Estimation and Prediction of Traffic Parameters under Mixed Traffic Conditions

Kawsar Ahmed Shah[1],Ashish Kumar[2], Dr. Sandeep Singla[3]
[1]M. Tech Scholar Civil Engineering, [2]Assistsnt Professor Civil Engineering, [3]Professor/Head of Department Civil Engineering
Department RIMT University Punjab India.

***Abstract**:* Approaches for mining this wealth of information to benefit the drivers and traffic authorities. The main reason is most likely related to fundamental challenges in dealing with big data of various types and uncertain frequency coming from diverse sources. Mainly, the issues of types of data and uncertainty analysis in the predictions are indicated as the most challenging areas of study that have not been tackled yet. Important issues such as the nature of the data, i.e., stationary or non-stationary, and the prediction tasks, i.e., short-term or long-term, should also be taken into consideration. Based on this observation, a data-driven traffic flow prediction framework within the context of big data environment is proposed in this thesis. The main goal of this framework is to enhance the quality of traffic flow predictions, which can be used to assist travelers and traffic authorities in the decision-making process (whether for travel or management purposes). Static models and simulations are commonly used in urban traffic management but none feature a dynamic element for near real-time traffic control. This work presents an artificial neural network forecaster methodology applied to traffic flow condition prediction. The spatially distributed architecture uses life-long learning with a novel adaptive Artificial Neural Network based filter to detect and remove outliers from training data. The system has been designed to support traffic engineers in their decision making to react to traffic conditions before they get out of control. We performed experiments using feed-forward backpropagation, cascade-forward back-propagation, radial basis, and generalized regression Artificial Neural Networks for this purpose. Test results on actual data collected from the city of Leicester, UK, confirm our approach to deliver suitable forecasts

Keyword: **traffic, signal, optimization, cost.**

I. INTRODUCTION

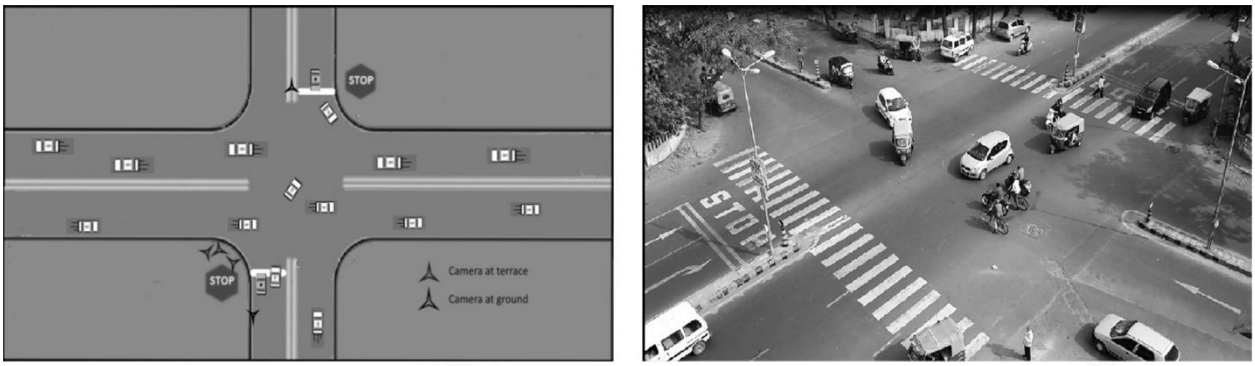
Heterogeneous traffic conditions prevail in all countries across the world, but the degree of heterogeneity is different in developed and developing countries. In the United States, heavy vehicles (trucks and buses) are mixed with passenger cars, whereas more than seven different categories of vehicles in India are observed in the traffic stream. Such heterogeneity is also observed at tollbooths, where dedicated lanes are provided for each vehicle category; however, during peak hours, lane discipline is broken, and the same lane is used by different categories of vehicles. Such mixed traffic conditions lead to wide variations in service time for the same vehicle category and, hence, reduce tollbooth capacity. To convert the mixed traffic flow into the equivalent passenger car, the tollbooth equivalency factor (TEF) is proposed in the present study. The TEF is based on the service time and clearance time of a vehicle with respect to a passenger car at the same tollbooth. The TEF is found not to be a fixed value but varies with the approaching mixed traffic volume and composition of traffic in the queue. Vehicles in the traffic stream are divided into seven categories, and simultaneous equations are developed to determine the service time of a vehicle type based on approach volume and traffic composition. These equations are further used to depict variations in TEF with varying approach volume and composition. The change in TEF values is explained on the basis of the relative change in service time of a vehicle type with respect to a standard car at different volume levels. The accuracy of the TEF values estimated through simultaneous equations is checked by comparing the estimated values with those calculated directly from the field data. TEF values obtained for each type of vehicle are multiplied by the number of that vehicle type to obtain the approaching mixed traffic volume in an equivalent homogenous mix (TEF/hTEF/h). This method will naturally require an estimation of TEF for each vehicle type from the field data. To avoid this exercise, the stream equivalency factor (SEF) at the tollbooth is proposed, which provides an overall multiplying factor for the entire traffic volume to convert heterogeneous traffic into the homogenous equivalent. The results presented in this paper will be useful for planners and tollbooth managers for the design and performance evaluation of toll plazas and to identify the number of lanes required during peak and nonpeak hours.

1.1 Traffic State Estimation and Prediction under Heterogeneous Traffic Conditions

The challenges faced by urban traffic in both developed and developing countries are infrastructure deficiency, congestion, accidents, and environmental and health damages due to pollution. Though all of these are important, the problem of traffic congestion is the most visible and affects a large number of motorists directly on a day-to-day basis. Urban areas in most of the developing countries are facing major challenges in traffic management and control in recent decades and India is no exception. It has witnessed a rapid growth in economy in recent years, resulting in vehicle ownership levels growing at a much faster rate. For example, the number of registered vehicles in India's six major metropolises went up by 7.75 times during 1981 to 2001, while the population increased only by 1.89 times. Thus, the growth of motor vehicles was almost four times faster than the growth of population . The World Bank reported that the economic losses incurred on account of congestion and poor roads alone run as high as $6 billion a year in India . Though there are various solution options like infrastructure expansion, Transportation System Management (TSM) measures and congestion pricing, technology applications like the Intelligent Transportation System (ITS) proved to be an efficient way to reduce congestion in developed countries like U.S.A. . Two of the major building blocks of ITS are the Advanced Traveler Information System (ATIS) and the Advanced Traffic Management System (ATMS). The ATIS offers users realtime traveler information enabling them to make better and more informed travel decisions that will lead to more efficient distribution of travelers to routes and modes. The ATMS detects traffic situations, transmits them to control center via a communication network, and then develops optimal traffic control strategies by combining

the available traffic information. Both ATIS and ATMS require the accurate estimates of the current traffic state and prediction of its short term evolution in future in order to ensure smooth traffic flow. The techniques for traffic state estimation and prediction can be grouped into data driven approaches and model based approaches. One of the major drawbacks of data-driven methods is that they correlate the mean (observed) traffic conditions to current and past traffic data, without explicitly incorporating the physical aspects of the traffic as model based approaches do. However only very few studies were reported which utilizes the macroscopic traffic flow models for online traffic state prediction and notably all of the studies were deployed under homogeneous lane discipline traffic conditions as discussed under the literature review section. There were no similar reported studies where the macroscopic traffic flow model is used for online traffic state estimation and prediction under heterogeneous traffic conditions as existing in India. Hence, the objective of the present study is to investigate the use of the existing Lighthill-Whith am Richards (LWR) macroscopic model for traffic state estimation and prediction in a busy arterial road of Chennai under heterogeneous traffic conditions.

1.2 Estimation of average space headway under heterogeneous traffic conditions.



(a) **(b)**

Fig.1.1 : Critical gap estimation

Space headway is defined as the distance between the same points of two consecutive vehicles following each other. Space headway values are essential in a wide spectrum of traffic engineering studies, varying from the classic overtaking maneuver problem to congestion prediction and automated highway systems control. The significance of this traffic parameter is mainly due to its direct link to traffic density and hence its ability to reflect the level of congestion. However, this is a parameter which is difficult to measure in field using any automated procedure and hence is usually estimated from other parameters. Also, a majority of the studies that had analysed headway were on finding the statistical distribution of the data. The present study attempts to estimate the average space headway using a model based approach with special reference to congestion prediction for Intelligent Transportation Systems (ITS) applications. The basic equations that were used are the conservation of vehicles equation and the fundamental traffic flow relation connecting flow, speed and density. The space headway was estimated from the density inside the section. The Extended Kalman Filter (EKF) is used in the estimation scheme. The model results are corroborated using the actual headway calculated using the actual traffic density measured from the field. Video graphic data collected from Rajiv Gandhi Road, Chennai, India, were used in this study.

1.3 Development of speed prediction model for mixed traffic conditions.

A micro level analysis of vehicular speed is done to develop model between two mean speeds of vehicles such as time mean speed (TMS) and space mean speed (SMS) under mixed traffic stream. Traffic volume and speeds were collected on a mid-block section of an urban arterial to analyse different stream parameters. Vehicular composition along with the individual speed characteristics of each vehicle class was extracted from the collected field data. Models are framed between the speed characteristics of vehicles estimated from the field. A comparative analysis is also done with the existing traditional model to check the applicability of the suggested model. The result shows that the suggested models are reliable in forecasting vehicular speeds with higher accuracy.

Fig:1.2 Speed and time headway distribution

II.PROPOSED METHODOLOGY

Methodology steps

Step1: Multiple Linear Regression is one of the most popular statistical techniques for fitting mathematical relationship between dependent and independent variables. This technique has been applied fruitfully in a number of transportation planning studies that can be used for predicting values of a dependent variable from one or more independent variables. The general form of the equation is:

Step2: Y = ao + a1 X1 + a2 X2 + a3 X3 + ... + an Xn ……(1) where Y is the dependent variable such as Saturation Flow Rate in the present study. X1, X2, X3 . . . . . Xn are the independent variables such as proportion of highly maneuverable vehicles, width of the approach under consideration etc. a1, a2, a3 . . . . . an are the coefficients of the respective independent variables X1, X2, X3 . . . . . Xn ao is the regression coefficient.

Step3: The Multiple linear regression analysis finds the values of ao, a1, a2, . … such that the error of estimation is minimum. In this equation, the regression coefficients represent the independent contributions of each independent variable to the prediction of the dependent variable. The regression line expresses the best prediction of the dependent variable (Y), given the independent variables (X). The deviation of a particular point from the regression line (its predicted value) is called the residual value. The smaller the variability of the residual values around the regression line relative to the overall variability, the better is the prediction.

Step4 : In most cases, the ratio would fall somewhere between these extremes, that is, between 0.0 and 1.0. 1.0 minus this ratio is referred to as R-square or the coefficient of determination. This value is immediately interpretable in the following manner. If the value of R-square is 0.4,it means that the variability of the Y values around the regression line is (1-0.4) times the original variance. In other words the equation can explain only 40% of the original variability, and the remaining 60% is the residual variability. The R-square value is an indicator of how well the model fits the data. Customarily, the degree to which two or more predictors (independent or X variables) are related to the dependent variable is expressed in the correlation coefficient R, which is the square root of R-square.

Step 5 : In multiple regression, R can assume values between 0 and 1. To interpret the direction of the relationship between variables, one looks at the signs (plus or minus) of the regression coefficients. It is assumed in multiple regression that the residuals (predicted minus observed values) are distributed normally (i.e., follow the normal distribution). The major conceptual limitation of all regression techniques is that one can only ascertain relationships, but never be sure about underlying causal mechanism.

Grey Wolf Optimization Algorithm (GWO)

The latest bio-inspired algorithm is the grey wolf optimization algorithm. This algorithm's main concept is simulating the behaviour of grey wolf living in a pack. They have serious hierarchy of social dominance. Alpha is known as the level leaders and is responsible for decision making in the pack. The wolf pack persistence is based on decision of alpha. Beta is known as the second level subordinate wolves. The beta operation is for helping in making decision for alpha or other activities.

Delta is known as the third level subordinate wolves. This category member consists of elders, scouts, hunters, caretakers and sentinels. For region boundary observation and in any danger case, scouts are liable for warning. The protection and pack's safety guarantee is given by sentinels. The expertise wolves are the elders, denoted as alpha or beta. Alphas and betas are helped by hunters while prey hunting and caring for the ill, weak, and wounded wolves by caretakers and providing food for pack. Omega is the lowest level. All dominant wolves with which omega wolves have to comply.

Grey wolves have the ability of memorizing the prey position and encircling them. The alpha as a leader performs in the hunt. For simulating the grey wolves hunting behaviour in the mathematical model, assuming the alpha (α) is the best solution. The second optimal solution is beta (β) and the third optimal solution is delta (δ). Omega (ω) is assumed to be the candidate solutions. Alpha, beta and delta guides the hunting while position should be updated by the omega wolves by these three best solutions considerations.

Encircling prey

Prey encircled by the grey wolves during their hunt. Encircling behaviour in the mathematical model, below equations is utilized.

$$\vec{A}(T+1) = \overrightarrow{A_P}(T) - \vec{X}.\vec{Z}$$

$$\vec{Z} = \left|\vec{Y}.\overrightarrow{A_P}(T) - \vec{A}(T)\right|$$

Where

T←iterative number

$\vec{A}$←grey wolf position

$\overrightarrow{A_P}$←prey position

$$\vec{X} = 2x.\overrightarrow{r_1} - x$$

$$\vec{Y} = 2\overrightarrow{r_2}$$

Where

$\overrightarrow{r_1}$and $\overrightarrow{r_2}$←random vector range[0,1]

The x value decrease from 2 to 0 over the iteration course.

$\vec{Y}$←random value with range [0,1] and is used for providing random weights for defining prey attractiveness.

Hunting

For grey wolves hunting behavior simulation, assuming α, β, and δ have better knowledge about possible prey location. The three best solutions firstly and ω (other search agents) are forced for their position update in accordance to their best search agents position. Updating the wolves' positions as follows:

$$\vec{A}(T+1) = \frac{\overrightarrow{A_1}+\overrightarrow{A_2}+\overrightarrow{A_3}}{3} \quad (1)$$

Where $\overrightarrow{A_1}, \overrightarrow{A_2}, and\ \overrightarrow{A_3}$ are determined,

$$\overrightarrow{A_1} = \left|\overrightarrow{A_\alpha} - \overrightarrow{X_1}.Z_\alpha\right|$$

$$\overrightarrow{A_2} = \left|\overrightarrow{A_\beta} - \overrightarrow{X_2}.Z_\beta\right|$$

$$\overrightarrow{A_3} = \left|\overrightarrow{A_\delta} - \overrightarrow{X_3}.Z_\delta\right|$$

Where $\overrightarrow{A_\alpha}, \overrightarrow{A_\beta}, and\ \overrightarrow{A_\delta}$← first three best solution at a given iterative T

Z_α, Z_β, and Z_ω are determined,

$$\overrightarrow{Z_\alpha} \leftarrow \left|\overrightarrow{Y_1}.\overrightarrow{A_\alpha} - \vec{A}\right|$$

$$\overrightarrow{Z_\beta} \leftarrow \left|\overrightarrow{Y_2}.\overrightarrow{A_\beta} - \vec{A}\right|$$

$$\overrightarrow{Z_\delta} \leftarrow \left|\overrightarrow{Y_3}.\overrightarrow{A_\delta} - \vec{A}\right|$$

The parameter x updating is the final process. The parameter x exploitation and exploration is updated linearly for ranging [2, 0] in every iteration.

$$x = 2 - t\frac{2}{maxI}$$

Where

T←iterative number

MaxI←total number of iteration

Algorithm
Step 1: Input Bugs in the form of PAVEMENT data. Step 2: Preprocessing of data test to remove the noisy data. Step 3: Extract the Bigrams and make the matrix Step 4: Apply n-layer convolution and mapping by Relu.

Step 5: Pooling of function and generalize the matrix.

Step 6: Apply the SVM

With optimization model $\boldsymbol{min_{\omega,\xi,Q}P(\omega,\xi)}$ we describe the model of SVM classification.

$$\boldsymbol{min_{\omega,\xi,Q}P(\omega,\xi_r) = \frac{1}{2}\omega^g\omega + \frac{1}{2}\gamma\sum_{r=1}^{n}\xi_r{}^2}$$

$$\boldsymbol{s_r[\omega^t\phi(u_r) + Q = 1 - \xi_r, r = 1, 2, \ldots\ldots, n}$$

$$\boldsymbol{\xi = (\xi_1, \xi_2, \ldots\ldots\xi_n)}$$

Where

$\boldsymbol{\xi_r}$←Slack variable

Q←Offset

$\boldsymbol{\omega}$←Support vector

$\boldsymbol{\gamma}$←Classification parameter for balancing the model complexity and fitness error.

Then, describing the classification decision function:

$$\boldsymbol{F(z_r) = sgn(\sum_{r=1}^{n}\alpha_r s_r L(q, q_r) + Q)}$$

Step 7: Make learning and testing Model.

Step 8: Analyze

III.RESULT AND ANALYSIS

In the section proposed result and comparison with different algorithm result is presented. This result is calculated on the heat generators and power generators on Genetic Algorithm, PSO, and Genetic with PSO.

Table 3.1: Generator comparison values

Approaches	Traffic flow(100)Cost	Traffic flow(200)Cost	Traffic flow(300)Cost
Genetic Algorithm	14.758	331.56	0.0321
PSO	125.48	69.0432	0.205
ANN-PSO	0.0169	778.27	0.0253

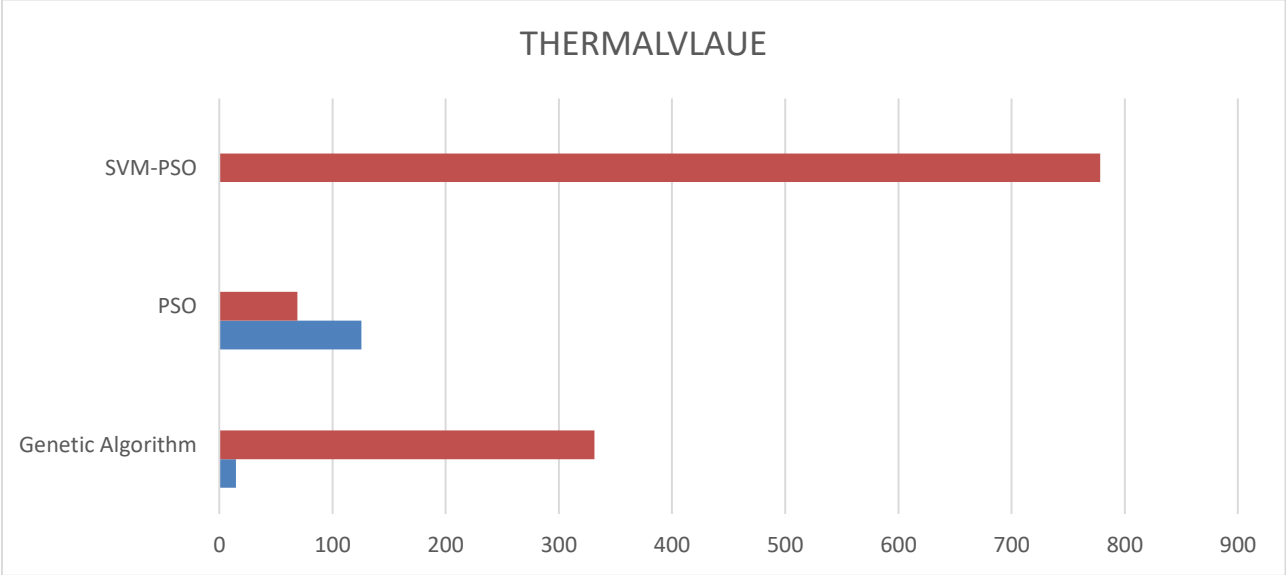


Figure 3.1Traffic flow on different algorithms

In figure 3.1 depicts the heat generator values on the different algorithm approaches. The x-axis represents the algorithms and y-axis represents the values of the generator. The Genetic with PSO gives the effective heat generator values.

Table 3.2: Different Algorithm Approaches

Approaches	Traffic Congestion(123)	Traffic Congestion(223)	Traffic Congestion(323)
Genetic Algorithm	49.11	79.7093	75.332
PSO	49.118	79.7093	75.332
ANN-PSO	49.11	79.7093	75.332

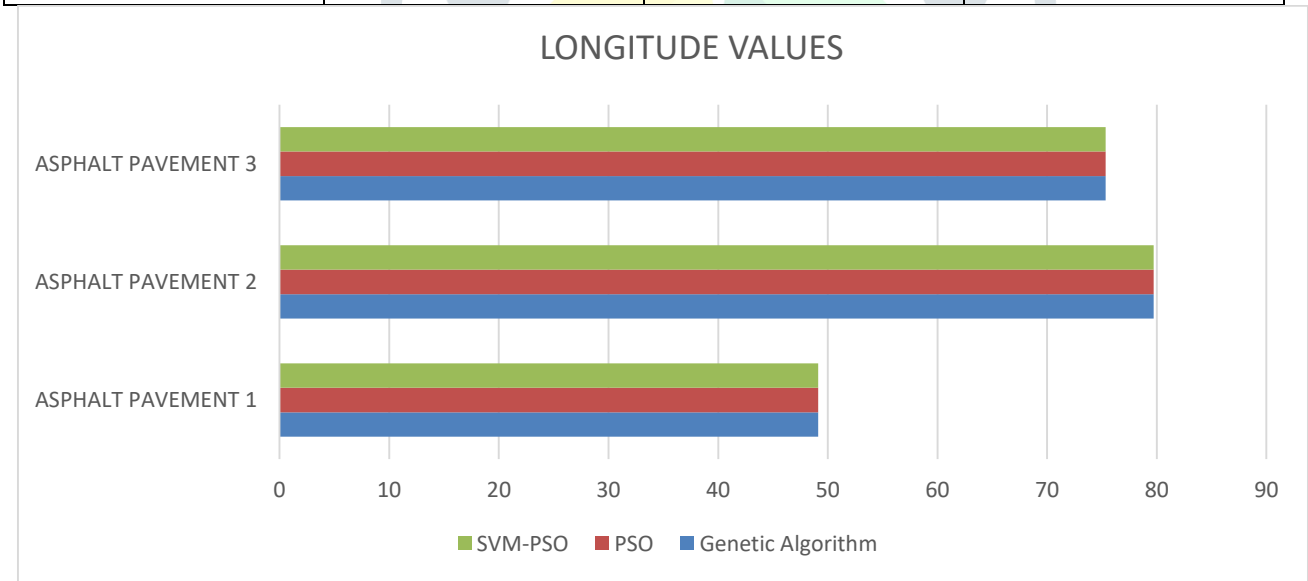


Figure 3.2: Traffic congestion on different algorithm

In figure 3.2 depicts the power generator values on the different algorithm approaches. The x-axis represents the algorithms and y-axis represents the values of the generator. The SVM- PSO gives the effective power generator values.

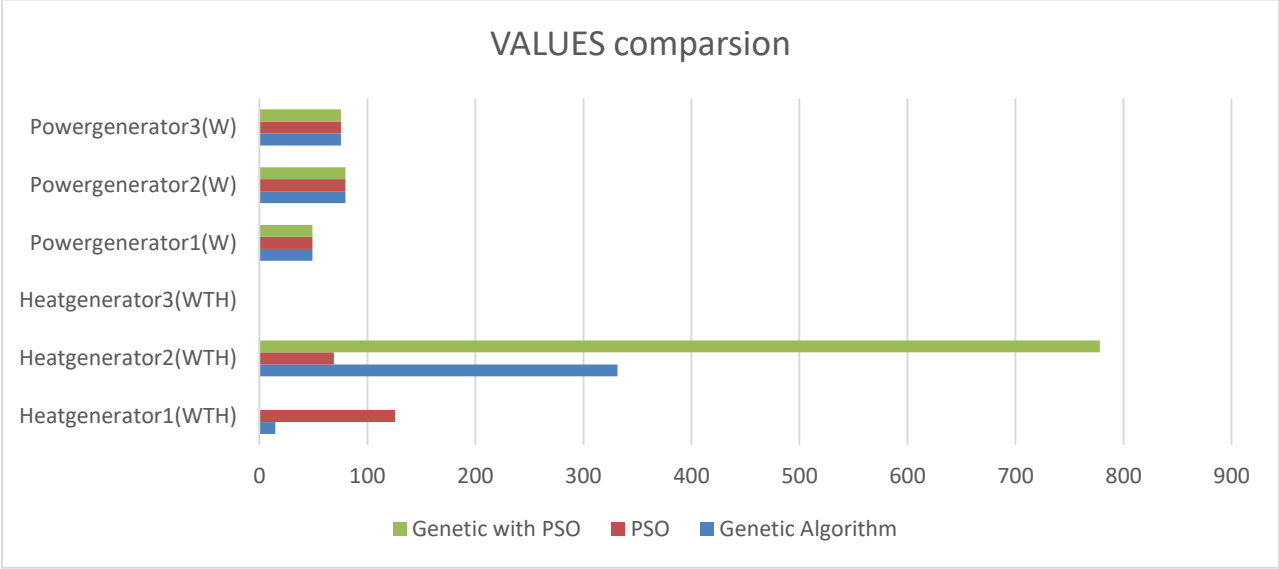


Figure 3.3: Generators Comparison

The figure 3.3 shows the comparison of the heat and power generators on the different algorithms. Here x-axis shows the values and y-axis shows the heat generators and power generators.

Table 3.3: Results on different parameters

Approaches	ASPHALT PAVEMENT1_**Cost(Rs)**	ASPHALT PAVEMENT1**cost(Rs)**	ASPHALT PAVEMENT**Total cost(Rs)**
Genetic Algorithm	725.09	1002.59	1727.68
PSO	625.03	472.74	1097.77
ANN-PSO	515.2	320.13	835.33

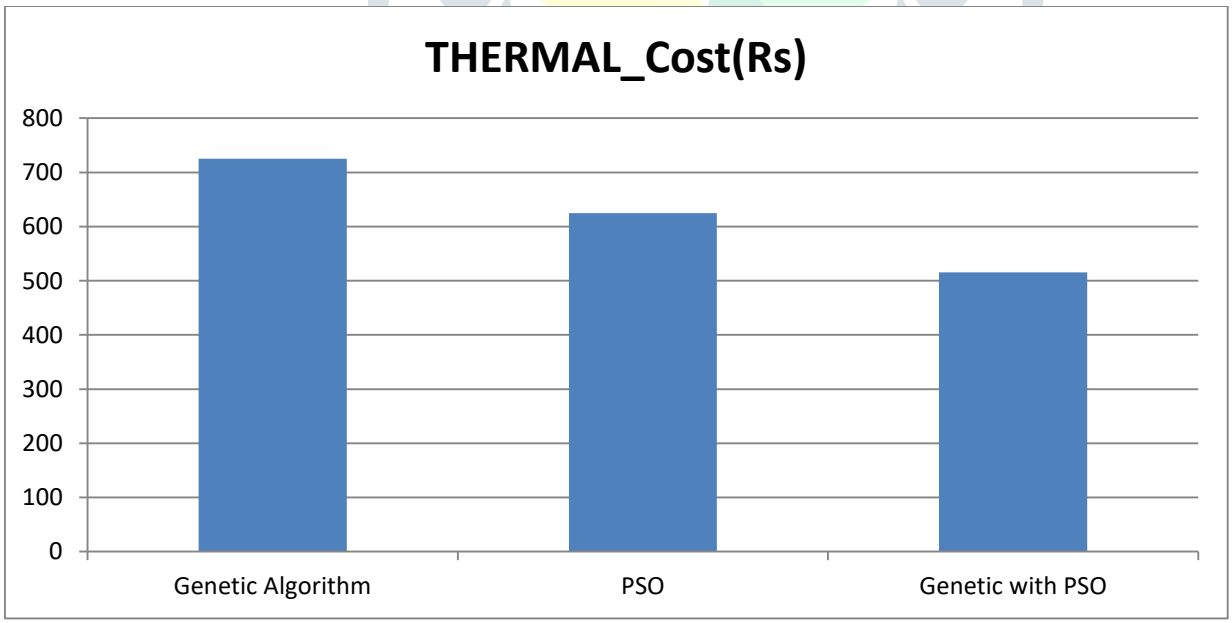


Figure 3.4: Heat_Cost on Different Algorithms

The figure 3.4 depicts the cost of heat in the different algorithms. In this graph x-axis represents the algorithmic approach and y-axis shows the value of cost.

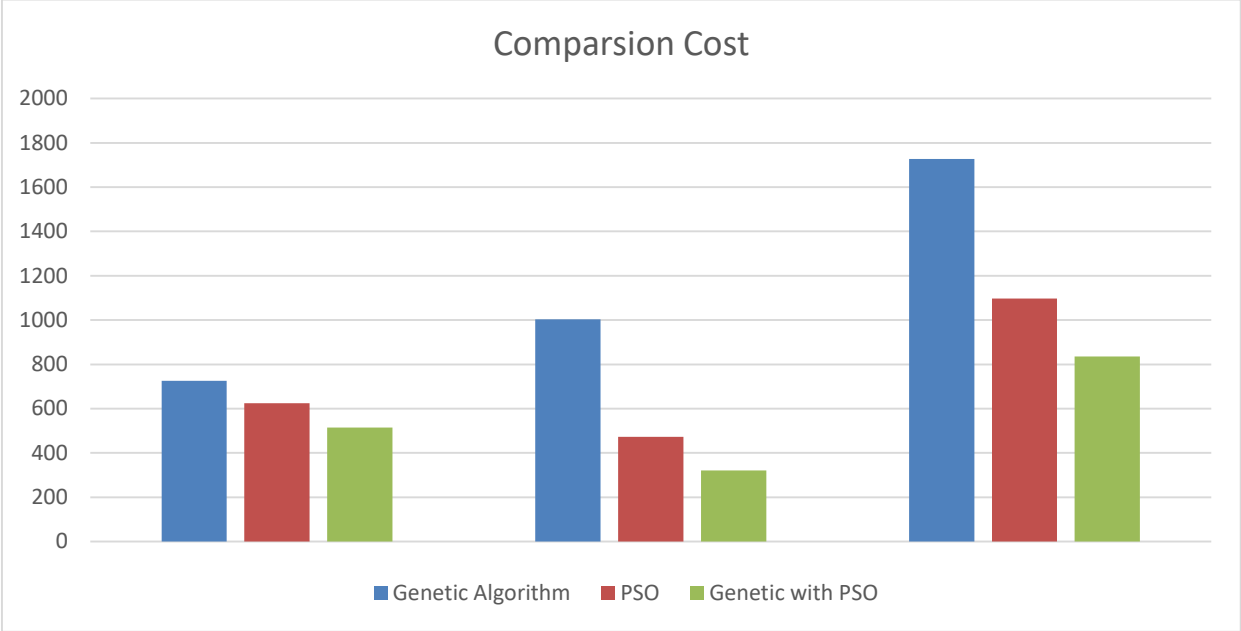


Figure 3.5: Cost Comparison on different algorithms

In figure 3.5 it depict the values of three algorithms that are Particle Swarm Optimization, G.A and Genetic Algorithm with Particle swarm optimization. The Blue bar represents the cost of Genetic algorithm, Red bar represents the Particle Swarm Optimization (PSO) and Green represents the proposed approach Genetic with PSO. The graph clearly describe the total cost is maximum on Genetic Algorithm and minimum on Genetic with PSO it is due to parallel working of both algorithm.

IV. CONCLUSION

Methodologies have been proposed and tested for the purpose of near real-time traffic forecasting, simulation and management. The forecasters are based on a novel spatially distributed ANN architecture that use iterative learning to adapt to the ever-changing urban environment and an adaptive ANN based filter to delete outliers in training data for more accurate forecasts. Four ANN techniques have been implemented and studied for the purpose of forecasting traffic flow: feed-forward backpropagation, cascade forward back-propagation, radial basis, and generalised regression ANNs. Furthermore, an adaptive ANN based training data filter has been introduced and tested. All methods have been tested over 13 months of real data from 20 roads in the city of Leicester, UK. A large number of experiments have been conducted to test the suitability of four distinct ANN techniques to forecast traffic flow conditions. The feed-forward back-propagation ANN obtained the best results closely followed by the cascade forward back-propagation ANN. The generalised regression ANN on the other hand obtained more consistently very good results. The radial basis ANN also provided good and consistent forecasters, but have not achieved the same level of forecasting as the other methods. The general regression ANN showed to be an excellent technique to build forecasters, given the right spread was chosen. However, the feed-forward back-propagation ANN in combination with the adaptive ANN based filter showed to be performing best while being relatively independent of the number of hidden nodes. This methodology is the authors' recommendation in the application of traffic flow forecasting. The introduced adaptive ANN based filter is capable of detecting outliers. The feed-forward back-propagation ANN as well as the cascade forward back-propagation ANN benefit heavily from using the adaptive training data filter. When applying the adaptive training data filter, the mean error of the feed-forward back-propagation ANN reduced on average by more than 48% and the cascade forward back-propagation ANN by more than 43%. The filter also consistently improved the radial basis and generalised regression ANNs performance although not to the same level. The statistical analysis of a large number of experimental tests confirmed the adaptive ANN based filter methodology to provide significantly better forecasters for all tested ANNs.

V. REFERENCES

1. Retrieved from www. publikasiilmiah.com on September 2014.

2. 2.Retrieved from https://ascelibrary.org on 6 June 2019.
3.Retrieved from www.semanticscholar.org on 1Feb 2011.

6. Wang, R., Li, Y. and Work, D.B., 2017. Comparing traffic state estimators for mixed human and automated traffic flows. *Transportation Research Part C: Emerging Technologies*, *78*, pp.95-110.

7. Dhyani, R. and Sharma, N., 2017. Sensitivity analysis of CALINE4 model under mix traffic conditions. *Aerosol and Air Quality Research*, *17*(1), pp.314-329.

8. Fountoulakis, M., Bekiaris-Liberis, N., Roncoli, C., Papamichail, I. and Papageorgiou, M., 2017. Highway traffic state estimation with mixed connected and conventional vehicles: Microscopic simulation-based testing. *Transportation Research Part C: Emerging Technologies*, *78*, pp.13-33.

9. Ahmad, A. and Rastogi, R., 2017. Regression model for entry capacity of a roundabout under mixed traffic condition–an Indian case study. *Transportation letters*, *9*(5), pp.243-257.

10. Beura, S.K., Chellapilla, H. and Bhuyan, P.K., 2017. Urban road segment level of service based on bicycle users' perception under mixed traffic conditions. *Journal of modern transportation*, *25*(2), pp.90-105.

11. Dhyani, R., Sharma, N. and Maity, A.K., 2017. Prediction of PM2. 5 along urban highway corridor under mixed traffic conditions using CALINE4 model. *Journal of environmental management*, *198*, pp.24-32.

12. Duret, A. and Yuan, Y., 2017. Traffic state estimation based on Eulerian and Lagrangian observations in a mesoscopic modeling framework. *Transportation research part B: methodological*, *101*, pp.51-71.

13. Fountas, G., Sarwar, M.T., Anastasopoulos, P.C., Blatt, A. and Majka, K., 2018. Analysis of stationary and dynamic factors affecting highway accident occurrence: A dynamic correlated grouped random parameters binary logit approach. *Accident Analysis & Prevention*, *113*, pp.330-340.

14. Li, T., 2018. Modeling uncertainty in vehicle trajectory prediction in a mixed connected and autonomous vehicle environment using deep learning and kernel density estimation. In *The Fourth Annual Symposium on Transportation Informatics*.

15. Jaikumar, R., Nagendra, S.S. and Sivanandan, R., 2017. Modal analysis of real-time, real world vehicular exhaust emissions under heterogeneous traffic conditions. *Transportation Research Part D: Transport and Environment*, *54*, pp.397-409.

16. Seo, T., Bayen, A.M., Kusakabe, T. and Asakura, Y., 2017. Traffic state estimation on highway: A comprehensive survey. *Annual Reviews in Control*, *43*, pp.128-151.

17. Zhang, W., Qi, Y., Yan, Y., Tang, J. and Wang, Y., 2017. A method of emission and traveller behavior analysis under multimodal traffic condition. *Transportation Research Part D: Transport and Environment*, *52*, pp.139-155.

18. Zhao, Z., Chen, W., Wu, X., Chen, P.C. and Liu, J., 2017. LSTM network: a deep learning approach for short-term traffic forecast. *IET Intelligent Transport Systems*, *11*(2), pp.68-75.