Individual Vehicle Speed Modeling on Urban Arterials in Mixed Traffic Conditions by Using Artificial Neural Network

Rangwala Ibrahim[1], Dr. L. B. Zala[2], A. A. Amin[3]

[1]*Student of Final Year, M. Tech (Transportation System Engineering), B.V.M. Engineering College, Vallabh Vidyanagar, Gujarat, India, ibhurang786@gmail.com*
[2]*Professor and Head, Civil Engineering Dept., B.V.M. Engineering College, Vallabh Vidyanagar, Gujarat, India,*
[3]*Assistant Professor, Civil Engineering Dept., B.V.M. Engineering College, Vallabh Vidyanagar, Gujarat, India,*

Abstract - *The traffic pattern in India is heterogeneous in nature so, development of speed model is a challenging task. This research paper focus on preparation of speed predicting model for individual vehicle category in the mixed traffic conditions on an urban arterial road. The field data collected from the selected stretch on S.G. Highway, Ahmedabad by manual as well as videography technique. The observed vehicles in the traffic flow were divided in different categories. The collected data were interpreted and analysed in MATLAB R2017a software using ANN tool by formulating a relationship between speed of a vehicle type to the volumes of all categories of vehicles. The developed ANN models were compared to regression-based models to determine the effectiveness of the ANN models. It has been found from the present study that the volume-based ANN models shows much greater correlations between input and output variables as compared to density and volume-based regression models. The average values of R for the ANN models were found out about 0.90 compared to regression models which estimated values of R between the range of 0.55-0.67*

Key Words: Speed model, Mixed traffic, Urban arterial, Artificial neural network, MATLAB R2017a, Regression based models, Coefficient of correlation R

1. INTRODUCTION

Speed is the most essential component that affects the performance of the traffic stream and one of the fundamental component which relates to traffic flow theory other than density and volume. Speed has also been recognized as one of the measures that designers can use to examine road consistency and driver expectancy on roadways. Highway Capacity Manual [5] used free flow speed as a measure for defining level of service on a road [1]. The speed data collection in the field is difficult, often infeasible and collected data cannot fulfill designer's requirements every time. In India all categories of vehicles are plying on the road which proves difficult to develop a speed model due to heterogeneous nature of traffic stream. All these issues related in speed data collection has inspired researchers to develop mathematical models for predicting speed. For speed modeling various methodologies are used however artificial neural networks (ANNs) are relatively new computational programs that have found large-scale utilization regard to theses speed modeling concepts [7].

In earlier studies, a multi-regime speed model has been developed and the influences on the traffic volume and its composition on vehicular speed on six lane divided urban arterials [9], in this study it is assumed that the speed is having linear relationships with classified volumes within each of the regimes. Hence for this condition, speed predicting modeling was done by considering overall classified vehicular volume only. In another study conducted related to speed prediction modeling developed simultaneous equations to predict speed for individual vehicle categories in a mixed traffic condition. Models developed in this study were based on the assumption that the speed is having a linear relationship with individual densities rather than volumes, this assumption is a drawback which inspires to utilize ANN based approach [1].

2. ARTIFICIAL NEURAL NETWORK

Artificial Neural Networks (ANNs) are relatively crude electronic models based on the neural structure of the brain, that functions like the brain by learning from the experience [7]. In a simple means they simulate the structure and the functioning of biological neurons. The function of the network is determined by the architecture of the network, the magnitude of the weights and the processing element's mode of operation. The artificial neural networks comprises of simple processing elements of input, hidden and output layers and interconnection between these layers are strengthened by weights as shown in the figure 1. Which is simple multi-input feed forward artificial neural network.

x and y in Fig 1, are the input and the output variables of the model respectively. I, H and O represents the input, hidden and output layers. Activation (transfer) function and the weightage factor are denoted by f and w respectively.

There are many types of neural networks. Such as,

- Single-layer feed forward network
- Multi-layer feed forward network
- Recurrent network

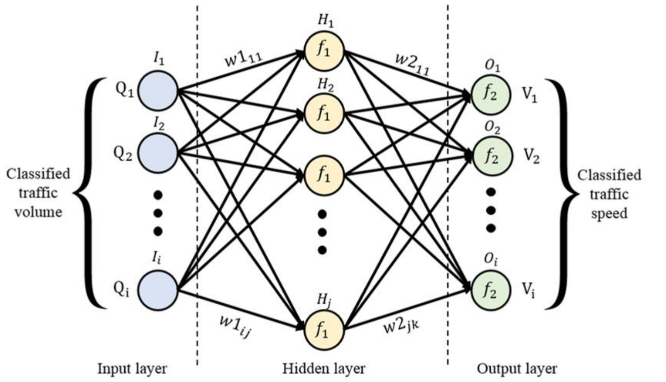


Fig -1: Typical multi-layer feed-forward artificial neural network

ANN employs two types of training supervised and unsupervised. In supervised learning both the inputs and outputs are provided. The network then compares generated outputs with the desired outputs [4]. This paper uses the supervised training for network preparation. While in unsupervised training the network is provided with inputs but not with desired outputs. Presently, the unsupervised learning is not well understood and there is a need to do a lot of research in this aspect.

The main working methodology of ANN network is that each of the inputs are multiplied by the weightage given to each connection. At the hidden layer these sums of the inputs are converted into outputs through some transfer functions, and same process applies from hidden layer to output layer. ANN offers very large scope of transfer functions, such as sigmoid, hard limit, tan hyperbolic, pure linear, etc. The results obtained are compared to target data which are fed to the network and collected from the field. This one complete process is called an epoch in ANN terminology. This process is repeated until the difference in the weightage value from the previous iteration comes under a certain range of values or it reaches at a predefined number of iteration. The backpropagation (BP) algorithm is the most commonly used training method for feed forward networks. The BP algorithm is based on two step process: forward activation and backward propagation. In forward activation network produces solution converting input into the outputs through the transfer functions, and these outputs are compared to target values and back propagated to modify the weights to minimize error in the next iteration.

3. STUDY STRETCH LOCATION

For this study purpose 6-lane divided mid-block section on S.G. Highway in Ahmedabad was selected. Few factors have been taken care of while selecting the road section. The selected stretch is located away from the influences of intersections. There is no parking places, bus stops or pedestrian cross-flow in the neighborhood of the section. Moreover, only the straight sections without any vertical and horizontal curvatures have been chosen for the data collection purpose. Selected stretch has a width of 10.5 m identical on both sides with paved shoulders of width 1.5 m in both the directions with proper road markings and proper median.

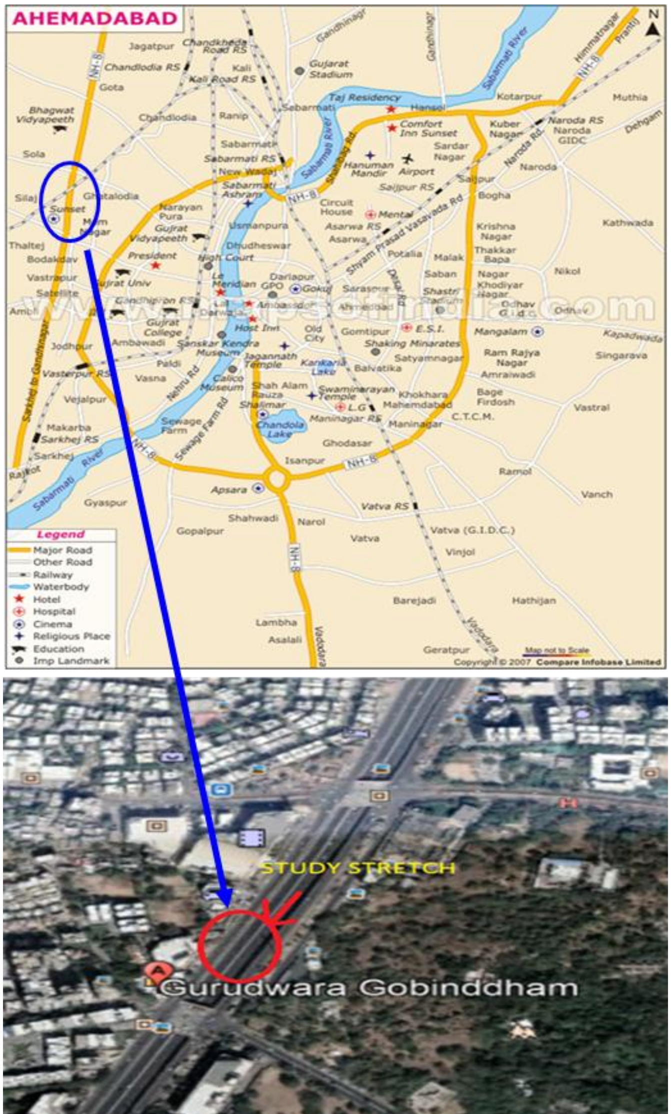


Fig -2: Study area location

4. DATA COLLECTION

Data collected precisely & accurately aids in correct analysis of the same. Traffic surveys are conducted to gather data on classified vehicular volume and speed for the individual category of vehicles. Manual as well as videography technique is used for the data collection.

4.1 CVC for peak hour determination

Traffic volume count survey was conducted on S. G. Highway in Ahmedabad at gurudwara underpass for 10 hours in both direction by taking Iskcon to Vaishnodevi as an upward direction and Vaishnodevi to Iskcon as a downward direction for reference purpose. The hourly variation in traffic volume is obtained from this survey

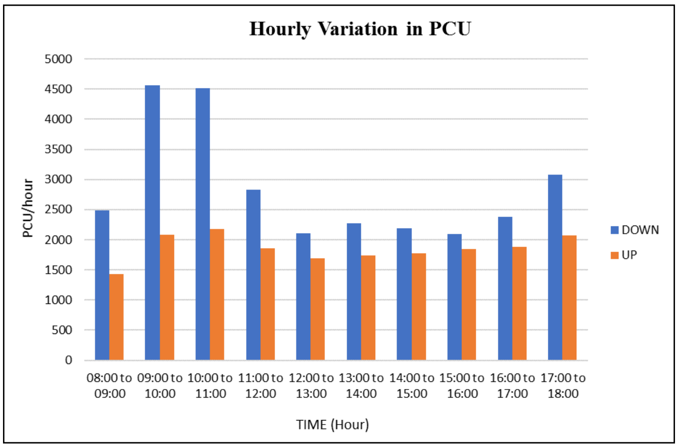


Chart -3: Hourly variation of traffic volume in down and up direction

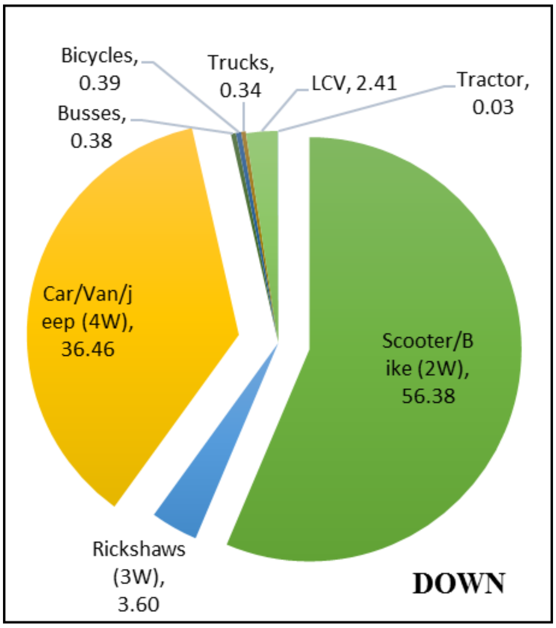


Chart -4: Composition in down direction

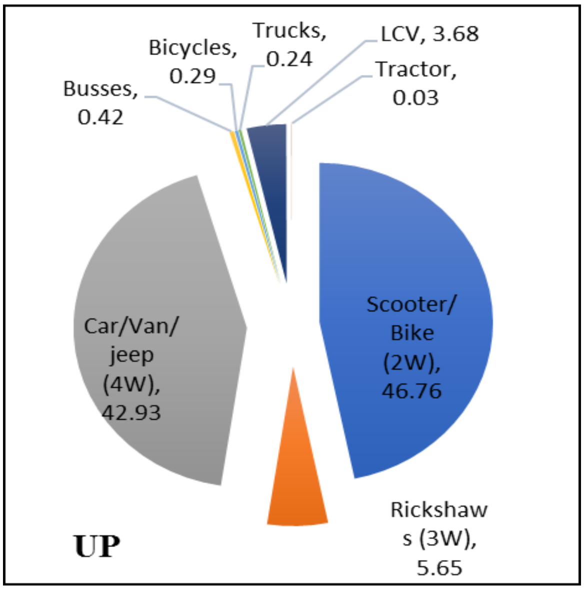


Chart -5: Composition in up direction

From 10 hours survey it is observed that 2W and cars consists of at least 90% of the total traffic volume. From survey further videography is conducted in both the direction during morning peak hours between 09:00 to 11:00 AM (09:00 to 11:00), afternoon off-peak hour between 01:00 to 02:00 PM (13:00 to 14:00) to cover different variations in the traffic flow during the daytime and evening peak hours between 04:00 to 06:00 PM (16:00 to 18:00).

4.2 Video recording during peak hours

Video graphic traffic survey is conducted for 5 hours in morning peak of 2 hours (09:00 to 11:00 am), afternoon off-peak of 1 hour (01:00 to 02:00 pm) and evening peak of 2 hours (04:00 to 06:00 pm) for determination of classified vehicular volume, vehicular composition and spot speed measurements on selected stretch. Classified traffic volume has been extracted by counting the number of vehicles crossing the predefined trap of 50 m in a time span of 5-min intervals. Based on the observation, all vehicles have been classified into five categories as per the table 1.

Table -1: Vehicle categories

Vehicle category	Vehicles included
Small car (SC)	Small cars
Big car (BC)	Larger sized SUV cars
Heavy vehicle (HV)	Light commercial vehicles, Busses and Trucks
3-wheeler (3W)	Auto-rickshaw
2-wheeler (2W)	Scooters and motorcycles

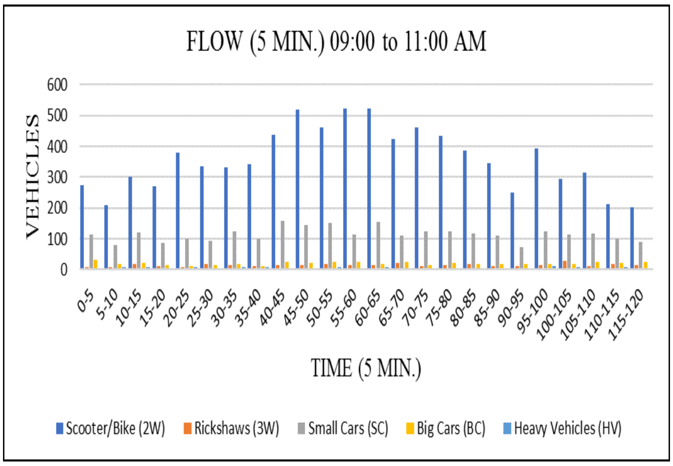
Spot speed on the other hand, has been measured by recording the entry and exit time of a vehicle from the video on the predefined 50 m longitudinal trap marked on the road surface during the survey. For 5-minute interval minimum 40 vehicles are selected randomly from recordings. Thus total 60 data sets have been formed for both down and up direction.

5. DATA ANALYSIS

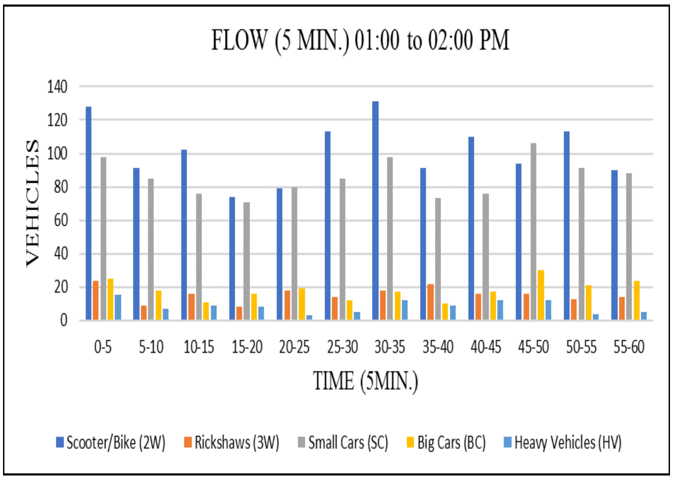
From the conducted video recordings traffic volume and spot speed for each category of vehicles in the 5-min. intervals are extracted.

5.1 Traffic volume

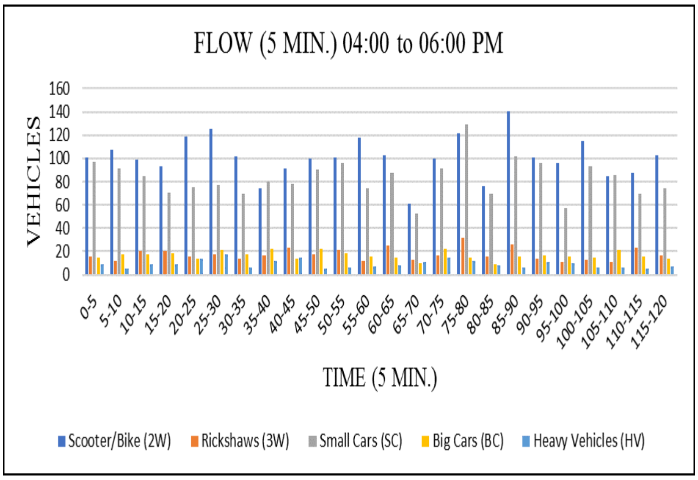
Traffic volume for 5-min. intervals are shown in chart 6 for down and in chart 7 for up direction for the three decided periods of recordings.



(a) 5 min. volume in morning peak

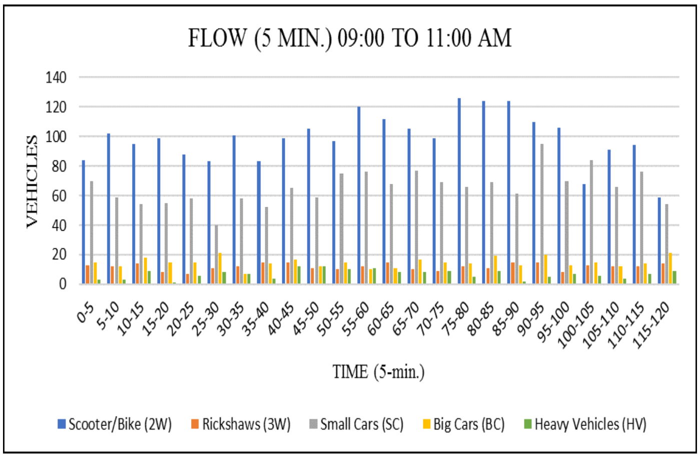


(b) 5 min. volume in afternoon off-peak

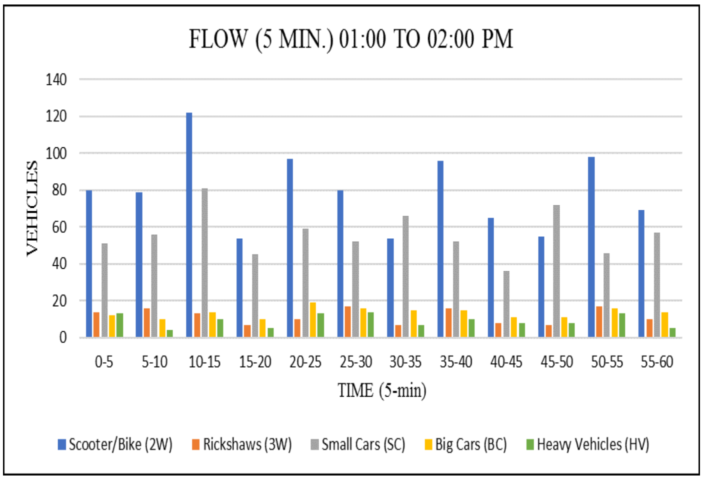


(c) 5 min. volume in evening peak

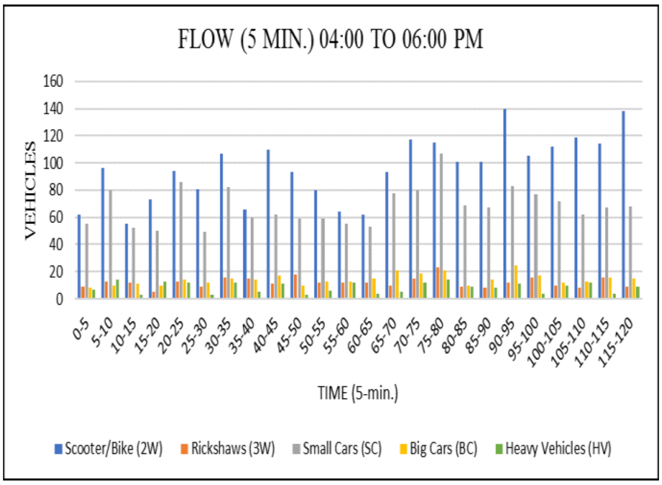
Chart -6: 5-min. volume count for each categories of vehicles in down direction



(a) 5 min. volume in morning peak



(b) 5 min. volume in afternoon off-peak



(c) 5 min. volume in evening peak

Chart -7: 5-min. volume count for each categories of vehicles in down direction

5.2 SPEED MEASUREMENTS

Spot speed for individual vehicle categories are extracted for the 5-min intervals, thus forming 60 data sets for model formulation in both down and up direction. In each 5-min interval minimum 40 vehicles are randomly selected for the spot speed measurements. Table 2 to 7 shows the variations in the speeds of the vehicles for the different periods of the day.

Table -2: Observed speed for morning peak in down direction

Vehicle Categories	Speed (km/hr.)		
	MAX.	MIN.	AVE.
2W	61.28	52.80	57.82
3W	47.27	40.77	44.12
SC	77.72	62.16	68.56
BC	74.50	58.58	66.35
HV	65.25	47.25	56.15

From table 2 to 4 it is observed that average speed of the 2W ranges between 51.56 km/hr. in the evening peak to 57.82 km/hr. in morning peak period, while the average speed of the SC and BC are nearly the same and 3W have the least average speed among the observed categories

Table -3: Observed speed for afternoon off-peak in down direction

Vehicle Categories	Speed (km/hr.)		
	MAX.	MIN.	AVE.
2W	57.24	50.03	52.36
3W	46.41	40.7	43.24
SC	74.51	66.47	69.15
BC	74.12	62.71	67.41
HV	58.32	51.32	55.19

Table -4: Observed speed for evening peak in down direction

Vehicle Categories	Speed (km/hr.)		
	MAX.	MIN.	AVE.
2W	55.9	48.19	51.56
3W	46.41	40.44	43.35
SC	77.03	65.56	71.17
BC	75.83	64.51	69.19
HV	64.12	50.69	55.87

Table 5 to 7 shows speed data in up direction for the three recording periods.

Table -5: Observed speed for morning peak in down direction

Vehicle Categories	Speed (km/hr.)		
	MAX.	MIN.	AVE.
2W	56.87	50.12	53.11
3W	47.2	40.58	42.91
SC	75.95	67.52	71.54
BC	81.15	60.73	69.34
HV	64.12	48.14	57.62

Table -6: Observed speed for afternoon off-peak in down direction

Vehicle Categories	Speed (km/hr.)		
	MAX.	**MIN.**	**AVE.**
2W	54.85	50.25	51.72
3W	49.6	39.26	43.63
SC	73.12	63.7	68.09
BC	76.26	61.06	68.47
HV	62.98	48.82	54.98

Table -7: Observed speed for evening peak in up direction

Vehicle Categories	Speed (km/hr.)		
	MAX.	**MIN.**	**AVE.**
2W	56.89	48.49	51.91
3W	47.36	39.82	42.79
SC	74.15	65.27	69.56
BC	81.25	59.52	68.42
HV	62.34	51.12	57.71

From table 5 to 7 it is observed that the average speed of the SC ranges from 68.09 km/hr. in afternoon off-peak to 71.54 km/hr. in the morning peak period, while the average speed of the other categories of vehicles are nearly same except heavy vehicles. In both of the table average speeds of the vehicles are higher in morning period compared to other periods.

6. MODEL DEVELOPMENT AND COMPARISONS

In the construction of the model, the input variables will be the traffic volume of number of five category of vehicles (table 1) observed in the field and output variables will be the spot speed of individual vehicles for the five categories (table 1) observed in the field. Here a total 60 data points have been formed for both the direction out of which 70% data have been randomly taken for training the ANN model, another 15% for testing the ANN model and the remaining 15% randomly taken for the validation of the model. While measuring the accuracy, the present study has relied on only two parameters; Coefficient of correlation R and Mean Square Error (MSE). One problem is associated with the formulation of an ANN model is to choose the optimum number of hidden neurons. For this, the network has been constructed repetitively for different number of neurons and the corresponding value of R and MSE is compared to decide optimum number of neurons for the preparation of actual models. The network that corresponds to the minimum MSE and the maximum R value, has been considered to serve the purpose of the study. Results have shown (chart 8) that **10** is the optimum number of neurons. Figure 3. represents artificial neural network formulated in MATLAB R2017a with 5 nodes in the input layer, 10 hidden neurons and a single output layer. Here tan hyperbolic is used as a transfer function from input layer to hidden layer and pure linear transfer function is used from the hidden layer to output layer [6]. Classified volumes for the five categories of vehicles is taken as the input variables and classified speed for the various categories of vehicles is taken as the output variables. In present study lavenberg-marquardt function is used as a training algorithm

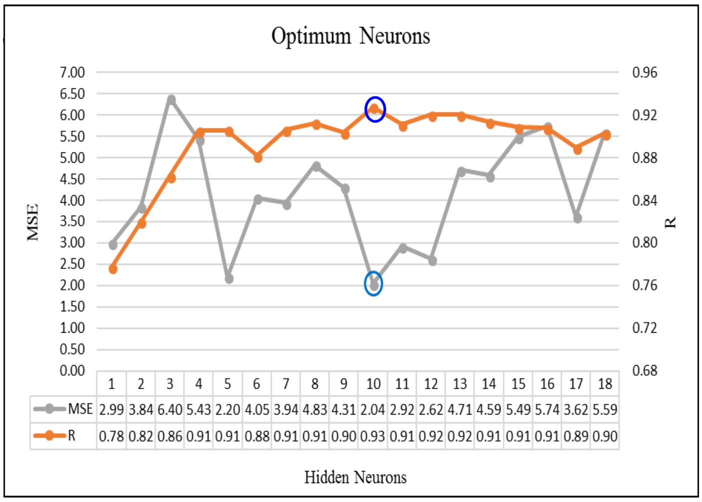


Chart -8: Optimum number of hidden neurons

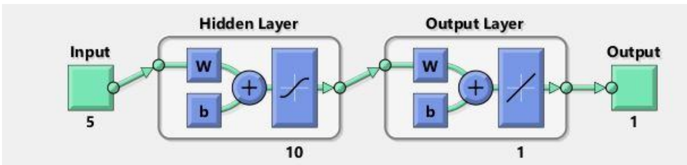
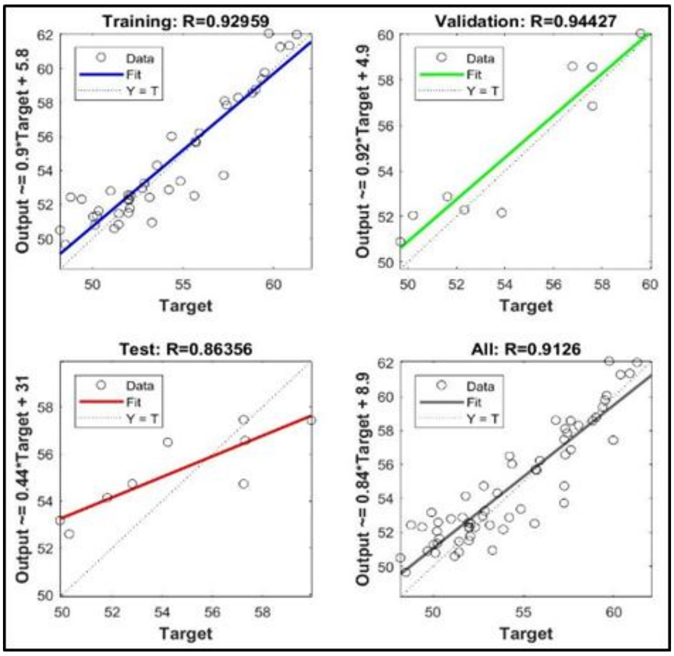


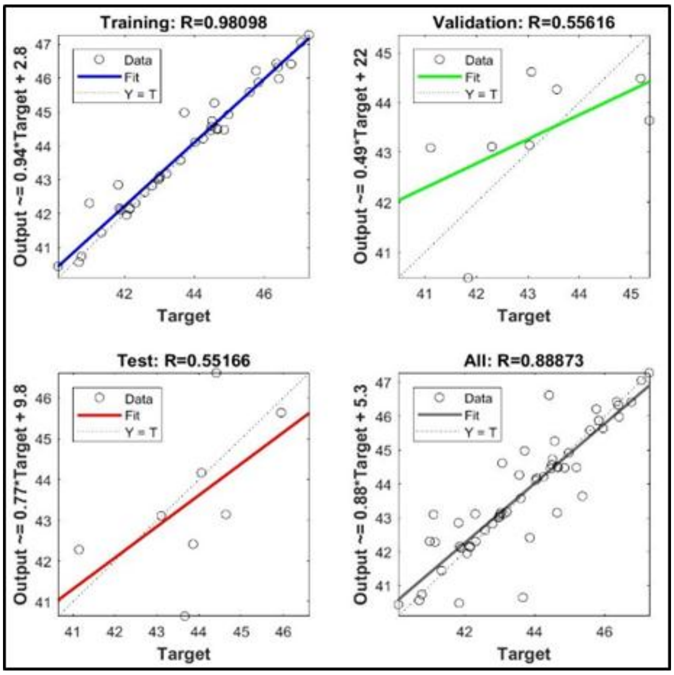
Fig -3: Neural network architecture as prepared in MATLAB

For five vehicles category ANN networks are prepared in MATLAB R2017a package. Regression plots for the prepared models are presented below:

6.1 Regression plots for down direction

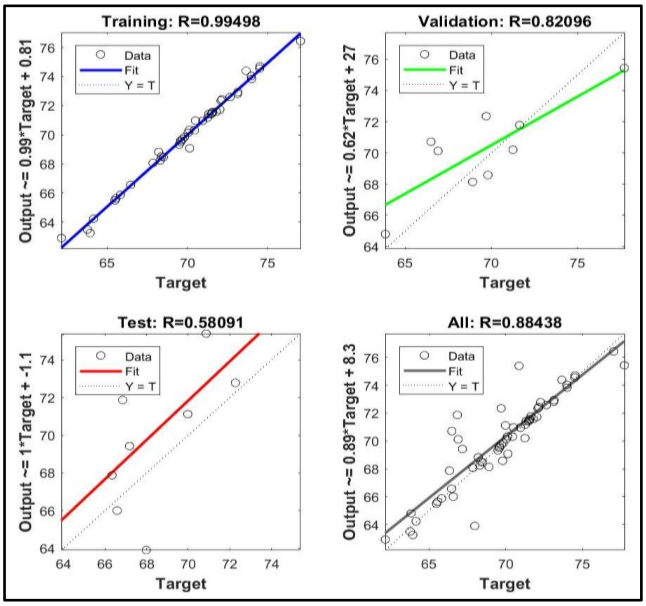


(a) 2W

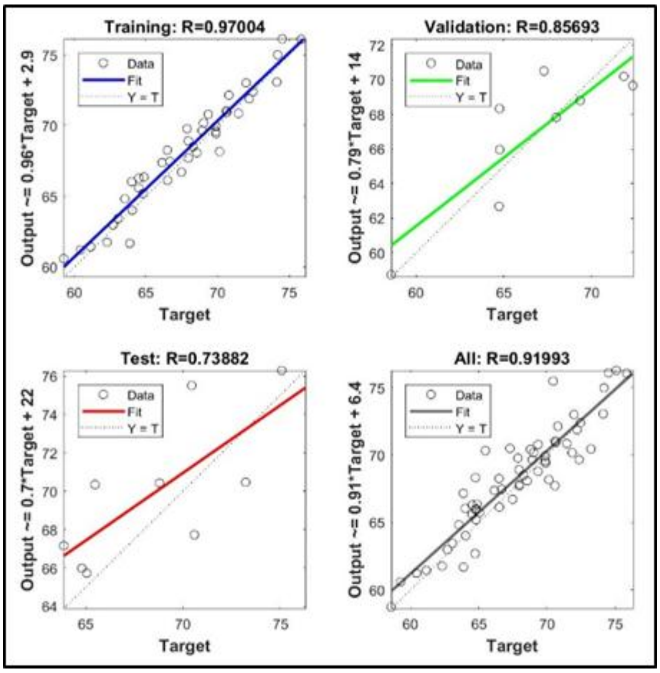


(b) 3W

From regression plots for the 2W and 3W category it is cleared that value of coefficient of correlation R for overall data is 0.9126 and 0.88873 for 2W and 3W respectively which shows good correlation between input and output variables. The value of R for the training, validation and testing data for 2W is good but for the 3W category it is slightly on the lower side.

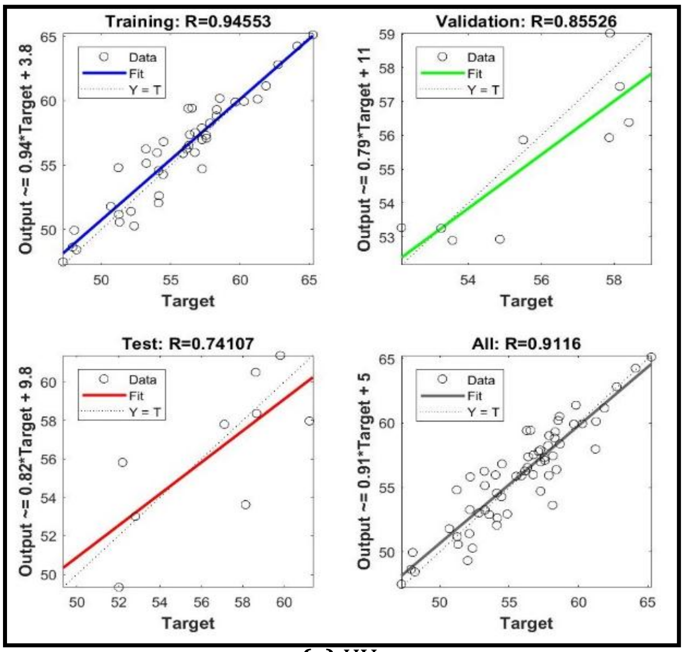


(c) SC



(d) BC

From regression plots for the SC and BC category it is cleared that value of coefficient of correlation R for overall data is 0.88438 and 0.91993 for SC and BC respectively which shows good correlation between input and output variables. The value of R for the training, validation and testing data for SC and BC category is good.

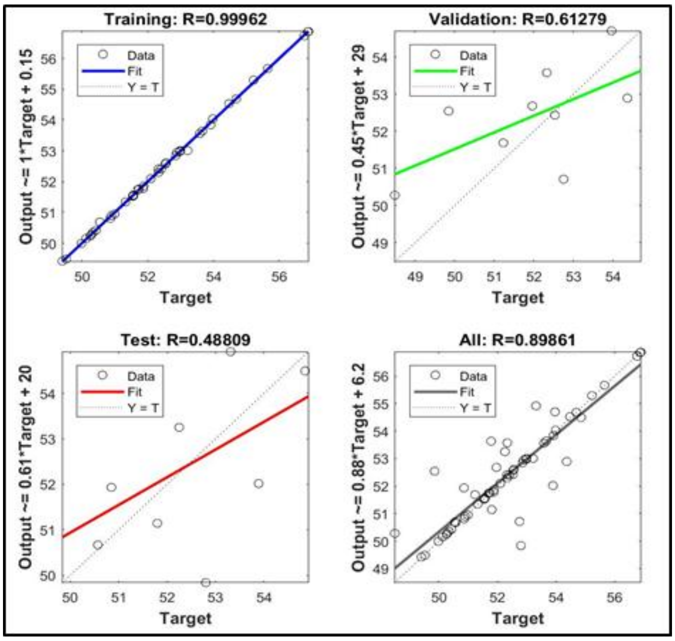


(e) HV

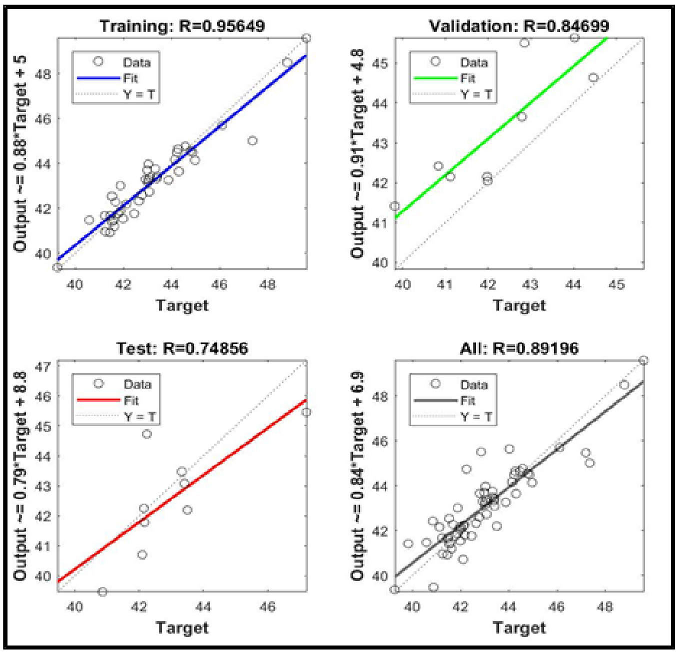
Chart -9: Regression plots for five categories of vehicles in down direction

From regression plots for the HV category it is cleared that value of coefficient of correlation R for overall data is 0.9116 which shows good correlation between input and output variables. The value of R for the training, validation and testing data is also on higher side

6.2 Regression plots for up direction

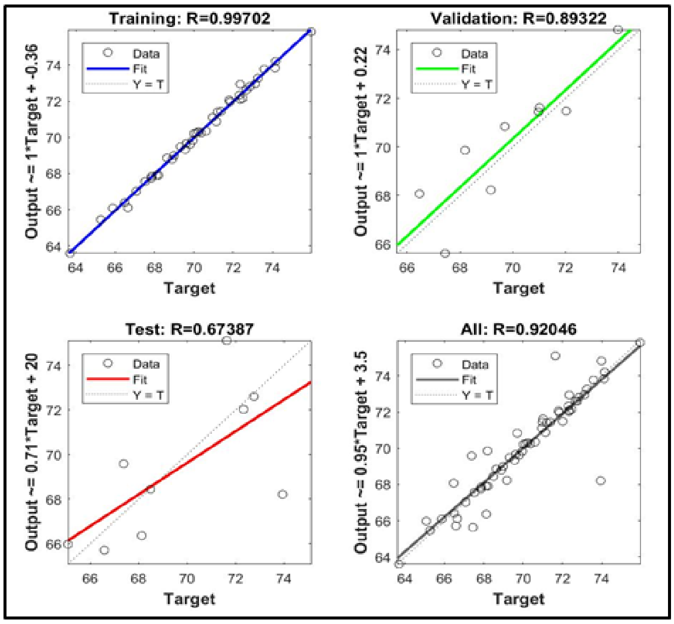


(a) 2W

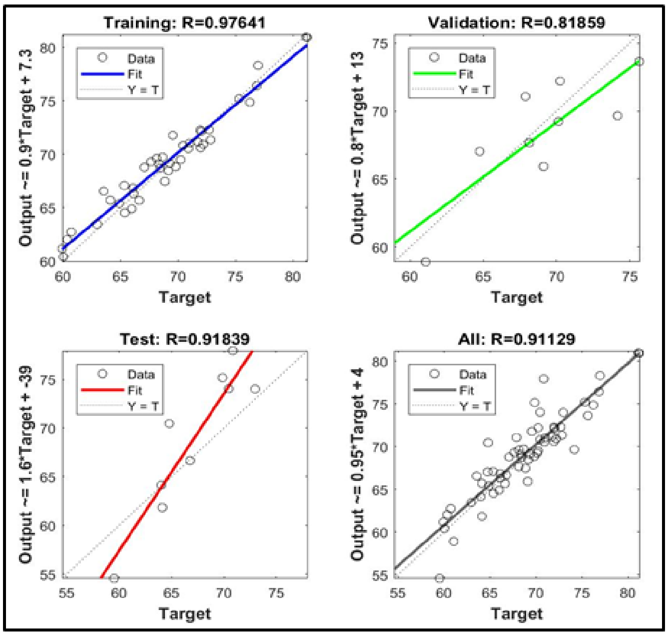


(b) 3W

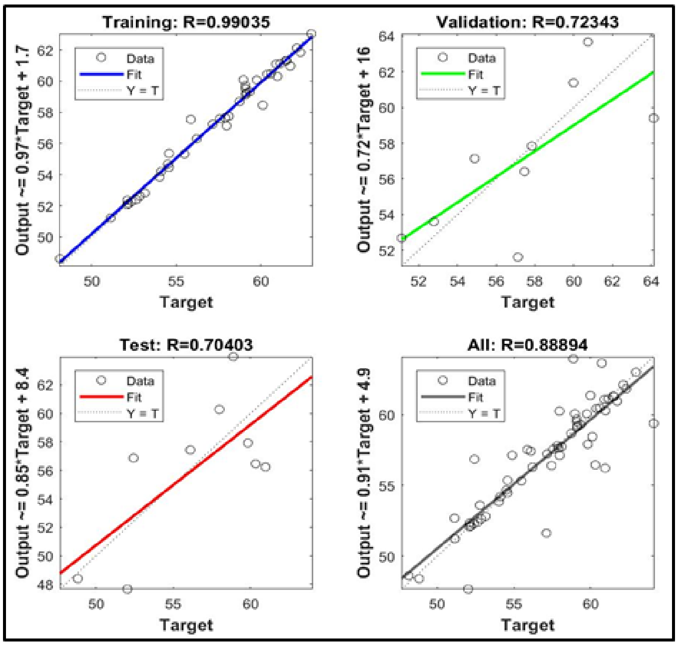
From regression plots for the 2W and 3W category it is cleared that value of coefficient of correlation R for overall data is 0.89861 and 0.89196 for 2W and 3W respectively which shows good correlation between input and output variables. The value of R for the training, validation and testing data for 3W is good but for the 2W category it is slightly on the lower side.



(c) SC



(d) BC



(e) HV

Chart -10: Regression plots for five categories of vehicles in down direction

From regression plots for the SC, BC and HV category it is cleared that value of coefficient of correlation R for overall data is 0.92046, 0.91129 and 0.88894 for SC, BC and HV respectively which shows good correlation between input and output variables. The value of R for the training, validation and testing data is also indicates strong relationship for the three respective category of vehicles.

These ANN models are than compared to regression models. Density based regression and volume-based regression models which are prepared for individual category of vehicles in SPSS software. In density-based regression models the assumption that the speed is having a linear relationship with individual densities is considered. Since the field data collection of density is difficult compared to that of traffic volume, the model was modified by expressing density as the ratio of volume and speed. Volume-based regression models are developed by using vehicular volumes of different categories of vehicles as independent variables and speed of individual category of vehicles as dependent variables. Table 8 and 9 shows comparisons of the R values between volume-based ANN models and density and volume-based regression models. it is clear that ANN model performs better than all other model in predicting speed for individual category of vehicles.

Table -8: Comparison of R values for down direction

TYPE OF VEHICLES	REGRESSION MODEL BASED ON VOLUME	REGRESSION MODEL BASED ON DENSITY	ANN MODEL BASED ON THE VOLUME
Two-Wheelers (2W)	0.71	0.75	0.90
Three-Wheelers (3W)	0.35	0.41	0.89
Small Cars (SC)	0.46	0.55	0.92
Big Cars (BC)	0.63	0.70	0.91
Heavy Vehicles (HV)	0.58	0.53	0.89
AVERAGE	**0.55**	**0.59**	**0.90**

Table -9: Comparison of R values for up direction

TYPE OF VEHICLES	REGRESSION MODEL BASED ON VOLUME	REGRESSION MODEL BASED ON DENSITY	ANN MODEL BASED ON THE VOLUME
Two-Wheelers (2W)	0.76	0.78	0.91
Three-Wheelers (3W)	0.47	0.68	0.89
Small Cars (SC)	0.75	0.72	0.88
Big Cars (BC)	0.67	0.47	0.92
Heavy Vehicles (HV)	0.57	0.66	0.91
AVERAGE	**0.64**	**0.66**	**0.90**

7. CONCLUSIONS & SCOPE OF THE FURTHER RESEARCH

This study uses traffic volume and its composition data for the determination of vehicular speed on urban arterials under mixed traffic conditions. An ANN based approach has been adopted in this study to arrive at a more accurate prediction of speed. The conclusions of the study are summarized below:

1. The parameters influencing the speed of the vehicles found out which are vehicular volume, composition of vehicles and density of vehicles.

2. The major composition in traffic stream is occupied by small cars and two-wheelers ranging between 30-45% and 40-60% for small cars and two-wheelers respectively.

3. The speed model developed considering an optimum number of hidden neurons using ANN methodology, from the comparison (table 8 & 9) of the above models it is clear that ANN models outperform regression models, while the density-based regression models perform slightly better than volume-based regression models.

4. The regression predicted values and ANN predicted values were found to be nearer to the actual observed values. This shows that these models are suitable for speed prediction of vehicles on selected road links.

Further the analysis on variation in speed of the vehicles due to the effect of varying traffic composition and traffic volume can also be carried out by utilizing ANN methodology. The methodology presented in this paper can be further used to observe the influences of speed on road geometry, side friction, weather condition and similar other factors that affects traffic stream performance.

REFERENCES

[1] Ashish Dhamaniya, et. al, (2013). "Speed Prediction Models for Urban Arterials under Mixed Traffic Conditions". 2nd Conference of Transportation Research Group of India (2nd CTRG).

[2] Balaji Ponnu, et. al, (2006). "Vehicle Class-wise Speed Volume model for Three-lane Undivided Urban Roads". Journal of Transportation Engineering, ASCE, Vol. 132, No. 6, June 1.

[3] Chetan R Patel, Dr. G. J. Joshi. (2012). "Capacity and LOS for urban arterial road in Indian mixed traffic condition". Proscenia - Social and Behavioural Science, pp.527-534.

[4] Haykin. S., "Neural Networks: A Comprehensive Foundation", 2nd Edition, Prentice Hall, 1999.

[5] Highway Capacity Manual (HCM 2010), Transportation Research Board.

[6] Howard Demuth, Mark Beale, "Neural Network Toolbox", Version 4, 2002.

[7] I.A.Basheer, M.Hajmee. (2000). "Artificial Neural Networks: Fundamentals". Journal of Microbiological Methods 43, 3-31.

[8] IRC (1990), "Guidelines for capacity of urban roads in plain areas". IRC 106:1990, Indian Road Congress, New Delhi, India.

[9] Thomas, J., et. al, (2012). "Vehicle class wise speed-volume models for heterogeneous traffic", Transport, Vol. 27(2), pp. 206-217.

[10] http://nptel.ac.in/courses/117105084/3 (NPTEL lecture on Neural Networks and applications)