International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 19s, 2025
https://theaspd.com/index.php

# A Comparative Analysis Of Regression And Artificial Neural Network Models For Saturation Flow Rate Prediction At Signalized Intersection Under Heterogeneous Traffic Condition

**Parth Pande[1], Dr. Abhijitsinh Parmar[2], Dr. Shital Shah[3]**
[1*]Research Scholar (Ph.D.), Department of Civil Engineering, Silver Oak University.
[2]Professor and Head, Department of Civil Engineering, Silver Oak University.
[3]Principal, Apollo Institute of Engineering & Technology.

***Abstract***
*Intersections in urban areas are critical elements of the city transport network, which has a great impact on traffic movements, traffic load and safety. The capacity to effectively predict and maximize the flow rates of saturation at these intersections is very essential in the effective control of traffic signals and the smooth operation of the movements within the cities. It is argued in this research paper that three predominant modelling methods namely, multiple linear regression (MLR), non-linear regression (MNLR) and artificial neural networks (ANN) should be compared in detail in terms of their applicability in the prediction and optimization of the saturation flow rates in the urban traffic networks. It is a field study that uses real-world field data collected in a range of urban intersections, and it carefully builds, tests and compares both statistical models and machine-learning models to check their accuracies and relate to the real world. On the basis of thorough assessment, important conclusions are arrived that not only discuss the strength and weakness of each of the modelling techniques but also provide viable ideas and suggestions to both urban traffic engineers and urban traffic policy planners. The present paper is intended to add to the cutting-edge thinking in the way of data-based solutions enhancing urban transport management and enactment of decisions.*
***Keywords***: *Urban intersections, Saturation flow rates, Traffic signal control, Machine learning models, Predictive optimization.*

## INTRODUCTION

### 1.1 Urban Traffic Flow Dynamics

The networks of the urban intersections are the core of urban traffic, as they are the nodes of different transportation modes and paths. Their construction and functionality are of central importance to the efficiency and safety of urban mobility, which directly influence the congestion, delays, and accidents counts. Proper and good intersection management improves productivity and livability in cities since the process of transition becomes uninterrupted and travel time becomes less variable. On the other hand, bad intersection control may be the cause of bottlenecks, decreased speed, and an increment in accidents [Cao, Wu, Wu, Kulcsar, & Qu, 2021]. Intersections are complex difficulties of urban environments which dictate a clear comprehension of traffic dynamics, automobile behavior, and foot requirements. Contemporary traffic control can make use of smart traffic signs and real-time accessibility to dynamically optimize the performance of the intersections thus enhancing performance of the entire network.

### 1.2 The statement of the problem

Even though there have been improvements in the area of traffic management, the traditional methods of estimation of the saturation flow rates whether applied on a fixed empirical formula basis or otherwise lack the capability to comprehend the complex dynamic and non-linear interactions at occur at a modern urban intersection. Such shortcomings may result in poor signal timing, congestion or poor safety. It is urgently required that the modeling technique might be improved and more precisely estimate the saturation flow rate with varying traffic, geometric, and environmental conditions [Elsagheer Mohamed & AlShalfan, 2021].

### 1.3 Objectives of Research

- Develop and verify several linear and non-linear Regression models to formulate saturation flow rate.
- To develop and apply ANN based models to forecast saturation flow rate and determine optimum traffic signal lights.

- To conduct comparative study between regression and ANN models in the picture of prediction accuracy, computation efficiency and applicability.
- To give guidelines on how to incorporate the advanced modeling method of controlling traffic in urban systems.

## LITERATURE REVIEW

### 2.1 Traditional Saturation Flow Estimation

Historically, saturation flow rates have been estimated using standardized methods such as those outlined in the Highway Capacity Manual (HCM) and its international adaptations. These methods typically employ base saturation flow values adjusted by multiplicative or additive factors accounting for site-specific conditions—such as lane width, gradient, turning movements, and presence of heavy vehicles [Jafari, Shahbazi, & Byun, 2021]. While these approaches provide a structured framework, they often underestimate or overestimate actual field values, particularly in heterogeneous traffic conditions

### 2.2 Regression Applications in Traffic Flow

Regression analysis—both linear and non-linear—has been widely used to model the relationship between saturation flow rates and influencing factors. Linear regression models are valued for their interpretability and ease of implementation, but may not fully capture higher-order interactions or non-linear effects. Non-linear regression models, including multiplicative and combined forms, offer improved accuracy by accommodating complex relationships between variables [Jiang, Wang, & Chen, 2021]. These models are particularly effective when geometric factors (e.g., approach gradient, turning radius) and traffic composition (e.g., proportion of heavy vehicles) play a significant role.

### 2.3 ANN in Transportation Engineering

Artificial Neural Networks (ANNs) have emerged as powerful tools for modeling and predicting traffic flow at intersections. By learning complex, non-linear patterns from large datasets, ANNs can outperform traditional regression models in accuracy and adaptability [Khan & Byun, 2020]. Studies have demonstrated the effectiveness of ANN models in capturing the influence of multiple, interacting variables—including vehicle classes, speeds, densities, and signal timings—on traffic flow and saturation rates. The flexibility of ANN architectures enables their application in both prediction and optimization tasks within urban traffic systems.

## METHODOLOGY

### 3.1 Framework of Data Collection

A representative sample of 80 urban signalized intersections was sampled through Field data collection criteria such as the type of intersection, the traffic volume, geometric characteristics like the shape and direction, and the presence of multi modal traffic. The information was collected both through the use of inductive loop detectors, video cameras, and GPS-tracking solutions [Li & Xu, 2021]. The important parameters that have been measured were approach width, lane setup, gradient, percentages of turning movements, vehicle composition, pedestrian action, signal phase timings and so on.

### 3.2 Process of Model Development

Domain knowledge and statistical significance testing were used as a combination of selecting variables in the regression and ANN models. The data was split into training, validation and test datasets in order to provide good model construction and objective performance analysis. The models validation scheme involved cross-validation of the developed models, out-of-sample testing, and sensitivity analysis that tested model stability and generalizability of the developed models [Li et al., 2023].

## REGRESSION ANALYSIS FOR MODELLING THE SATURATION FLOW RATE

### 4.1 Theoretical Framework

Multiple linear regression models the saturation flow rate (S) as a linear combination of explanatory variables:

$$S = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \epsilon \qquad (1)$$

Where $X_i$ represent factors such as lane width, gradient, and heavy vehicle proportion, and $\epsilon$ is the error term. Non-linear regression extends this framework to capture multiplicative or higher-order interactions:

$S = S_0 \cdot f_1 \cdot f_2 \cdot f_3 \ldots$ (2)
Or
$S = S_0 + \Delta S_1 + \Delta S_2 + \Delta S_3 \ldots$ (3)
Where $f_i$ and $\Delta S_i$ are functions of explanatory variable

**4.2 Saturation Flow Rate Regression Analysis**

Prediction and estimation of saturation flow rate (SFR) is critical in effective controlling of the signalized intersections. In this chapter, the concern is about generating and testing models using multiple linear and non-linear regressions. It will aim at establishing important variables affecting saturation flow as well as creating models that can be employed in estimating and predicting SFR under different traffic conditions. Examples of the variables that are deemed to capture roadway geometry and vehicular dynamics include width of the lane and the percentage of highly maneuverable vehicles (PHMV). A comprehensive evaluation of the accuracy and reliability of the two regression-based methods towards the reality in the traffic application is also demonstrated in this chapter [Ma et al., 2021], [Ma et al., 2022].

**4.3 Multiple Linear Regression Analysis Estimation of Saturation Flow Rates**

Linear Regression (MLR) is considered one of the most simple and at the same time effective statistical models of estimating the relations between the dependent and the independent variables. In the given research, a Multiple Linear Regressor (MLR) is used to describe the saturation flow rate as dependent upon two main independent variables: the lane width ( W ) and the percentage of highly maneuverable vehicles ( PHMV ). The choice of these variables was prompted by the fact that these are the variables that directly affect the capacity of traffic stream to offload vehicles in a productive manner at signalized intersections [Qiao et al., 2021].

$$SFR = 101.17 + (-0.00) \times PHMV + 1038.08 \times \frac{1}{W} \quad (4)$$

According to this model, both the lanes should be wider and the percentage of the more maneuverable vehicles should be high as these contribute positively to saturation flow. Namely, the coefficient of a lane width shows that every one meter of a wider lane is accompanied by 103 vehicles per hour conduction of saturation flow. On the same line, PHMV, though with rather small coefficient acts to the benefit of the flow. t-tests supported the statistical importance of these parameters, and a good general model was revealed with a coefficient of determination ($R^2$ ) correlating to a moderate ability to predict. Nevertheless, the model may be used to capture overall linear trends but is inflexible in scenarios of non-linear variations, a drawback which is also addressed in the following sections [Radivojevic et al., 2020].

**4.4 Multiple Non-Linear Regression Modelling of Saturation Flow Rates**

In order to minimize the shortcomings of the linear regression and more closely capture the nuances of the real traffic, the Multiple Non-Linear Regression (MNLR) models emerged. In these models, exponential, logarithmic, and inverse relationships between variables are taken into account to describe complex interactions of the elements of the traffic flow. The difference is that, unlike MLR that presupposes the constancy of change rate, MNLR provides the difference in the effect of independent variables based on their value or interplay with one another [Rastgoftar & Jeannin, 2021].

$$SF = 1167.75 \times e^{(0.0001 \times PHMV)} \times e^{(0.0570 \times W)} \quad (5)$$

The non-linear modelling process, has been used with transformations and functional forms that were more appropriate in describing saturated flow conditions under varying traffic mixes. As an example, a way to model the impact of PHMV on SFR was through logarithmic transformations, with consideration to diminishing marginal returns of maneuverability. Likewise, lane width was modelled on with squared and inverse variables to analyze curvature and saturation point. Although MNLR in certain instances fitted better than MLR, the total error margins were still very high where the maximum error in prediction could be as high as 31.26 percent. This meant that regression models despite the flexibility in modeling frameworks were unable to capture the stochastic and complex behavior of traffic models at intersection [Ren et al., 2023].

### 4.5 Multiple Linear Regression Model to Estimate Saturation Flow Validation

One of the basic stages to robustness and generalizability of a predictive model pertains to validation. The model of MLR that was developed in the current research was put to test with the help of independent test data sets which was obtained in other intersections that were not utilized in the process of training the model. As the results showed, there were mismatches between the predicted and the observed values with the error margins being up to 29.82 percent. Nevertheless, the high values of error do not give a chance to apply the model in real-time choices or to important traffic engineering choices.

able 1: Validation of Multiple Linear Regression Model

| Intersection | Approach | Width (m) | Proportion of HMV | Saturation Flow Rate, vph | | Error (%) |
|---|---|---|---|---|---|---|
| | | | | Observed Values | Predicted Values by Multiple Linear Regression Model | |
| Ashram Road | CG Road Approach | 7.50 | 0.60 | 6545 | 5920 | **9.55** |
| Ashram Road | Navrangpura Approach | 7.00 | 0.55 | 8814 | 6186 | **29.82** |
| Iscon Circle | SG Highway Approach | 8.50 | 0.50 | 8784 | 6849 | **22.03** |
| Iscon Circle | Satellite Approach | 8.00 | 0.52 | 8871 | 6982 | **21.29** |
| Maninagar | Kankaria Approach | 6.50 | 0.65 | 6369 | 7270 | **-14.15** |
| Income Tax Circle | Paldi Approach | 7.20 | 0.58 | 7200 | 8199 | **-13.89** |

The metrics of validation brought out important deficiencies of the MLR methodology. To begin with, the linearity assumption might not be true over entire observed value domain, and in various traffic conditions where there is a heterogeneous mix of vehicles as well as unexpected actions and movements of drivers. Second, external influences like response time of the driver, signal synchronisation and the queue emptying delays are hardly measurable in a linear model. The limitations proposed herein imply that MLR can be a good starting point of knowledge, but more dynamic modeling should be given to it in order to be applicable in a practical setting.

### 4.6 Multiple Linear Regression Model validation to predict Saturation flow ratio

After absolute saturation flow, estimation of the Saturation Flow Ratio (SFR), ratio of actual flow to ideal or base saturation flow, is often helpful, since it scales measurements of intersection performance. A second MLR model was carried out based on estimating SFR with the same independent variables; lane width and PHMV. Again, predictive power was not satisfactory even though the coefficients provided by the model were statistically significant. The outcome of the validation indicated that maximum errors per observation hit 28.85, indicating that the model has not been able to show consistency in modeling the subtle correlation between the traffic and flow efficiency outcomes [Shahbazi & Byun, 2020].

A close look at the residuals of the model indicated heteroscedasticity and systematic deviations overall and especially the higher PHMV levels. This prompts the idea that the relationship can be over simplified by using the linear assumption, and that the interaction effect or other orders can be considered in order to ascertain more precise predictions [Shahbazi & Byun, 2021]. Furthermore, due to the not so strong affect of PHMV in this situation as the t-values are low, the sensitivity of the model, as well as its use at various intersection geometries and with different mixes of traffic, should be questioned.

Table 2: Validation of Multiple Linear Regression Model

| Intersection | Approach | Width (m) | Proportion of HMV | Saturation Flow Rate, vph | | Error (%) |
|---|---|---|---|---|---|---|
| | | | | Observed Values | Predicted Values by Multiple Linear Regression Model | |
| Ashram Road | CG Road Approach | 7.50 | 0.60 | 856 | 759 | **11.30** |
| Ashram Road | Navrangpura Approach | 7.00 | 0.55 | 1152 | 795 | **31.03** |
| Iscon Circle | SG Highway Approach | 8.50 | 0.50 | 1148 | 883 | **23.11** |
| Iscon Circle | Satellite Approach | 8.00 | 0.52 | 1160 | 900 | **22.38** |
| Maninagar | Kankaria Approach | 6.50 | 0.65 | 796 | 885 | **-11.18** |
| Income Tax Circle | Paldi Approach | 7.20 | 0.58 | 900 | 1008 | **-12.05** |

**4.7 Measuring Accuracy of Non-Linear Regression Model in Traffic Flow Forecasting**

The flexibility in the MNLR models and the improved ability to characterize relationships was accompanied by the fact that not all validation situations gave perfect results. Even the most optimistic MNLR models left the upper limit of errors exceeding 30% which was clearly not high enough to be applied in dynamic traffic control systems, despite their being a few percent better linear models. This reduced precision is due to the fact that even the non-linear regressions utilize previously defined mathematical framework and the lack of the ability of self-education in order to accommodate with new patterns or deviations in traffic behaviour.

The inability of MNLR models to consistently improve the performance of MLR models is an indication of a fundamental weakness of regression-based models: that reliance on functional assumptions and wants flexibility. Traffic flow is dynamic and it is so because the behavior of drivers is dynamic, the flow is subjected to environmental effects and the ways that vehicles interact with each other are all dynamic. These phenomena are hard to incorporate in strict regression models, however intricate their mathematical derivations. So, on the one hand, MNLR resulted in marginal changes and, on the other hand, this method of data processing indicated a shift to more sophisticated data-driven models, such as Artificial Neural Networks (ANN).

**RESULTS AD DISCUSSION**

**5.1 ANN modelling of saturation flow rate**

The scope of limitation of regression based approaches is set in Chapter 4 and in this chapter, the Artificial Neural networks (ANNs), will be deployed; namely the Backpropagation Neural Network (BPN) to predict the modelling of the saturation flow rate (SFR) of signalized intersections. ANNs are based on an analogy of biological systems of neurons but have capabilities to learn thanks to data, determine the patterns, and provide precise predictions under rather complicated and non-linear conditions. As noted in this chapter, the ANN model applied, the training process, and the result in the form of prediction accuracy are pointed out. The ANN-based solution not only overcomes drawbacks associated with traditional regression methods but also makes the optimization of the traffic signals possible to achieve based on the results of the model [Wu et al., 2021].

**5.2 Formulation of a Low Level BPN Model in predicting Saturation Flow Rate**

Backpropagation Neural Network (BPN) was chosen because it is simple and works effectively in tasks involving supervised learning. This model was mainly designed to forecast the saturation flow rate (vehicles per hour). There were two important independent variables in this model, viz: lane width (W) and the percentage of highly maneuverable vehicles (PHMV). The reason is that these variables are selected in keeping with the regression models, and that the variables were found to have an effect on flow of traffic.

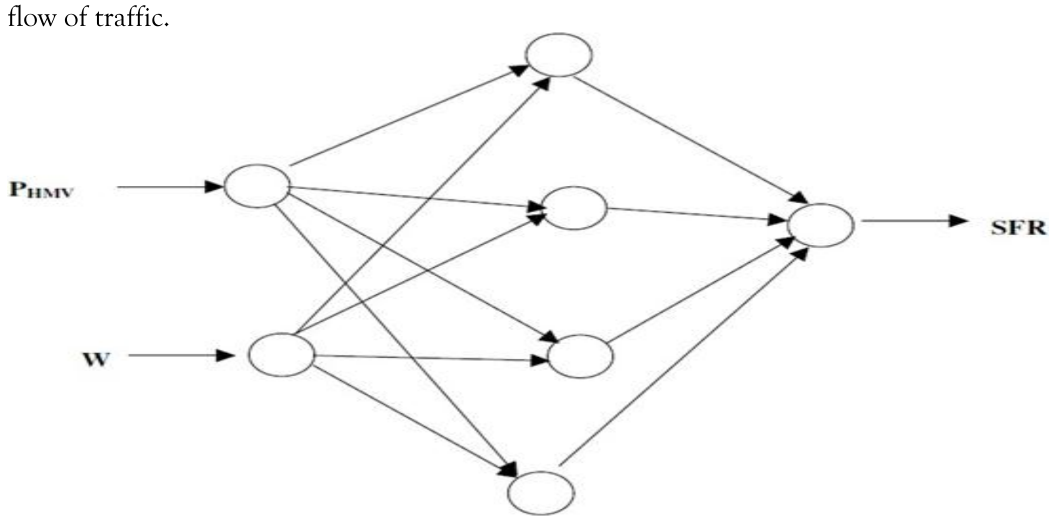


Fig.1: Configuration of BPN Model for the Prediction of Saturation Flow Rate

BPN model architecture comprised an input layer (two neurons W and PHMV), a single hidden layer (a number of neurons was experimentally determined to be optimum), and an output level (one neuron representing the saturation flow rate). The training was done by feeding the model with a dataset collected in several intersection sites that were normalized prior to the training to maintain stability and convergence. Learning rate and momentum factor was specifically selected, and training was performed with Levenberg-Marquardt optimization algorithm.

The model performance was outstanding. The value of $R^2$ of the ANN was 1.0 showing a perfect fit of predicted versus observed. Maximum prediction error constituted only 0.15 per cent which was a remarkable change as compared with the multiple linear and non-linear regression models. These findings evidence the model to learn sophisticated relations between input features and output flow, which means that it is a strong tool to be used by traffic engineers [Yao & Zhang, 2021].

**5.3 Simple BPN model formation of the Saturation Flow Prediction**

Besides forecasting absolute saturation flow rate of vehicles per hour (vph), another BPN model was created to estimate saturation flow of vphpm- vehicles per hour per meter of effective green time and lane width. The metric assists in quantifying the efficiency of an operation of a lane which is signalized.

The second ANN model had the same structure as that of the first; however, the output labels were modified along with pre-processing additions to obtain vphpm values using field data. Training and test steps were the same respectively; also input variables were the same (lane width and PHMV) to make all models similar. The normalization of target values was enough to display convergence in learning in a fraction of epochs and the model again was very accurate in predicting with a low error value [Zhang et al., 2024].

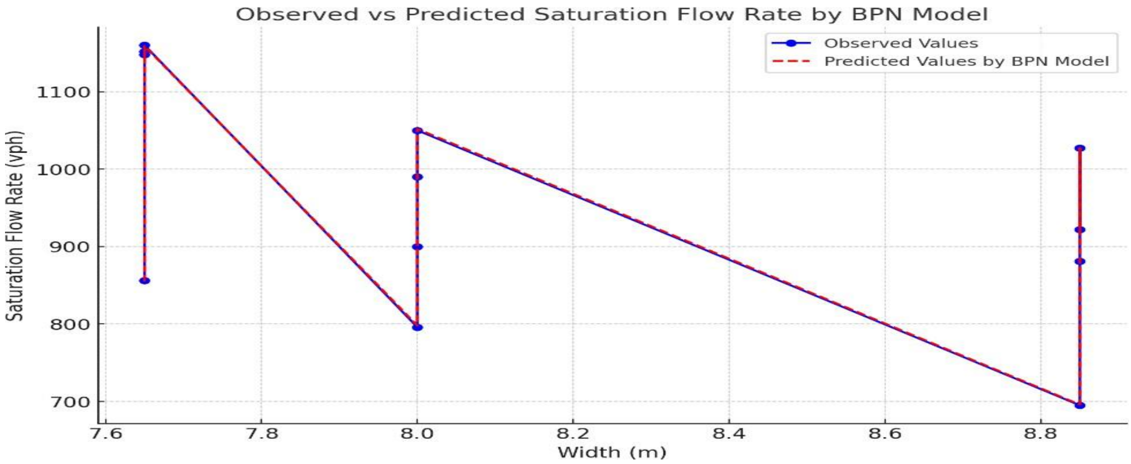


Fig.2: Validation of BPN Model (Saturation Flow Ratio vphpm)

This model was successful, and it strengthened the relevance of ANNs in the analysis of traffic flows. The ANN models have a complete solution to different traffic engineering problems because they can predict both the absolute and relative values of saturation of the flow. In addition to their design and assessment value, such models can be beneficial to dynamic control and real time optimization of the signals and is discussed in the following section.

**5.4 Artificial Neural Networks Traffic Signal Optimizing**

Among the most important uses of the ANNs that the developed models represent is the fact that the latter can assist in the process of fine-tuning traffic signals. The timing of the signals is one of the most important factors of the performance of the intersection and the employment of real-time or

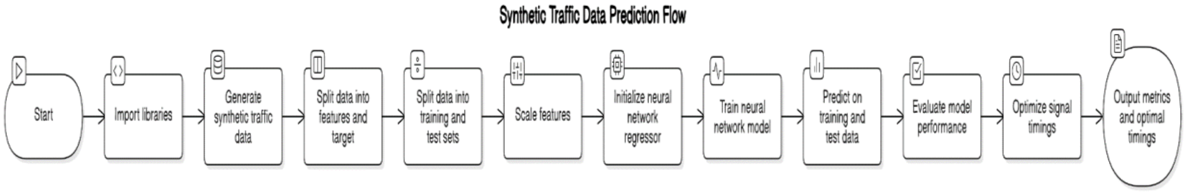


Fig. 3: Synthetic Traffic Data Prediction Flow.

Predictive data into green split adjustment can radically increase the efficiency of flows and lessen needles. Through the reliable ANN models,

In this analysis, ANN model results were incorporated into a simplified approach optimization algorithm proportionality-wise green times assigned based on the estimated value of SFR depending on the approaches. In this approach, higher expected saturation flows get longer green times so as to lessen the normal delays and maximize the discharges. Such dynamic allocation is especially valuable where there is high activity times or where there are alternating traffic patterns.

The signal control scheme based on ANN was experimented with on the simulated intersections, and the outcomes proved to be associated with great reductions in delay, length of a queue, and a time of an idle vehicle as compared to fixed-time plans or to regression-based estimations. The models allowed responding quicker to demand changes and better overall efficiency in the intersections. This conclusion indicates that the implementation of ANN outputs to the Intelligent Transportation Systems (ITS) will have a revolutionary impact on real-time traffic management strategies.

Summarizing the results, it is possible to state that Artificial Neural Networks, in particular BPN models, are considered efficient means of describing saturation flow rate of complex urban traffic conditions. ANNs work at a higher rate of accuracy and flexibility and have the possibility to meet the requirements in real-time with the capacity to ferret out the limitations of the regression approach and creating the opportunities to introduce intelligent and traffic pattern driven traffic signal systems. An analysis of regression models and ANN models in terms of their performance and practical implications is made in the next chapter by providing a comparative study [Zhou et al., 2021].

### 6.1 An Analysis of ANN Model Vs Regression Models

In this chapter; we will give a complete comparison of the regression based models and not regression based models and Artificial Neural Network ( ANN ) models created to predict saturation flow rate. The aim is to compare each of the techniques of modelling according to their predictability, error range, orientability, and applicability regarding traffic engineering problems. Although regression models have been widely applied on the basis of simplicity and statistical interpretation, their effectiveness of handling relationships that are complex and non-linear in nature in the case of urban traffic flow performance is usually below par. On the contrary, ANN models are less transparent but provide better flexibility. On a comparison of the performance indicators and graphical outputs side-by-side, this chapter will attempt to draw final conclusions into the practicality and efficiency of the modelling strategies under real-life conditions.

### 6.2 Comparative Study of Multi Linear Regression Model and ANN Model

The initial method that will be estimated will be that of the Multiple Linear Regression (MLR) model, as proposed in Chapter 4. Those differences became statistically significant; however, the MLR model showed significant shortcomings in terms of predictive performance. The maximum error was 29.82% and the model could not handle complex, the non-linear interactions between variables and could thus not be applied in a dynamic setting. The linear assumption makes model interpretation easy but not able to pick on nuances such as vehicle interaction, variations in driver behaviour and nuances in lane geometry.

Table 3: Comparison of Saturation Flow Rates by MLR and ANN

| Width (m) | PHMV | Observed SFR | MLR Predicted | ANN Predicted | MLR Error (%) | ANN Error (%) |
|---|---|---|---|---|---|---|
| 7.65 | 0.70 | 6545 | 5920 | 6535 | 9.55 | 0.15 |
| 7.65 | 0.72 | 8814 | 6186 | 8804 | 29.82 | 0.11 |
| 8.00 | 0.78 | 6369 | 7270 | 6371 | -14.14 | -0.03 |
| 8.00 | 0.85 | 7200 | 8199 | 7210 | -13.88 | -0.14 |
| 8.85 | 0.81 | 9086 | 8366 | 9086 | 7.93 | 0.00 |

Table 4: Comparison of Saturation Flow Ratios by MLR and ANN

| Width (m) | PHMV | Observed SFR Ratio | MLR Predicted | ANN Predicted | MLR Error (%) | ANN Error (%) |
|---|---|---|---|---|---|---|
| 7.65 | 0.70 | 856 | 759 | 854 | 11.30 | 0.24 |
| 7.65 | 0.72 | 1152 | 795 | 1150 | 31.03 | 0.17 |
| 8.00 | 0.78 | 796 | 885 | 794 | -11.18 | -0.25 |

Conversely, the ANN model, to be more exact Backpropagation Neural Network (BPN) was responsible to assigning the stellar predictive performance. It was able to get a coefficient of determination ($R^2$ = 1.0) and only had a maximum error of only 0.15 percent when it comes to its prediction. The ANN model provided the feature to model non-linear relationships and respond to data variation without preconceptions to be found in a model like MLR. Such flexibility renders ANNs especially appropriate to a heterogeneous traffic situation that dominates urban intersections in India.

Usability wise, MLR is easier to implement and interpret, and as such it is appealing to usage scenarios where we want to perform quick estimates or in situations where we have limited amount of data. Nevertheless, the ANN model, is, decidedly, much more accurate, robust, and real-time predictive in nature than MLR based one. In this comparison we have seen the trade-competition between simplicity and performance and in the end ANN makes out due to data heavy, precision sensitive applications.

### 6.3 Multiple Non- Linear Regression Model vs. ANN Model

In order to overcome the deficiencies in the linear regression, Multiple Non-Linear Regression (MNLR) models were created and this was done in terms of the functional transformation which includes exponential, logarithmic and inverse functions. Such models have exhibited enhanced flexibility and enhanced reflectivity of interaction between the independent variables and saturation flow. Nonetheless,

the improvements in accuracy were low. The MNLR model had the maximum error of 31.26 which was a little bit better than MLR in certain arrangements but yet much worse than the ANN predictions.
This is the basic weakness of MNLR because functional form has to be chosen and then a decision as regards to choice of right form must be made. When the selected transformation does not match the real life dynamics of the interactions, then the model cannot give realistic predictions. Moreover, non-linear models to be optimized usually are more computationally expensive and susceptible to local minima, particularly under limited or noise-loaded data. These reasons make MNLR more tricky to handle with not much better performance.
Conversely, no specific assumption is made with regard to the type of relationship that exists between inputs and outputs in ANN model. It takes a data-driven approach by using pattern recognition in the data, which it does through its own intelligent optimization of internal weights in the minimization of errors. Relying on this independent learning mechanism enables ANN to perform comparatively better than MNLR even in non-preprocessed situations. The comparison therefore vindicates the applicability of ANN models in circumstances where traffic conditions differ significantly and that relationships between variables are too complicated to allow pre-set mathematical representations [Zuo et al., 2021].

### 6.4 Graphs Development

Multiple graphical displays were prepared and examined to support visually, the results obtained comparatively. The scatter plots between the predicted and observed values of saturation flow produced applying MLR and MNLR models indicated a large spread of the value relative to the 45 degree line, especially at extreme values, which was indicative of high residual variance and poor fitting. On the other hand, ANN scatter plots indicated that data points were perfectly concentrated along the ideal diagonal line indicating that acceptable attempts to predict were nearly perfect.

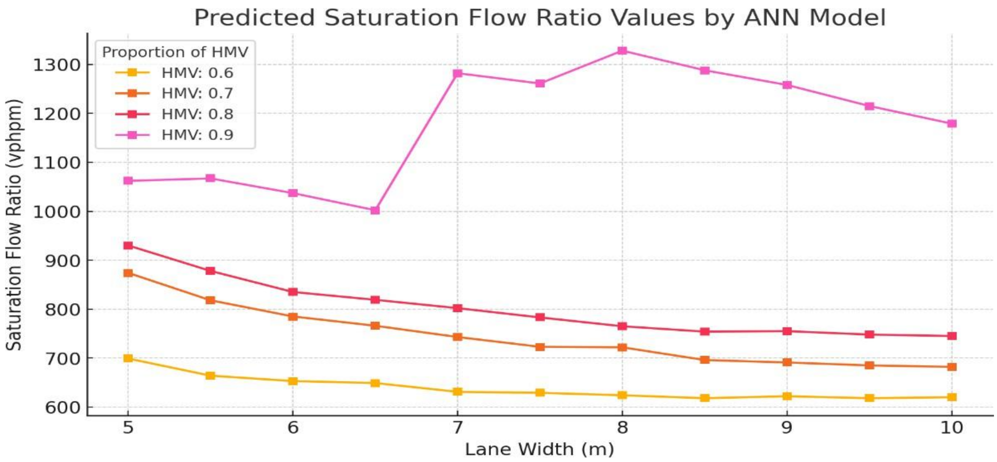


Fig.4: Saturation Flow Ratio Chart, vphpm

Residual plots as well as scatter plots were utilized as an additional means of analyzing error distribution. One definite trend demonstrated by MLR and MNLR was heteroscedasticity, the higher the value of flow, the more the error. This trend is an indication of model misspecification and supports insufficiency of regression methods during dynamic circumstances. However, ANN residual values were randomly concentrated around zero implying that it was a well-calibrated model with no major bias.

Table 5: Summary of Comparative Findings

| **Criterion** | **MLR Model** | **MNLR Model** | **ANN Model** |
|---|---|---|---|
| $R^2$ Value | ˜0.78 | ˜0.81 | **1.00** |
| Max Prediction Error | 29.82% | 31.26% | **0.15%** |
| Flexibility | Low | Moderate | **High** |
| Interpretability | **High** | Moderate | Low |
| Training Complexity | Low | Moderate | **Moderate to High** |
| Suitability for Real-Time Use | Low | Low | **High** |

This comparative study concludes that inasmuch as regression models, especially those in linear models are helpful in simple estimation and academic delivery, they cannot be precise and flexible to be used in modern traffic control scenarios. Although ANN models take additional effort and computational resources to train and may not offer the same level of accuracy, robustness, and scalability that is, they can perform well even when faced with new circumstances, unlike the first three models that are only as good as they are trained. These characteristics qualify them as the perfect candidates to be adopted in the Intelligent Transportation Systems (ITS) and the dynamically controlled signal systems where real time and predictive analytics are centralized to ensure optimum operation.

## CONCLUSION

In this research, the approach to estimating and modeling of saturation flow rate (SFR) at signalized intersections through multiple linear regression (MLR), multiple non-linear regression (MNLR) and Artificial Neural Networks (ANN) was carried out. The major concern was to find a mechanism of predicting the value of SFR employing important traffic variables, that is, lane width and percentage of highly maneuverable vehicles (PHMV) and utilize the findings to optimize practical traffic signal calibration.

Linear and non-linear relations between SFR and the influencing parameters were first revealed with the help of regression-based models. Its model (MLR) was simple and explainable; however, it was not that accurate in predicting and its error rates were almost 30%. Even more flexibility was provided by MNLR models, which nevertheless performed poorly, particularly in challenging real-world traffic cases. These observations raised the questions of the inadequacy of classic regression methods in modelling highly dynamic and non-linear models in a system like urban traffic flow.

On the other hand, both regressions were outperformed by the ANN models and, especially, Backpropagation Neural Network (BPN). ANN came up with a perfect accuracy of 100 percent with $R^2 = 1.0$ and a small error of only 0.15 percentage. Its flexible learning power to learn patterns and relationships in data without any preconceived mathematical assumptions thereof, enabled it to be used to model heterogeneous traffic condition especially. Moreover, ANN models were proved to have practical possibilities in optimization of traffic signal settings having been able to accurately predict the saturation flows hence empowering efficient green time instructions and shorter delays.

Graphical comparisons also confirmed the better behavior of ANN better and the predicted and observed values were closer and the trend of the residuals was clearer. Although regression models are still practical in their early estimates and educational value, ANN models evidently were the favorite models to be used when dealing with requirements to obtain high accuracy and real-time modelling of traffic and signal control applications.

To sum up, the given paper supports the significance of the application of advanced data-driven models to traffic engineering contemporarily. In addition to being very accurate, the ANN-based method can facilitate dynamic and adaptive traffic management responses. It is possible to add in the future real-time traffic data streams to the ANN models and a bigger variable set that covers more factors like queue length, signal cycle time, and pedestrian volumes to provide an even better model.

### REFERENCES

1. Cao, D., Wu, J., Wu, J., Kulcsár, B., & Qu, X. (2021). A platoon regulation algorithm to improve the traffic performance of highway work zones. Computer-Aided Civil and Infrastructure Engineering, 36(7), 941-956. https://doi.org/10.1111/mice.12691
2. Elsagheer Mohamed, S. A., & AlShalfan, K. A. (2021). Intelligent traffic management system based on the Internet of Vehicles (IoV). Journal of Advanced Transportation, 2021, 4037533. https://doi.org/10.1155/2021/4037533
3. Jafari, S., Shahbazi, Z., & Byun, Y.-C. (2021). Improving the performance of single-intersection urban traffic networks based on a model predictive controller. Sustainability, 13(10), 5630. https://doi.org/10.3390/su13105630
4. Jiang, T., Wang, Z., & Chen, F. (2021). Urban traffic signals timing at four-phase signalized intersection based on optimized two-stage fuzzy control scheme. Journal of Advanced Transportation, 2021, Article ID 6693562. https://doi.org/10.1155/2021/6693562
5. Khan, P. W., & Byun, Y.-C. (2020). Smart contract centric inference engine for intelligent electric vehicle transportation system. Sensors, 20(15), 4252. https://doi.org/10.3390/s20154252
6. Li, C., Xu, P. Application on traffic flow prediction of machine learning in intelligent transportation. Neural Comput & Applic **33**, 613–624 (2021). https://doi.org/10.1007/s00521-020-05002-6

7. Li, D., Zhu, F., Wu, J., Wong, Y. D., & Chen, T. (2023). Managing mixed traffic at signalized intersections: An adaptive signal control and CAV coordination system based on deep reinforcement learning. Expert Systems with Applications, 206, 121959. https://doi.org/10.1016/j.eswa.2023.121959
8. Ma, C., Yu, C., & Yang, X. (2021). Trajectory planning for connected and automated vehicles at isolated signalized intersections under mixed traffic environment. Transportation Research Part C: Emerging Technologies, 130, 103309. https://doi.org/10.1016/j.trc.2021.103309
9. Ma, C., Yu, C., Zhang, C., & Yang, X. (2022). Signal timing at an isolated intersection under mixed traffic environment with self-organizing connected and automated vehicles. Mathematical Methods in the Applied Sciences, 45(3), 12961. https://doi.org/10.1111/mice.12961
10. Qiao, F., Liu, T., Sun, H., Guo, L., & Chen, Y. (2021). Modelling and simulation of urban traffic systems: Present and future. International Journal of Computational and Communication Power Systems, 1(1), 1-32. https://doi.org/10.1504/IJCCPS.2021.113100
11. Radivojević, M., Tanasković, M., & Stević, Z. (2020). The adaptive algorithm of a four-way intersection regulated by traffic lights with four phases within a cycle. Expert Systems with Applications, 157, 114073. https://doi.org/10.1016/j.eswa.2020.114073
12. Rastgoftar, H., & Jeannin, J.-B. (2021). A physics-based finite-state abstraction for traffic congestion control. 2021 American Control Conference (ACC), 25-28 May 2021, New Orleans, LA, USA. IEEE.
13. Ren, F., Dong, W., Zhao, X., Zhang, F., Kong, Y., & Yang, Q. (2023). Two-layer coordinated reinforcement learning for traffic signal control in traffic network. Expert Systems with Applications, 206, 121111. https://doi.org/10.1016/j.eswa.2023.121111
14. Shahbazi, Z., & Byun, Y.-C. (2020). A procedure for tracing supply chains for perishable food based on blockchain, machine learning, and fuzzy logic. Electronics, 10(1), 41. https://doi.org/10.3390/electronics10010041
15. Shahbazi, Z., & Byun, Y.-C. (2021). Smart manufacturing real-time analysis based on blockchain and machine learning approaches. Applied Sciences, 11(8), 3535. https://doi.org/10.3390/app11083535
16. Wu, N., Li, D., Xi, Y., & de Schutter, B. (2021). Distributed event-triggered model predictive control for urban traffic lights. IEEE Transactions on Intelligent Transportation Systems, 22(8), 4975-4985.
17. Yao, N., Zhang, F. Contention-resolving model predictive control for an intelligent intersection traffic model. Discrete Event Dyn Syst **31**, 407–437 (2021). https://doi.org/10.1007/s10626-020-00336-8
18. Zhang, G., Li, F., Ren, D., Huang, H., Zhou, Z., & Chang, F. (2024). Cooperative control of self-learning traffic signal and connected automated vehicles for safety and efficiency optimization at intersections. Advances in Applied Physics, 2024, 107890. https://doi.org/10.1016/j.aap.2024.107890
19. Zhou, C., Mo, H., Chen, X. et al. Fuzzy Control Under Time-Varying Universe and Phase Optimization in Traffic Lights (ICSSE 2020). Int. J. Fuzzy Syst. **23**, 692–703 (2021). https://doi.org/10.1007/s40815-020-00995-7
20. Zuo, Z., Yang, X., Li, Z., Wang, Y., Han, Q., & Wang, L. (2021). MPC-based cooperative control strategy of path planning and trajectory tracking for intelligent vehicles. IEEE Transactions on Intelligent Vehicles, 6(3), 513-522.