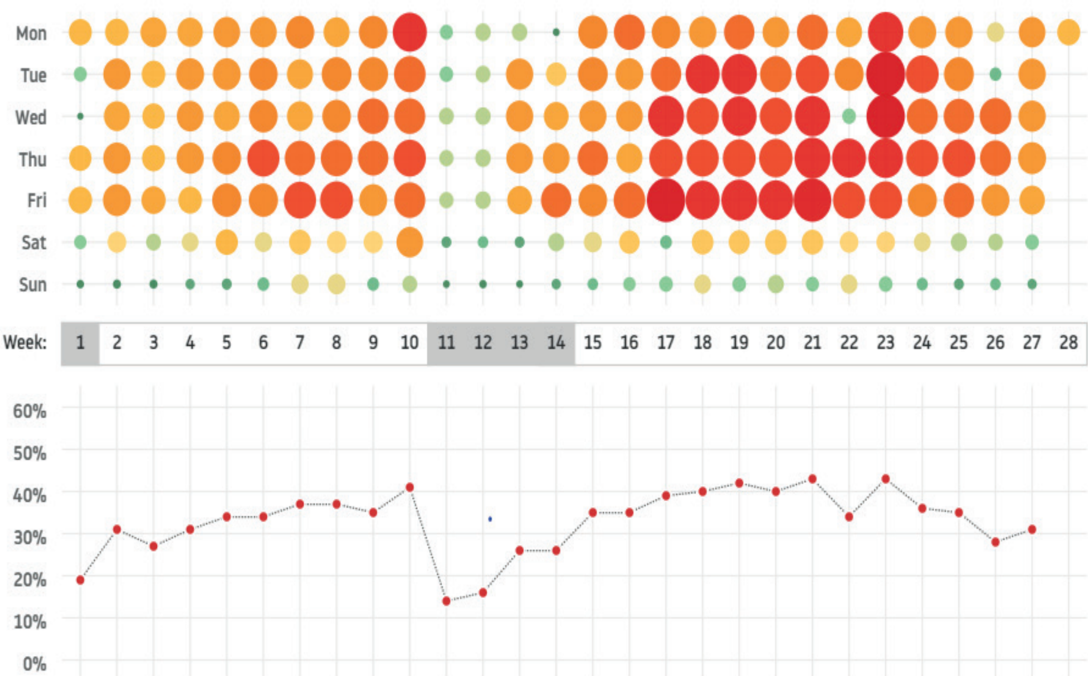

Figure 2 *Italy levels of traffic congestion, [33]*
Note: ***Deep Red****-Heavy congestion,* ***Green****-No congestion,* ***Yellow****-Rising congestion levels*

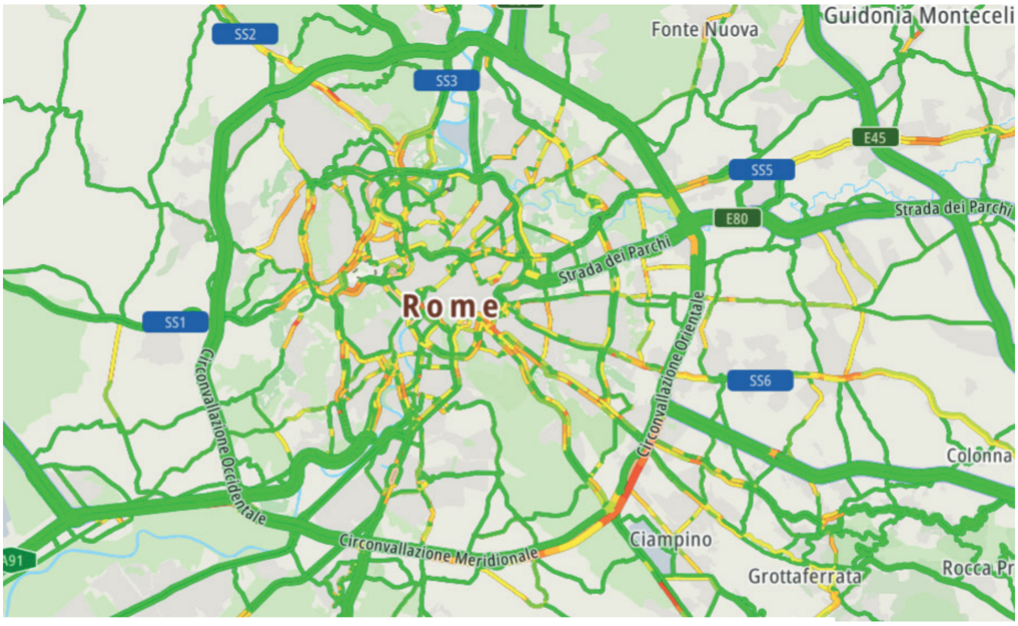

Figure 3 *A graphical illustration of the traffic congestion levels in Rome, Italy [33]*

Table 1 *Characteristics and descriptions of the Highways*

Highways	Dates	Distance (m)	Direction	Number of lanes	Speed limit (Km/h)	Number of Vehicles (Vehicles/h)	Pavement Type
Highway 1	03/02/2019-10/07/2019	14.5	Northbound	2	130	1,097152	Dense Asphalt
Highway 2	03/02/2019-10/07/2019	10.40	Northbound	2	130	12,240260	Dense Asphalt
Highway 3	03/02/2019-10/07/2019	9.0	Northbound	3	130	13,448023	Dense Asphalt
Highway 4	03/02/2019-10/07/2019	5.70	Northbound	2	130	15,051124	Dense Asphalt
Highway 5	03/02/2019-10/07/2019	6.30	Northbound	3	130	14,262048	Dense Asphalt

each week starts on Monday and ends on Sunday. The figure below shows the different levels of congestion that exist in the Italy Road Transportation System.

Traffic Data were collected from five highways in Rome connecting to the Rome highway (Rome is one of the most traffic congested cities in Europe regarding road transportation connectivity and vehicular mobility). These five highways have different distances and different number of lanes. It is only highway 3 and 5 that have three lanes. These highways have a speed limit of 130 km/h and determination of the number of vehicles on these highways is per hour- see Table 1. A graphical illustration of the traffic congestion levels in the Rome transportation system is shown in Figure 2. The Traffic data used in this research were obtained by using the sophisticated data collection equipment such as video cameras and inductive loop detectors. This equipment was applied at the five highways for a period of seven days from Monday to Sunday between a 24-hours period with a 5-hour determination of on-peak and off-peak hours. Figure 3 illustrates the geographical google image of Rome Transportation network, with green colour signifying free flowing traffic, yellow indicating gradual decrease in free-flowing vehicular traffic flow and red signifying traffic jam or congestion.

Table 1 describes the distance, direction, number of lanes and number of vehicles on the highways in which the traffic data was collected. It also shows the speed limit of each highway and the type of pavement on the highways.

2.2 Traffic data extraction

The traffic flow parameters applied in this research are extracted from the Italian Road Transportation System, this is shown with the aid of a flowchart in Figure 4. The traffic density is the overall number of vehicles on the road at a particular time of the day, it is dependent on the traffic volume of vehicles, because

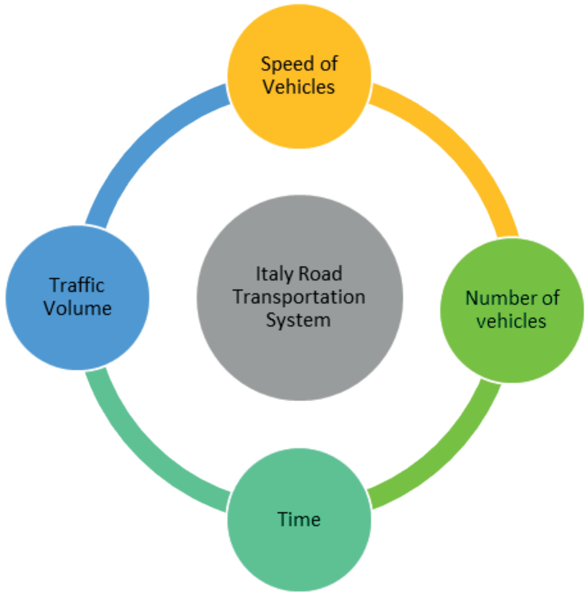

Figure 4 *Traffic flow variables*

Table 2 *Division of the period of the day into Off-peak and On-peak hours*

Time Classification	Time of day	Off-peak/On-peak period
1	00:00:00 am - 04:59:59 am	Off-peak hours
2	05:00:00 am - 09:59:59 am	On-peak hours
3	10:00:00 am - 14:59:59 pm	On-peak hours
4	15:00:00 pm - 19:59:59 pm	On-peak hours
5	20:00:00 pm - 23:59:59 pm	Of-peak hours

the higher the traffic density of vehicles on the road, the higher the traffic volume resulting in traffic congestion. The traffic data collection equipment used for the traffic data collection such as video cameras were installed on the side of each highway and were installed in such a way that the recording of the video camera of each vehicle passing the highway does not interfere with the traffic flow on each highway and the efficiency of the video cameras were tested to make sure that they were able to record any vehicular activities on the highways. The research was carried out based on an assumption that the traffic flow is fluid-like in nature and the directions of the vehicles are Northbound. As the video cameras installed in each highway was recording, these recordings have been displayed and replayed on a screen monitor at the Italian Ministry of Transportation to determine which of the traffic data is applicable in the traffic flow prediction. The traffic data gotten from the traffic data collection equipment were analyzed in the Microsoft excel environment using the excel analytical tools. The analysis of the traffic data in the excel sheet were carried out by dividing the overall number of vehicles for the highways with the time it takes them to travel all the highways. An accuracy of 0.03 seconds was adjudged to be the time accuracy it takes each vehicle to travel the length of the highways. Five hundred and twenty-six (526) traffic datasets were collected between seven days of the week (3rd February to 9th February 2019) including weekends (Saturday and Sunday). However, because of a significant amount of raw and unfiltered traffic data, the research focused on the traffic data collected within seven days. These traffic data were divided, based on the 24-hour period in which they are collected, by applying a 5-hour period to determine the off-peak and on-peak periods.

The classification of different periods of the day is shown in Table 2 illustrating the On-peak and Off-peak hours during a 24-hour period. The traffic datasets comprise of the number of vehicles, time, traffic volume and speed of vehicles. The Traffic volume datasets used in this research study were collected through manually by counting the number of vehicles crossing a fixed section of each of the five highways using the video camera. The average time taken by each vehicle type to travel the length of the highway was measured using the time displayed on the video camera with an accuracy of 0.003 s. The vehicular speed was determined by Speed by measuring the time taken by a vehicle to travel the length of the highway. Entry and exit times of vehicles at the beginning and endpoints of the highway were tabulated as in-time (T_1) and out-time (T_2). Each vehicle category's speed in both directions was calculated by operating on time difference values ($T_1 - T_2$). About 526 traffic datasets were generated from the five highways considering the off-peak and the on-peak hours. The extracted traffic data used in this research were entered

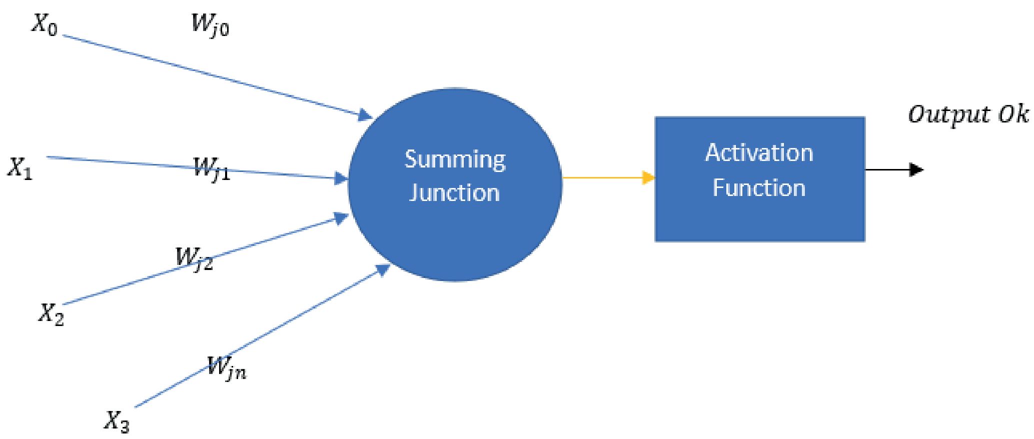

Figure 5 *The non-linear model of a neuron*

into the Microsoft Excel environment for additional evaluation and analysis.

2.3 Artificial neural network

Artificial neural network models are widely known for their application in the field of transportation. It has a long application history in the mining of datasets for modelling difficult relationships among datasets. Its merit over conventional statistical models is its ability to interpret the non-linear relationships between the input and output datasets even if there is no non-linear relationship between the datasets. However, considering other data mining methods, artificial neural network model exhibits optimal performance when applied on both categorical and continuous datasets. The artificial neural network model has been created in a way that it is analogous to the neural system of human beings. This means that ANN trains datasets within their hidden layer like the way the human being's neural system functions. Previous literature reviews have discussed the ANN methods, especially the multilayer perception networks, spectral-based neural networks, radial basis function networks, resource allocating networks, recurrent networks and wavelet networks. This study used a multilayer feedforward neural network to model traffic flow using the traffic datasets from the Italian transportation system. A multilayer feedforward neural network comprises three significant layers, namely: input, hidden and output. The significant feature about these layers is that a node denotes each independent variable. These nodes are interconnected to other nodes in subsequent layers and each interconnection is represented by the weight-related to it, this is shown in Figure 5. The weights reflect the strength of the connections among the neurons are updated regularly [30]. A backpropagation method is mainly used to estimate an appropriate set of weights in the neural network. According to this method, prediction values are updated in each epoch [31-32]. The neural network error undergoes successive iterations to decrease errors as much as possible; this is called Artificial neural network model training. The ANN model training applies a statistical optimization method to reduce the errors. However, due to the process nonlinearity in the ANN, there is a non-open form remedy to the problem of minimization. For the calculation of a weighted sum in the ANN model, there is a need for a linear mathematical equation to be created in conjunction with the independent variables of datasets and their related weights. The ANN model's hidden layer receives the weighted sum by applying a transfer function that functions by processing the sum in a layer. In an ANN network, there are various types of transfer functionalities, namely log-sigmoid and tan-sigmoid; these transfer function can be used for the training and testing of datasets. During the process, the output layer of the ANN model accepts the new weighted sum via the hidden neurons and applying another transfer function, the output layer mathematically calculates the sum and the transfer function process continues. The ANN model toolbox of the MATLAB 2020a was applied to train and validate the ANN model used for this research study. Figure 6 shows the structure of the used LM-ANNM model.

Furthermore, different types of algorithms in an artificial neural network can be applied in solving the problem of minimization in the ANN model. These algorithms are Levenberg-Marquardt (LM), Bayesian Regularization and Scaled Conjugate Gradient. In this research study, the Levenberg-Marquardt algorithm was applied because of its accuracy features in terms of computationally and the ability to achieve an efficient result compared to Bayesian Regularization and Scaled Conjugate Gradient algorithms. To achieve an appropriate structure for the final estimation problems in a neural network, the Log-sigmoid transfer function (Equation (2)) and hyperbolic tangent sigmoid (Equation (3)) were used in the hidden layer of the ANN architecture of this research study. The Mean squared error (MSE) was used in achieving an optimal training performance of the ANN model, which is explained by equations below:

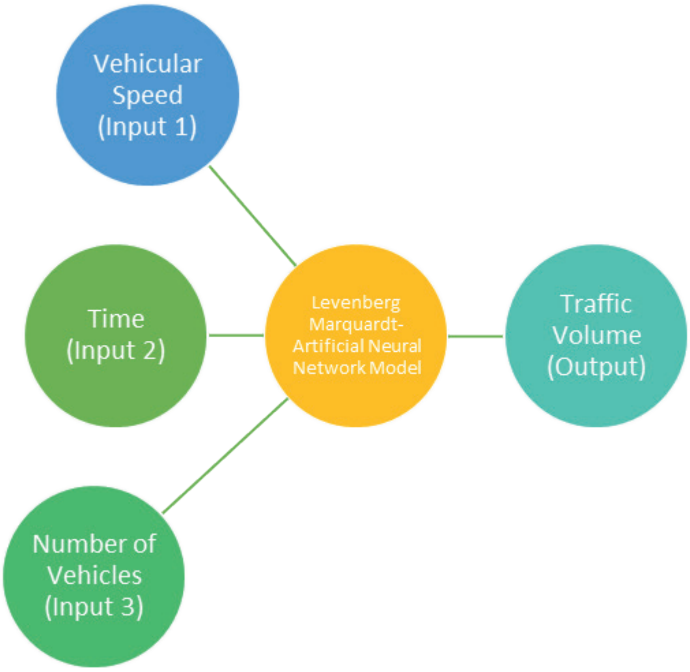

Figure 6 *Structure of the Levenberg-Marquardt Artificial Neural Network Model Inputs and Output*

$$f(x) = \frac{2}{(1+e^{-2x})-1}, \tag{1}$$

$$f(x) = \frac{1}{1+e^{-x}}. \tag{2}$$

The input layers of the ANN model were connected with the hidden layers, which were then connected to the output layers. However, layers in the neural network accept constant input parameters from the bias. The neural network input to a unit j is mathematically expressed as:

$$X_j = \sum_j W_{IJ} X_J + b_i, \tag{3}$$

where:
X_i is the output from the preceding layer,
W_{ij} s denotes the connection weight between layer i and j and b_i is bias.

The artificial neural network weights were selected randomly and initialized throughout the network model adhering to a range before the start of the ANN training process. For each mathematical calculation of input pattern (X_{ij}), the analysis of output at the output units will be dependent on the output used in the hidden units in Equation (4). The output vector (Y_{ij}) was compared to the chosen (i.e., target values) input vector (X_{ij}). The error function parameters of the MSE were selected as the difference between X_{ij} and Y_{ij}, This as:

$$MSE = \frac{1}{N}\sum_{i=1}^{n}((X_{ij} - Y_{ij}))^2. \tag{4}$$

3 Results and discussions

3.1 Italy vehicular traffic flow

The vehicular traffic flow of one of the four highways is explained in Figures 9-15, representing the vehicular traffic flow from Monday (03/02/2019) to Sunday (09/02/2019). The vehicular traffic flow system in Italy is reliant on the traffic volume of vehicles taking into consideration the time of the day. Figure 7 to 13 shown below is used for the explanation of the on-peak and the off-peak hours. To buttress this explanation, let's use highway 1 as an example, the period between 1 (00:00:00- 04:59:59), the traffic volume gradually increases. This is a period between midnight and the opening hours of places of work in Italy. The traffic volume of vehicles in the Italy road transportation systems is usually high between 2(05:00:00-09:59:59) and 4 (15:00:00-19:59:59), these periods of the day are often when it is the on-peak period. These periods of the day are when the traffic volume of vehicles is at the maximum on the highway. The most striking observation about this traffic flow illustrations (Figures 7 to 13) is that weekends produce low traffic volume of vehicles especially between 05:00:00-09:59:59 because it is a non-working day for most Italians. However, during the weekends, the traffic volume of vehicles gradually increases in the afternoon 4 (15:00:00 and 19:59:59). Another important observation about the traffic flow of vehicles in Italy is that the higher the traffic flow of vehicles on a highway it will resulted in a high traffic density. Conclusively, the Italian Vehicular traffic flow

Highway 1

Model Development

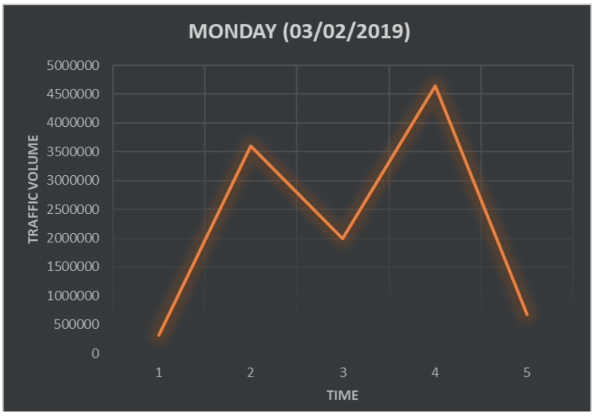

Figure 7 *Day 1*

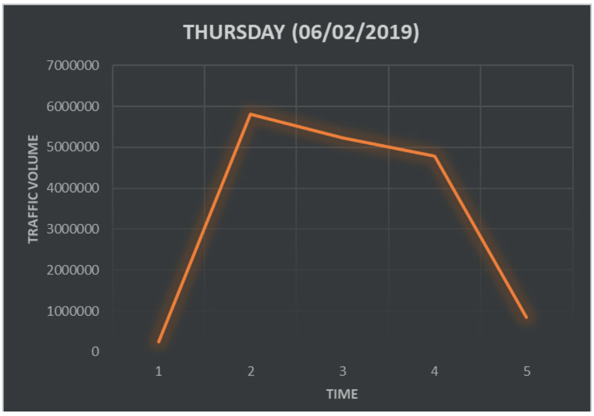

Figure 10 *Day 4*

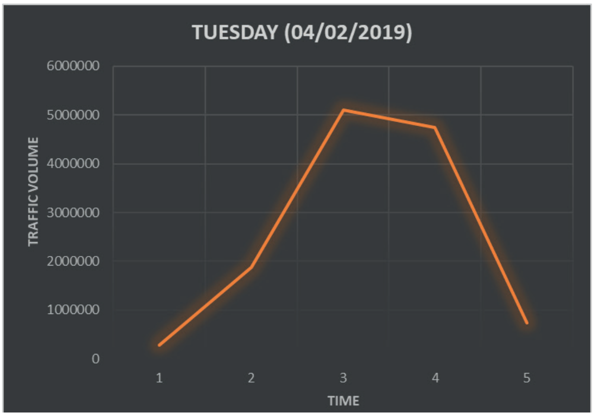

Figure 8 *Day 2*

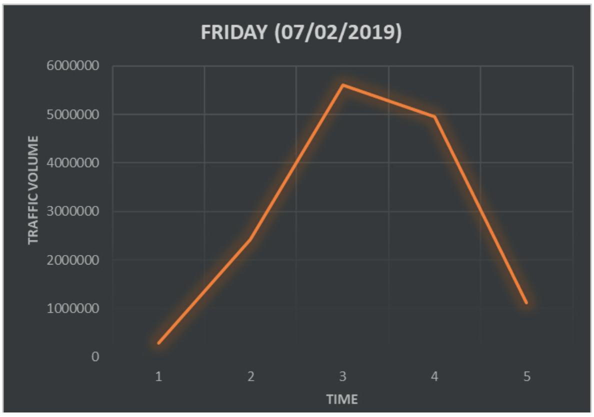

Figure 11 *Day 5*

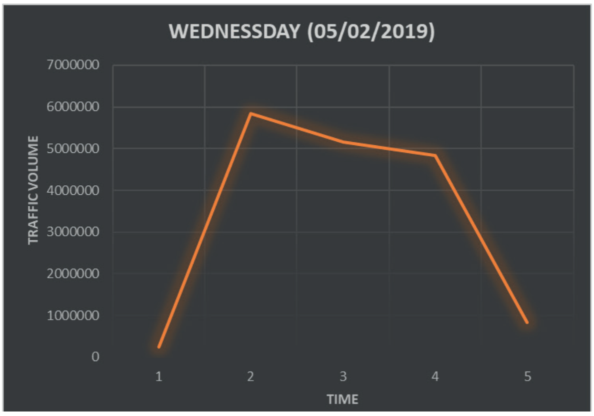

Figure 9 *Day 3*

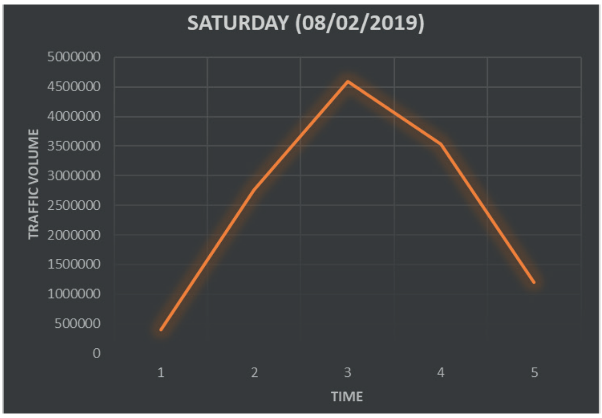

Figure 12 *Day 6*

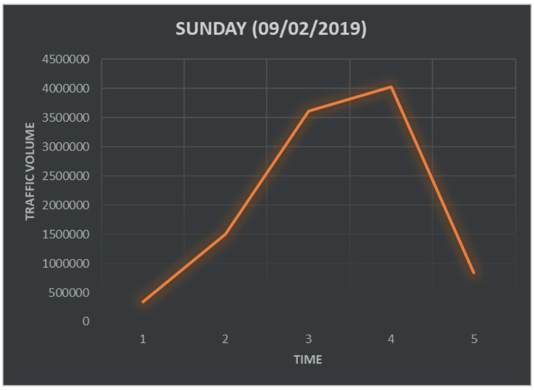

Figure 13 *Day 7*

Table 3 *The artificial neural network (ANN) of the traffic datasets with different architectures MSE = Mean Square Error; V = Validation; T_1= Training; T_2 = Testing; R = Regression*

	Train 1	Train 2	Train 3	Train 4	Train 5	Train 6
Number of hidden layers	1	1	1	1	1	1
Number of hidden neurons	5	6	7	8	9	10
Transfer function	Sigmoid	Sigmoid	Sigmoid	Sigmoid	Sigmoid	Sigmoid
Learning algorithm	Levenberg Marquardt	Levenberg Marquardt	Levenberg Marquardt	Levenberg Marquardt	Levenberg Marquardt	Levenberg Marquardt
Number of epochs	143	4	2	2	3	12
R(T_1)	0.25429	0.94936	-0.24814	0.0229067	0.823872	0.98919
R(V)	0.24567	0.87512	0.022083	0.289241	0.830123	0.995484
R(T_2)	0.18974	0.89143	-0.26913	0.00902591	0.791652	0.996148
R(All)	0.233306	0.92914	-0.21366	0.049181	0.81726	0.99714

has shown that having an insight into the traffic flow of vehicles on a particular highway, it is imperative to have a fundamental understanding on the traffic volume and density and the period of the day.

Artificial neural network models possess many features such as flexibility, efficient parallelism and generalization capacity. The interrelationships between variables are created automatically and fitting is created naturally in the neural network architecture. The appropriate number of inputs or outputs will be chosen depending on the Levenberg-Marquardt Neural network modelling. However, there is no pre-existing technique for creation of an ANN model architecture, a trial-and-error method is always used most of the time. Initiatives can be applied in the designing of the overall ANN structure. Difficulties in the ANN only arise in terms of high order heterogeneity and the problem we want to model with ANN. Primary characteristics such as ANN structure, input and output variables, transfer function activation and choosing a training algorithm are important when designing an artificial neural network model. The Levenberg Marquardt algorithm was used in this study due to its efficiency when achieving optimal results as compared to scaled conjugate gradient and Bayesian regularization. In this research, a Multilayer Perceptron (MLP) network was applied on the traffic flow datasets from the Italian transportation system. For the generation of the LM-ANN model, the number of vehicles, time of day, traffic volume, the corresponding average speed of vehicles were inserted in the required format in the MATLAB environment. Consequently, the five hundred and twenty-six traffic data gotten from the Italy transportation system were divided into 368, 85, 73 for the LM-ANN model training, cross-validation and testing, respectively.

Network Inputs: Number of Vehicles, time and Speed of vehicles.

Network Outputs: Traffic Volume

Training Algorithm: Levenberg Marquardt

Five hundred and twenty-six (526) traffic data were used for the LM-ANNM training, testing and validation. The traffic data were divided into 7% (368), 15% (85) and 15% (73) appropriately for both the training, testing and validation of the model, respectively. All these were carried out in the MATLAB environment. In Table 3, train 6 was adjudged to be the LM-ANNM optimal performance and Figure 14 shows the results of the training, testing and validation results of train 6. The overall optimal validation performance of the LM-ANNM is shown in Figure 15. A regression value of R^2 of 0.99714 (3-10-1) is presented. These results clearly show that the traffic data's inputs and outputs are well correlated.

Figure 15 presents the overview graph of the gradient epoch and validation performance of the ANN model. Table 3 shows the results obtained from the analysis of the traffic performance evaluation of R^2 parameters for training, testing and validation of the LM-ANNM using various architectures. According to the results, it was discovered that the best optimum model training and testing performance was achieved when the number of hidden neurons is 10. It is indicative from Table 3 that the LM-ANNM parameters, the number of hidden neurons and of epochs is essential in determining the performance of the traffic data collected from the Italian road transportation system. Another significant observation of the result of the research occurs when an LM-ANNM achieves the maximum correlation coefficient (R^2) and the minimum MSE, it can be said that the ANN model is capable of modelling. The current study found that the LM-ANNM is of optimum performance and simulation can be used for entering new inputs. Another significant finding is that the testing regression value (R-value = 0.99615)

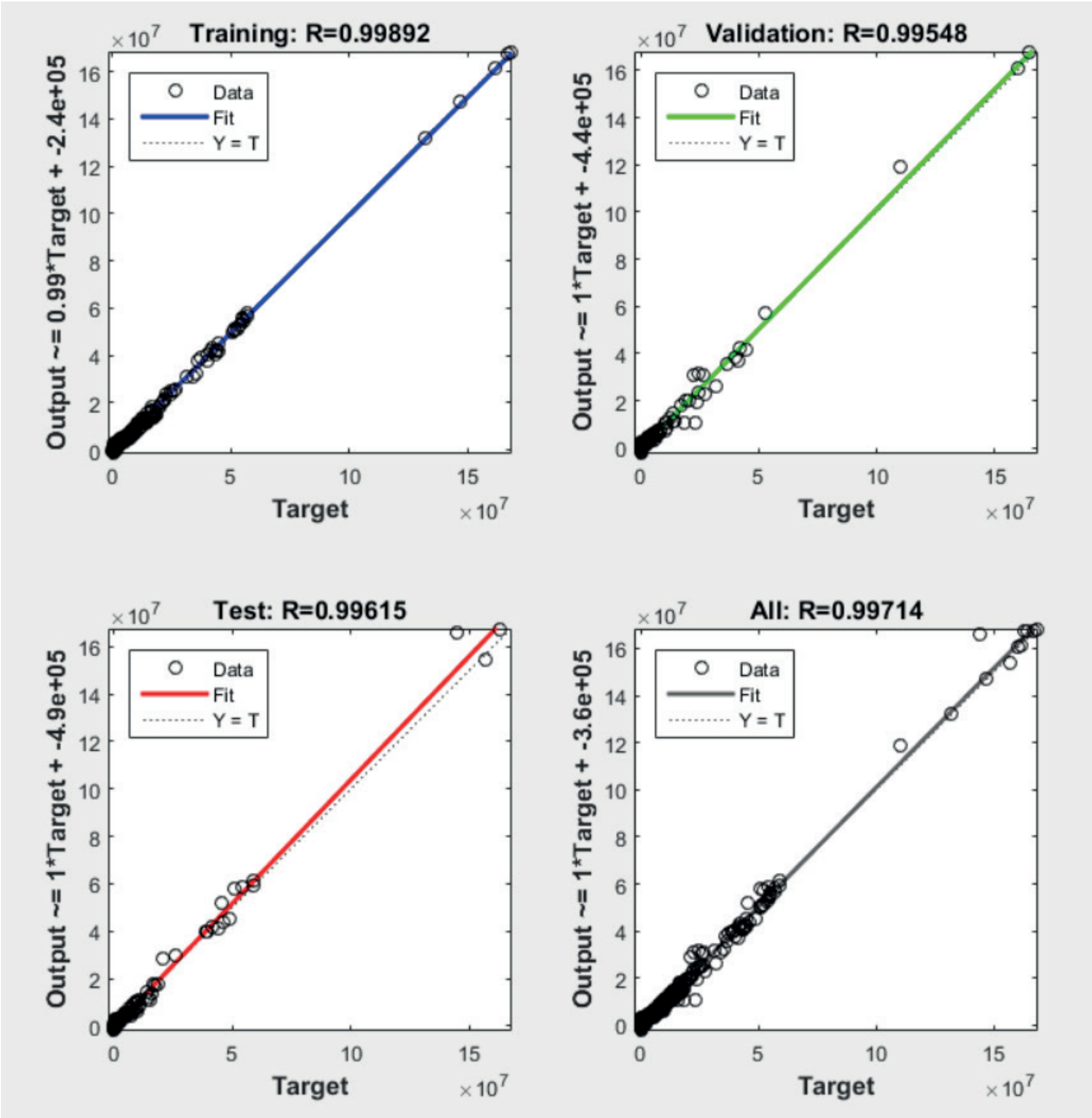

Figure 14 *Train 6*

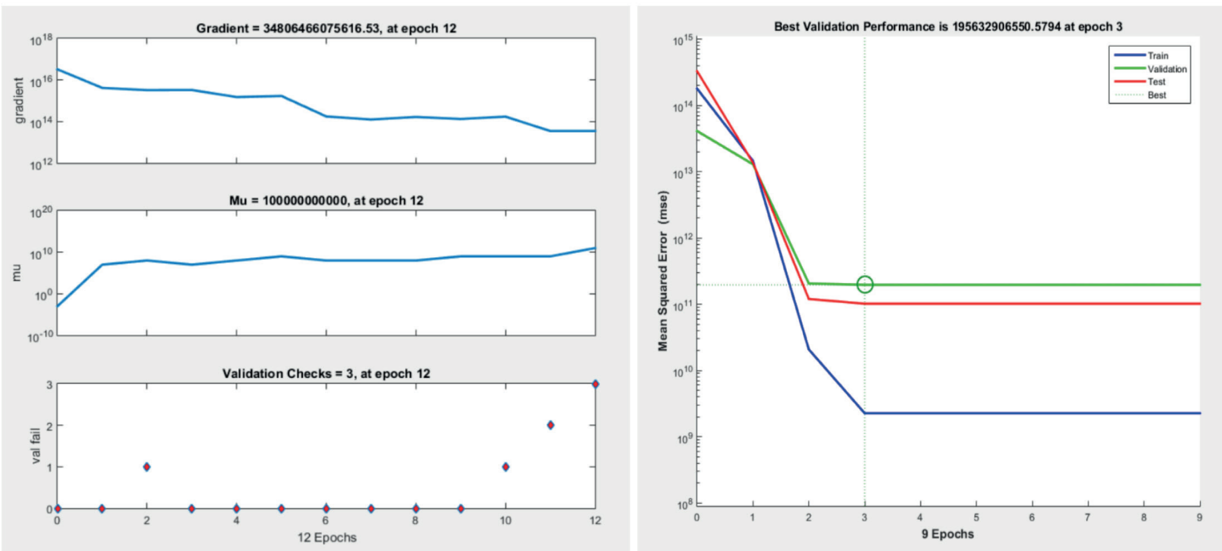

Figure 15 *Gradient epoch and validation performance of the ANN model*

indicates that the inputs and target are well correlated. These results show that when the Regression value is closer to one, there is an efficient linear relevance between the traffic data inputs and a target.

4 Conclusions

The study focused on application of the Levenberg-Marquardt neural network model to model the flow of vehicular traffic in the Italian transport system, with particular reference to the city of Rome, using some of the parameters obtained from the survey instruments. The evidence of this research suggests the following:

- The back propagation neural network analyzes the prediction error and propagates to the following layers to modify the weights, leading to an efficient accuracy of prediction training.
- The training of the ANN model can be stopped after a specific number of epochs if there is no further improvement in prediction accuracy.
- One of the most significant findings from this research is that the Levenberg-Marquardt neural network model (LM-ANNM) has emerged as a reliable predictive model of the vehicular traffic flow of the Italian transport system. This study found that, in general, the Levenberg-Marquardt neural network model showed a superior understanding of traffic flows obtained from the Italian road transport network with a test performance of R^2 0.99714.
- The second significant finding was that vehicle speed, traffic volume, time and number of vehicles are integral to e understanding of light vehicle traffic flow.
- The evidence from this research study suggests that the Italian vehicle traffic flow system will be useful for advanced traveler information systems, short- and long-term traffic flow forecasts.
- This research contributes to the existing knowledge of the traffic flow modelling, providing ways in which traffic datasets can be used and providing further evidence on how the Levenberg-Marquardt neural network model can be used in modelling the non-autonomous vehicles on highways. The results of this research provide the following insights for the future research:
- It would be interesting to evaluate the effects of traffic accidents on vehicular traffic flow using the soft computing techniques.
- Further research regarding the role of traffic volume, traffic density and other different classes of vehicles in modelling the traffic flow would be useful and would contribute to the knowledge of road transport.

Limitations of the present research are related to the case study examined and non-replicability. In particular, the traffic datasets used for this study were limited to a small number of motorways (four motorways) in Italy and the traffic data were collected on specific days. Another limitation is that the current research was not explicitly designed to evaluate the factors influencing the collection of traffic data; factors such as weather conditions and extreme road conditions (such as road accidents) were not considered.

Data availability statement

Some or all data, models, or code generated or used during the study are available from the corresponding authors by request.

Acknowledgements

The authors would like to thank the Italy Ministry of Transportation for making their Traffic data available for this study.

References

[1] FUSCO, G., COLOMBARONI, C., COMELLI, L., ISAENKO, N. Short-term traffic predictions on large urban traffic networks: applications of network-based machine learning models and dynamic traffic assignment models. In: 2015 International Conference on Models and Technologies for Intelligent Transportation Systems MT-ITS: proceedings [online]. IEEE. 2015. p. 93-101. Available from: https://doi.org/10.1109/MTITS.2015.7223242

[2] LIU, L., JIA, N., LIN, L., HE, Z. A cohesion-based heuristic feature selection for short-term traffic forecasting. *IEEE Access* [online]. 2018, **7**, p. 3383-3389. ISSN 2169-3536. Available from: https://doi.org/10.1109/ACCESS.2018.2889814

[3] CHAN, K. Y., DILLON, T. S., SINGH, J., CHANG, E. Neural-network-based models for short-term traffic flow forecasting using a hybrid exponential smoothing and Levenberg-Marquardt algorithm. *IEEE Transactions on Intelligent Transportation Systems* [online]. 2011, **13**(2), p. 644-654. ISSN 1524-9050, eISSN 1558-0016. Available from: https://doi.org/10.1109/TITS.2011.2174051

[4] HE, S., HU, C., SONG, G.-J., XIE, K.-Q., SUN, Y.-Z. Real-time short-term traffic flow forecasting based on process neural network. In: International Symposium on Neural Networks: proceedings. Springer. 2008. p. 560-569.

[5] WANG, Y., PAPAGEORGIOU, M., MESSMER, A. A real-time freeway network traffic surveillance tool. *IEEE Transactions on Control Systems Technology* [online]. 2005, **14**(1), p. 18-32. ISSN 1063-6536, eISSN 1558-0865. Available from: https://doi.org/10.1109/TCST.2005.859636

[6] HOU, Q., LENG, J., MA, G., LIU, W., CHENG, Y. An adaptive hybrid model for short-term urban traffic flow prediction. *Physica A: Statistical Mechanics and its Applications* [online]. 2019, **527**, p. 121065. ISSN 0378-4371. Available from: https://doi.org/10.1016/j.physa.2019.121065

[7] KARLAFTIS, M. G., VLAHOGIANNI, E. I. Statistical methods versus neural networks in transportation research: differences, similarities and some insights. *Transportation Research Part C: Emerging Technologies* [online]. 2011, **19**(3), p. 387-399. ISSN 0968-090X. Available from: https://doi.org/10.1016/j.trc.2010.10.004

[8] YANG, H.-J., HU, X. Wavelet neural network with improved genetic algorithm for traffic flow time series prediction. *Optik* [online]. 2016, **127**(19), p. 8103-8110. ISSN 0030-4026. Available from: https://doi.org/10.1016/j.ijleo.2016.06.017

[9] OLAYODE, I. O., TARTIBU, L. K., OKWU, M. O. Application of artificial intelligence in traffic control system of non-autonomous vehicles at signalized road intersection. *Procedia CIRP* [online]. 2020, **91**, p. 194-200. ISSN 2212-8271. Available from: https://doi.org/10.1016/j.procir.2020.02.167

[10] OLAYODE, I. O., TARTIBU, L. K., OKWU, M. O., UCHECHI, U. F. Intelligent transportation systems, un-signalized road intersections and traffic congestion in Johannesburg: a systematic review. *Procedia CIRP* [online]. 2020, **91**, p. 844-850. ISSN 2212-8271. Available from: https://doi.org/10.1016/j.procir.2020.04.137

[11] 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: 2021 IEEE 12th International Conference on Mechanical and Intelligent Manufacturing Technologies ICMIMT: proceedings. IEEE. 2021. p. 287-292.

[12] YANG, M., ZHANG, X. A novel travel adviser based on improved back-propagation neural network. In: 2016 7th International Conference on Intelligent Systems, Modelling and Simulation ISMS: proceedings. IEEE. 2016. p. 283-288.

[13] ZHANG, Z.-G., YIN, J.-C., WANG, N.-N., HUI, Z.-G. Vessel traffic flow analysis and prediction by an improved PSO-BP mechanism based on AIS data. *Evolving Systems* [online]. 2019, **10**(3), p. 397-407. ISSN 1868-6478, eISSN 1868-6486. Available from: https://doi.org/10.1007/s12530-018-9243-y

[14] ZHAO, X., HAN, M., DING, L., CALIN, A. C. Forecasting carbon dioxide emissions based on a hybrid of mixed data sampling regression model and back propagation neural network in the USA. *Environmental Science Pollution Research* [online]. 2018, **25**(3), p. 2899-2910. ISSN 0944-1344, eISSN 1614-7499. Available from: https://doi.org/10.1007/s11356-017-0642-6

[15] CHEN, D. Research on traffic flow prediction in the big data environment based on the improved RBF neural network. *IEEE Transactions on Industrial Informatics* [online]. 2017, **13**(4), p. 2000-2008. ISSN 1551-3203, eISSN 1941-0050. Available from: https://doi.org/10.1109/TII.2017.2682855

[16] ZHANG, J., NI, L., YAO, J., WANG, W., TANG, Z. Adaptive bare bones particle swarm inspired by cloud model. *IEICE Transactions on Information Systems* [online]. 2011, **94**(8), p. 1527-1538. ISSN 0916-8532, eISSN 1745-1361. Available from: https://doi.org/10.1587/transinf.E94.D.1527

[17] TANG, J., LIU, F., ZOU, Y., ZHANG, W., WANG, Y. An improved fuzzy neural network for traffic speed prediction considering periodic characteristic. *IEEE Transactions on Intelligent Transportation Systems* [online]. 2017, **18**(9), p. 2340-2350. ISSN 0916-8532, eISSN 1745-1361. Available from: https://doi.org/10.1109/TITS.2016.2643005

[18] TANG, J., ZOU, Y., ASH, J., ZHANG, S., LIU, F., WANG, Y. Travel time estimation using freeway point detector data based on evolving fuzzy neural inference system. *PloS One* [online]. 2016, **11**(2), e0147263. eISSN 1932-6203. Available from: https://doi.org/10.1371/journal.pone.0147263

[19] SHAMSHER, R., ABDULLAH, M. N. Traffic congestion in Bangladesh-causes and solutions: a study of Chittagong metropolitan city. *Asian Business Review* [online]. 2013, **2**(1), p. 13-18. ISSN 2304-2613, eISSN 2305-8730. Available from: https://doi.org/10.18034/abr.v2i1.116

[20] SHANG, J., ZHENG, Y., TONG, W., CHANG, E., YU, Y. Inferring gas consumption and pollution emission of vehicles throughout a city. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD´14: proceedings [online]. 2014. ISBN 978-1-4503-2956-9, p. 1027-1036. Available from: https://doi.org/10.1145/2623330.2623653

[21] KUANG, L., YAN, H., ZHU, Y., TU, S., FAN, X. Predicting duration of traffic accidents based on cost-sensitive Bayesian network and weighted K-nearest neighbor. *Journal of Intelligent Transportation Systems* [online]. 2019, **23**(2), p. 161-174. ISSN 1547-2450, eISSN 1547-2442. Available from: https://doi.org/10.1080/15472450.2018.1536978

[22] KUANG, L., YAN, X., TAN, X., LI, S., YANG, X. Predicting taxi demand based on 3D convolutional neural network and multi-task learning. *Remote Sensing* [online]. 2019, **11**(11), 1265. eISSN 2072-4292. Available from: https://doi.org/10.3390/rs11111265

[23] GAO, H., ZHANG, K., YANG, J., WU, F., LIU, H. Applying improved particle swarm optimization for dynamic service composition focusing on quality of service evaluations under hybrid networks. *International Journal of Distributed Sensor Networks* [online]. 2018, **14**(2), 1550147718761583. ISSN 1550-1477, eISSN 1550-1477. Available from: https://doi.org/10.1177/1550147718761583

[24] YIN, Y., CHEN, L., WAN, J. Location-aware service recommendation with enhanced probabilistic matrix factorization. *IEEE Access* [online]. 2018, **6**, p. 62815-62825. ISSN 2169-3536. Available from: 10.1109/ACCESS.2018.2877137

[25] YIN, Y., CHEN, L., XU, Y., WAN, J., ZHANG, H., MAI, Z. QoS prediction for service recommendation with deep feature learning in edge computing environment. *Mobile Networks Applications* [online]. 2020, **2**, p. 391-401. ISSN 1383-469X, eISSN 1572-8153. Available from: https://doi.org/10.1007/s11036-019-01241-7

[26] YIN, Y., AIHUA, S., MIN, G., YUESHEN, X., SHUOPING, W. QoS prediction for Web service recommendation with network location-aware neighbor selection. *International Journal of Software Engineering Knowledge Engineering* [online]. 2016, **26**(4), p. 611-632. ISSN 0218-1940, eISSN 1793-6403. Available from: https://doi.org/10.1142/S0218194016400040

[27] GAO, H., HUANG, W., YANG, X., DUAN, Y., YIN, Y. Toward service selection for workflow reconfiguration: an interface-based computing solution. *Future Generation Computer Systems* [online]. 2018, **87**, p. 298-311. ISSN 0167-739X. Available from: https://doi.org/10.1016/j.future.2018.04.064

[28] CHEN, Y., DENG, S., MA, H., YIN, J. Deploying data-intensive applications with multiple services components on edge. *Mobile Networks Applications* [online]. 2020, **25**(2), p. 426-441. ISSN 1383-469X, eISSN 1572-8153. Available from: https://doi.org/10.1007/s11036-019-01245-3

[29] DENG, S., HUANG, L., XU, G., WU, X., WU, Z. On deep learning for trust-aware recommendations in social networks. *IEEE Transactions on Neural Networks Learning Systems* [online]. 2016, **28**(5), p. 1164-1177. Available from: https://doi.org/10.1109/TNNLS.2016.2514368

[30] RAJU, M. M., SRIVASTAVA, R., BISHT, D., SHARMA, H., KUMAR, A. Development of artificial neural-network-based models for the simulation of spring discharge. *Advances in Artificial Intelligence* [online]. 2011, **2011**, 686258. eISSN 2582-9793. Available from: https://doi.org/10.1155/2011/686258

[31] OLAYODE, I. O., TARTIBU, L. K., OKWU, M. O. Application of Fuzzy Mamdani Model for effective prediction of traffic flow of vehicles at signalized road intersections. In: 2021 IEEE 12th International Conference on Mechanical and Intelligent Manufacturing Technologies ICMIMT: proceedings [online]. IEEE. 2021. p. 219-224. Available from: https://doi.org/10.1109/ICMIMT52186.2021.9476201

[32] 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. *Applied Sciences* [online]. 2021, **11**(18), 8387. eISSN 2076-3417. Available from: https://doi.org/10.3390/app11188387

[33] TOMTOM INDEX. https://www.tomtom.com/en_gb/traffic-index/. Accessed on 23rd July 2021