ARTICLE | OPEN ACCESS

Indian journal of
Engineering

To Cite:
Akeke MU, Okafor FO, Nnaji CC, Ikeagwuani CC. Forecasting Urban Traffic Congestion using Bayesian-Driven Models. *Indian Journal of Engineering*, 2026, 23, e6ije1714
doi: https://doi.org/10.54905/disssi.v23i59.e6ije1714

Author Affiliation:
[1]Department of Civil Engineering, University of Nigeria, Nsukka, Enugu State, Nigeria
[2]Faculty of Engineering and Built Environment, University of Johannesburg, South Africa
[3]Department of Civil Engineering, Akwa-Ibom State University, Ikot-Akpaden, Akwa-Ibom State, Nigeria
[4]Department of Civil Engineering, Alex-Ekwueme Federal University, Ndufu-Alike, Ikwo, Nigeria; Orcid id: https://orcid.org/0000-0001-5037-552X; Email: chijioke.ikeagwuani@unn.edu.ng

[*]**Corresponding author:**
Akeke MU,
Department of Civil Engineering, University of Nigeria,
E-mail: mikkyukandi@yahoo.com

Peer-Review History
Received: 24 September 2025
Reviewed & Revised: 16/October/2025 to 12/April/2026
Accepted: 25 April 2026
Published: 07 May 2026

Peer-Review Model
External peer-review was done through double-blind method.

Indian Journal of Engineering
pISSN 2319–7757; eISSN 2319–7765

DISCOVERY
SCIENTIFIC SOCIETY

Forecasting Urban Traffic Congestion using Bayesian-Driven Models

Akeke MU[1*], Okafor FO[1], Nnaji CC[1,2], Ikeagwuani CC[1,3,4]

ABSTRACT

One of the challenges of urbanizing African cities in terms of mobility is traffic congestion. Congestion, long queues, and unpredictable travel times in Calabar, Nigeria, are a daily experience, especially in commercial areas and along major roads. In this study, congestion states are predicted using Gaussian Process Regression (GPR) and a Relevance Vector Machine (RVM), based on field data from 20 large corridors. The findings indicate considerable variation in congestion rates across places. Post hoc Tukey HSD tests indicate that congestion is intermittent and is instigated by localized pinpoints of congestion at the Harbor, Tinapa, Main Avenue, Goldie, and Ettagbor crossings. Ekpo Abasi, Uwanse, and Goldie, on the other hand, were doing fairly well regarding delay and queue length. Vehicle arrivals were the most important feature, with the strongest linear correlations with total delay (r = 0.669) and queue length (r = 0.858), and traffic demand showed a positive linear association with the volume-to-capacity ratio (r = 0.568). GPR was more effective than RVM in all the congestion indicators. GPR obtained R^2 values of 0.876 (total delay), 0.966 (queue length), and 0.735 (VCR) on the test set with well-calibrated 95% credible interval coverage (94.1%, 93.0%, 90.2%). The locations of congestion were detected as clusters in the spatial maps around markets, schools, and transport termini. These findings support congestion hotspot-specific interventions and endorse probabilistic modeling, particularly GPR, for traffic prediction in data-sparse African cities. This study provides evidence-based recommendations for transportation policy, route planning, and signal optimization in Calabar.

Keywords: Calabar, Gaussian, Learning, Machine, Regression, Relevance, Traffic, Vector

1. INTRODUCTION

Urban transport has been increasing at an alarming rate across the globe, with cities struggling with the challenge of incessant traffic snarls, soaring fuel prices, increasing car emissions, and falling living standards. Surprisingly, these problems have even led to the creation of strategies that offer more secure, efficient, and quicker means of transport (Ullah et al., 2025). This has led to a very significant area of research studying the development of models to predict traffic congestion within this larger domain of inquiry. It forms the basis of Intelligent

Transportation Systems partly due to the fact that it helps in route guidance as well as information in making traffic management decisions (Rahmani et al., 2024; Haghighat et al., 2024). With the increase in the significance of precise and dependable prediction, scientists have steadily created and enhanced an assortment of modeling techniques, all aimed at enhancing the dependability with which one can predict congestion. Historical developments in the development of methods encompass classical forms of statistics (such as the Kalman Filter and ARIMA), classical machine learning algorithms and the ensemble approach, evolutionary systems of rules, and, more recently, deep learning systems.

Over the last ten years, there has been a growing interest in probabilistic Bayesian models, such as Gaussian process regression and relevance vector machines. The waves of innovation have, in turn, improved the capability to deal with challenges in traffic forecasting: nonlinearity, spatial and temporal dependencies, the necessity to quantify uncertainty, and the variability of data across infrastructural settings. During the initial stages of this development, traffic prediction was primarily done by the use of classical statistical methods. These involve the Kalman filter, on which the estimation of the vehicle densities and the congestion states was done (Momin & Hamim, 2022; Attioui & Lahby, 2025; Kumar, 2017; Kashyap et al., 2022; Alghamdi et al., 2021), and the use of Autoregressive Integrated Moving Average (ARIMA) models in the prediction of the traffic volumes. It quickly became obvious that they are ill-adapted when it comes to capturing the non-linear, fast-changing dynamics of dense urban traffic (Toba et al., 2025).

Empirical data show that ARIMA models tend to be weak during traffic peaks, as the stationarity assumptions underlying them cannot capture the volatility of actual traffic flows. The same can be said of linear regression, which has been reported to struggle to forecast a sudden congestion event driven by oscillating demand. Moreover, in cities, the classical concepts of linearity and stationarity are often inapplicable, and complex interactions among drivers and non-homogeneous traffic flows are common (Khin, 2024; Pawel et al., 2025; Satria, 2021).

Satria (2021) employed a multi-layer perceptron (MLP) artificial neural network to create a model to forecast urban traffic congestion that relies on automatically preprocessed CCTV footage. This model estimates the number of cars, buses, and trucks, with a root mean square error (RMSE) of 1.88, and can be used to efficiently forecast short-term traffic. Although Support Vector Machines (SVMs) have shown a lot of potential to model nonlinear relationships with kernel functions, there are grave limitations to their functionality. Despite the solid theoretical foundation of SVM techniques in statistical learning, these methods are rather difficult to train and do not apply to large data sets. Also, SVMs are prone to local optima and poor performance since they are highly sensitive to hyperparameter tuning as well as the choice of kernel. These limitations involve a lot of preprocessing, such as feature scaling or sensitive parameter choices to avoid overfitting, when dealing with data that is highly noisy or overlapping classes (Attioui & Lahby, 2025). In addition, another research direction was evolutionary algorithm-based rule mining (Li et al., 2025). With the advent of deep learning, enabled by the proliferation, a significant increase in the capability of predicting congestion was observed in a large number of traffic sensors.

Long Short-Term Memory networks were shown to be able to capture a temporal dependence in traffic data with considerable effectiveness (Essien & Giannetti, 2020; Liu et al., 2017). Building on this, Hybrid CNN-LSTM architectures are highly effective because they utilize Convolutional Neural Networks (CNN) to automatically extract spatial features (such as topological relationships in networks or local patterns in images). Long Short-term Memory (LSTM) networks are used to learn both temporal dependence (i.e., sequence-based changes across time) and achieve improved spatiotemporal modeling performance. These shortcomings have, in recent years, led to a resurgence of the use of probabilistic Bayesian models, providing an alternative trade-off between accuracy of prediction, interpretability, and the possibility of measuring uncertainty (Bhattacharya et al., 2025). Much of the foundation for this line of inquiry was laid by Rasmussen and Williams (2006), who developed Gaussian Process Regression as a nonparametric Bayesian model. Liu et al., (2020) confirmed its predictive capability and showed that GPR's ability to provide uncertainty estimates makes it especially useful for strategic planning applications. Wang et al., (2021) and (Rodrigues & Pereira, 2018) proved that the capability of GPR to offer uncertainty estimates makes it especially useful in the context of strategic planning (Cui et al., 2020).

In a more recent study, Akinlana et al., (2024) also verified the performance of GPR by use of data fusion applications. Similarly, another probabilistic model suggested by Tipping (2001) is the Relevance Vector Machine. Of particular recent interest is a recent model by Yang et al., (2024) that recommended a customized RVM model which had RMSE minimization rates of 61.68% compared with GPR, 55.54% compared with SVM, 40.97% compared with artificial neural networks, and 14.00% compared with conventional RVM. Diker & Nasibov (2012) and Emmanuel & Mohammed (2025) have recently measured traffic congestion in cities in Nigeria, concluding that fast growth in the number of vehicles, without corresponding growth in the infrastructure, has created high levels of congestion; volume-to-capacity ratios of 0.65. They are reported and reflect a service C and they raise the probability of congestion. Kazaara, (2025) on the

problems of transport data in African cities, it was found that the counts of transport made by hand varied greatly, and the necessity to normalize the data to minimize the impact of outliers. Sulaiman et al., (2024), given these limitations in data availability and quality, it is apparent that there is a need to develop methods for detecting highly correlated traffic patterns within larger datasets. One method for generating meaningful subsets of data is clustering. This is especially appealing to Density-Based Spatial Clustering of Applications with Noise, as it automatically clusters based on density, replaces arbitrary-shaped clusters, and scales reasonably well (Lahijanian, 2024).

The current research was conducted in Calabar Metropolis, a location that represents many of the issues facing medium-sized African cities. Population growth in Calabar has been immense over the past decades, and so has the number of vehicles on the road and the unending congestion on major arterial roads, such as the Murtala Mohammed Highway, IBB Way, and Marian Road. Most importantly, the city does not have automated sensor infrastructure; manual counts, thus, represent the main source of available data. In this research, the data of manual traffic volumes were gathered in 20 strategic areas in the metropolis during 12 weeks. Peak hours in the morning and evening were counted, as well as off-peak hours, with a 15-minute resolution to record vehicle classes, directional flows, and queue lengths. These data were run through two probabilistic machine learning methods (Gaussian Process Regression and Relevance Vector Machines).

By comparing the performance of GPR and RVM under controlled conditions using the same datasets, the study aims to determine the best modeling paradigm for predicting congestion in an environment where manual data collection is the most commonly used methodology. This type of comparative work is quite necessary, as current studies have been inclined to investigate single modeling methods, and few have evaluated methods cross-methodologically (Mihaita et al., 2024). The lack of information on traffic in cities such as Calabar and the heterogeneity of traffic in these cities make this investigation especially timely. Hopefully, the results will guide evidence-based transportation planning not only in Calabar but also in similar situations in Nigeria, Africa, and other parts of the world, where data scarcity is the rule rather than the exception.

2. MATERIALS AND METHODS

2.1. Study Area

Calabar Metropolis is the capital of Cross River State in south-eastern Nigeria and is located along the Calabar River on the coast of the Niger Delta. It is geographically located at approximately 4.95 ° N and 8.33 ° E. The city, covering a 406 km2 area, is a significant administrative, commercial, and educational hub, and its economy is based on government, services, trade, tourism, and informal trade. The main traffic intersections that comprise the Calabar transport system are Uwanse-Mount Zion, Watt Market Roundabout, Tinapa Junction-Technical Roundabout, and Mobil MCC Junction. Traffic jams are a common occurrence, especially during rush hours in the morning and evening, caused by limited road capacity, poor road pavements, and widespread manual road control. These, along with the low and unsteady traffic statistics, make Calabar a suitable place to apply advanced traffic modeling techniques such as Gaussian Process Regression (GPR) and Relevance Vector Machines (RVM). These methods are suitable for modeling the nonlinear, dynamical, and uncertain characteristics of traffic flows, offering a more realistic approach to predicting traffic congestion and controlling traffic in cities.

2.1.1. Data Collection

The field team was trained to use systematic data sheets to conduct manual traffic counts at the selected intersections. The field assistants were also trained (2 days) on vehicle classification and timing. There were two independent observers who counted 6 hours a day in each site; the average of the two counts was below 7 percent, which shows that there was good inter-observer reliability. Counts were performed during the morning peak (7:00-9:00), evening peak (16:00-18:00), and off-peak (12:00-14:00) hours, on all days of the week, over 12 weeks. This method has been used because it records rich, detailed characteristics like the vehicle types, turning, occupancy, and loading, which would otherwise not be recorded by mechanical counters.

Since our Calabar traffic environment is diverse, it was necessary to use a manual counting approach to data collection, which would have yielded a more accurate and contextually sensitive data set. It involved a short scan of the city at the beginning to track traffic and its density. Then, major crossings in the city were monitored, including Uwanse-Mount Zion, Ettagbor-Main Gate, Mobil-MCC, Stadium-IBB Way, Watt Market Roundabout, Ekpo Abasi-Yellow Duke, and other major crossings that form the city's main road network. These places were chosen for their critical role as traffic hubs and recurrent congestion. Passenger Car Units per Hour (PCU/h) was used to standardize traffic volumes to represent traffic volumes that are comparable in terms of vehicles.

The empirical basis for implementing machine learning models, namely Gaussian Process Regression (GPR) and Relevance Vector Machine (RVM), was the collected data used to explore congestion patterns and forecast traffic behavior. The dynamic road signs and static road features were logged to enable robust modeling. These included: road name, width, length of the segments, the number of lanes, the presence of traffic lights, the number of potholes, the capacity of the highway, the day and the time of observation, time interval, the length of the queues, the arrival of the vehicles, the green time, Time demand, the volume of movement and change in the length of the queue. Together, these parameters gave an overall picture of the working conditions necessary to perform an advanced machine learning analysis.

2.2. Problem Formulation and Feature Representation

The traffic congestion prediction was developed as a supervised nonlinear regression problem. Let the complete observed dataset be defined as:

$$D = \{(x_i, y_i)\}_{i=1}^{N_{total}} \quad (1)$$

Where, D denotes the complete dataset containing all observed traffic instances, $x_i \in \mathbb{R}^p$ is the p-dimensional input feature vector representing the traffic and roadway conditions for the $i-th$ observation. y_i is the target congestion variable for the $i-th$ observation, and N_{total} is the total number of observations in the dataset.

The target variable y_i represents one of the following congestion indicators: queue length (m), total delay (s), or volume-to-capacity ratio. Thus, the underlying relationship between the input features and the target variable is modeled as:

$$y_i = f(x_i) + \varepsilon_i, \qquad \varepsilon_i \sim N(0, \sigma^2) \quad (2)$$

The unknown smooth, nonlinear function, $f(.)$, that compares the input features to the target variable, ε_i is a noise, which is independent and identically distributed (i.i.d.), meaning that error and other unobservable effects are independent and randomly distributed, and σ^2 s is the noise variance to be estimated.

2.2.1. Feature Selection and Relevance

The research employed two complementary procedures to determine predictors that have the greatest impact on traffic congestion. They applied Pearson and Automatic Relevance Determination (ARD) to a Gaussian Process Regression (GPR) model. The two analyses have been performed based on the training set (2352 observations) only in order to prevent the leakage of data into the test set. Under the Pearson correlation technique, correlations were first determined between each feature and three indicators of congestion (total delay, queue length, and volume-to-capacity ratio). This provided a linear measure of association, enabling the initial determination of the variables and yielding strong monotonicity. However, because traffic congestion often exhibits nonlinear interactions, correlation alone is insufficient for feature selection. For this reason, the Pearson correlation technique was complemented by Automatic Relevance Determination (ARD).

ARD was implemented using a GPR with a squared-exponential kernel, assigning a separate length scale to each input feature. The length scale of a feature in the ARD framework is negatively proportional to its predictive relevance. A small length scale means that the function changes swiftly with that feature and is thus very relevant, whereas a large length scale means that the feature does not significantly affect the output. Learning of the length scales maximized the log-marginal likelihood with the L-BFGS algorithm (with five random restarts). The relevance of each feature was a measure of the inverse of the learned length scale (1 / length scale).

2.3. Data Preprocessing and Partitioning

The raw data D was processed through a typical preprocessing pipeline. This included cleaning (utilization of missing values), dealing with outliers, and standardizing all the numerical features to achieve numerical stability in the model training. The resulting processed data was randomly divided into a training set D_train(70% of data) to develop the model and a testing set D_test (30% of data) to evaluate the final results, where:

$$D = D_{train} \cup D_{test}$$

This partitioning enabled an unbiased assessment of model generalization performance.

2.4. Model Development

2.4.1. Gaussian Process Regression (GPR)

GPR is a non-parametric, Bayesian approach to regression. The core assumption is that the latent function. $f(x)$ follows a Gaussian Process (GP) prior, defined as:

$$f(x) \sim GP\left(0, k(x, x')\right) \tag{3}$$

Where $k(.,.)$ is a positive-definite covariance (kernel) function. In this study, the Squared Exponential (SE) kernel with Automatic Relevance Determination (ARD) was employed:

$$k(x_i, x_j) = \sigma_f^2 \exp\left(-\frac{1}{2}\sum_{m=1}^{p}\frac{(x_{i,m} - x_{j,m})^2}{l_m^2}\right), \tag{4}$$

Where, σ_f^2 is the signal variance controlling the overall scale of the function, and l_m is the characteristic length-scale for the $m-th$ input dimension ($m = 1, 2, \ldots, p$). Larger l_m values imply lower relevance of the corresponding input feature.

The set of GPR hyperparameters is defined as:
$\theta_{GPR} = \{\sigma_f^2, l_1, \ldots, l_p, \sigma^2\}$, where σ^2 is the noise variance from Eq. (2)

Given the training dataset $D_{train} = \{(x_i, y_i)\}_{i=1}^{N_{train}}$, the joint distribution of the observed target values, $y = [y_1, y_2, \ldots, y_{N_{train}}]^T$ and the predicted function value $f_* = f(x_*)$ at a new point test x_* is Gaussian. And the posterior predictive distribution is:

$$p(f_*/x_*, D_{train}) = N(\mu_*, \sigma_*^2), \tag{5}$$

Where μ_* denotes the point prediction (expected congestion level) for the input x_*, and σ_*^2 is the predictive variance, quantifying the uncertainty in the prediction. Kernel hyperparameters were estimated by maximizing the log marginal likelihood of the training data.

2.4.2. Relevance Vector Machine (RVM)

The Relevance Vector Machine was applied as a sparse Bayesian regression model for congestion prediction. The RVM expresses the response variable as:

$$y(x) = \sum_{i=1}^{m} w_i\, k(x, x_i) + \varepsilon \tag{6}$$

Where,w_i, are model weights and $k(.,.)$ is a kernel function, m is the number of relevance vectors, and ε, is the Gaussian noise.

A Gaussian kernel was used and is defined as:

$$k(x, x_i) = \exp\left(-\frac{\| x - x_i \|^2}{2\varsigma^2}\right) \tag{7}$$

To enforce sparsity, an Automatic Relevance Determination (ARD) prior was placed over the weights as:

$$p(w_i/\alpha_i) = N\left(0, \alpha_i^{-1}\right) \tag{8}$$

Where α_i It is a precision hyperparameter. During training, many α_i values tend to infinity, forcing the corresponding weights to zero and yielding a sparse model with a limited number of relevance vectors.

2.5. Model Calibration and Validation

Calibration of the models was only carried out on the D_train training data. In the case of GPR, the calibration was done by estimating the hyperparameter set θ_GPR that maximized the log marginal likelihood. In the case of RVM, the process included an evidence-maximization approach, performed iteratively to locate {w, α, ς^2}. Overfitting was guarded by using K-fold cross-validation (K=5) in this stage. In the case of GPR, the marginal likelihood was maximized in terms of the log using the L-BFGS algorithm with 5 random restarts. Length-scale bounds were set to [1e-2, 1e2], and noise-level bounds to [1e-3, 1e1]. For RVM, the Gaussian kernel width was selected via 5-fold cross-validation over a logarithmic grid from 0.1 to 10 using the evidence-maximization procedure (Tipping, 2001).

Final model performance was evaluated on the independent test set D_test. Let {(x_j y_j)}_(j=1)^(N_test) denote the test samples. For each test input x_j, each model produces a point prediction y ˆ_j. Performance was quantified using the following metrics: Root Mean Square Error (RMSE), Normalized Root Mean Square Error (NRMSE), and the coefficient of determination. (R^2).

$$RMSE = \sqrt{\frac{1}{N_{test}} \sum_{j=1}^{N_{test}} \left(y_j - \hat{y}_j\right)^2}, \quad (9)$$

$$R^2 = 1 - \frac{\sum_{j=1}^{N_{test}} \left(y_j - \hat{y}_j\right)^2}{\sum_{j=1}^{N_{test}} \left(y_j - \bar{y}_{test}\right)^2}, \quad (10)$$

To enable scale-independent comparison across different congestion indicators (queue length, delay, and VCR), the Normalized Root Mean Square Error (NRMSE) was computed as:

$$NRMSE = \frac{RMSE}{\sigma_y} \quad (11)$$

Where σ_y is the standard deviation of the observed target variable in the test set, defined as:

$$\sigma_y = \sqrt{\frac{1}{N_{test}} \sum_{j=1}^{N_{test}} \left(y_j - \bar{y}_{test}\right)^2}, \quad (12)$$

2.6. Scenario-Based Application

The calibrated models were implemented in a scenario-based analysis to show practical applicability. Key control variables: signal green time (G_i), lanes (N), and traffic volume (V_i), were varied systematically within realistic ranges. In every scenario, s, which is characterized by a given set of inputs $x^{(s)}$ The models produced a congestion prediction. y ^((s)). Predictions across scenarios provided quantitative estimates of congestion sensitivity to these variables, providing direct, model-driven information for traffic planning and policy formulation in Calabar Metropolis.

3. RESULT AND DISCUSSION

Table 1 shows the descriptive statistics of traffic congestion measures. The indicators indicate that traffic in Calabar Metropolis is at average levels, but it is highly varied both in time and space. The volume to capacity ratio (VCR) is between 0.01 and 4.18, with an average of 0.437, which means that the roads are not fully utilized during off-peak hours. Values on VCR that exceed one, on the other hand, represent intermittent but intense congestion, such as at morning and evening rush periods and at major road intersections like Watt Market Roundabout, Uwanse-Mount Zion, Tinapa-Technical Roundabout, and Mobil Junction, where local demand can easily be higher than the capacity of the road. The average and large queue of 223.91 m and 545 m, respectively, are indicative that normal and extreme congestion takes place, and spillback and network-bound disruption.

Table 1. Descriptive Statistics of Traffic Congestion Metrics

	N	Min	Max	Mean	Std. Deviation
Volume to Capacity Ratio (VCR)	3360	0.01	4.18	0.44	0.44
Queue Length (m)	3360	0.00	545.00	223.91	133.28
Total Delay (s)	3360	0.00	163500.00	50632.17	40648.91
Total	3360				

A large standard deviation (133.28 m) is typical of Calabar traffic, where small lanes, roadside incidents, and manual traffic control cause delays during the busy times. The same trend is further attested to by the total delay values, with the highest delay of 163,500 seconds and a mean delay of 50,632 seconds, indicating that long delays are inherent in cases of congestion. These radical delays are in line with the situation in Calabar, where accidents, poor signal coordination, and a lack of traffic management infrastructure severely affect travel time reliability. In general, the findings reveal that congestion in Calabar is not sustained but highly episodic, driven by

both local bottlenecks and temporal increases in demand rather than over-capacity. The highly fluctuating nature of this suggests the nonlinearity and uncertainty of traffic dynamics within the city. It supports the appropriateness of probabilistic modeling designs in practical congestion prediction and control.

Table 2. One-Way ANOVA Results for VCR, Queue Length, and Total Delay Across 20 Road Locations

ANOVA						
		Sum of Squares	**Df**	**Mean Square**	**F**	**Sig.**
Volume to Capacity Ratio	Between Groups	93.511	19	4.922	29.614	.000
	Within Groups	555.078	3340	.166		
	Total	648.588	3359			
Queue Length (m)	Between Groups	1541973.047	19	81156.476	4.663	.000
	Within Groups	58127019.701	3340	17403.299		
	Total	59668992.748	3359			
Total Delay (s)	Between Groups	95542708952.946	19	5028563629.102	3.079	.000
	Within Groups	5454645653330.357	3340	1633127441.117		
	Total	5550188362283.304	3359			

Table 2 presents the results of a one-way ANOVA for VCR, Queue Length, and Total Delay across 20 road locations. The analysis evaluated whether differences among locations were statistically significant for the Volume-to-Capacity Ratio (VCR), Queue Length (m), and Total Delay (s). For VCR, the between-groups sum of squares was 93.511 (df = 19, mean square = 4.922), and the within-groups sum of squares was 555.078 (df = 3340, mean square = 0.166). The F-value was 29.614, and the p-value was highly significant (p = 0.000). It means that VCR averages will be highly dispersed over the locations, so not all of them will be saturated as compared to the capacity: some roads will be almost at capacity, and others will be in free flow. For Queue Length, the between-groups sum of squares was 1,541,973.047 (df = 19, mean square = 81,156.476), and the within-groups sum of squares was 019.701 (df = 3340, mean square = 17,403.299). The F-value was 4.663, and the p-value was significant (p = 0.0). Similarly, for Total Delay, the between-groups sum of squares was 95,542708,952.946 (df = 19, mean square = 5, 028,563,629.102), and the within-groups sum of squares was 5,45464565330.357 (df =3340, mean square = 1, 633, 127,441.10). The overall ANOVA findings suggest that there are significant differences in the level of traffic congestion, length of queue, and total delay among the 20 road locations.

The findings of Tukey HSD (Appendix tables A1- A4) show that the congestion in Calabar is not a continuous phenomenon but due to the local bottlenecks along the Harbor, Tinapa, Main Avenue, Goldie, and Ettagbor crossings. Ekpo Abasi, Uwanse, and Goldie (with high VCR) performed relatively well in terms of delay and queue, suggesting effective signal control. These results therefore support targeted interventions in areas with the worst performance and the use of probabilistic models to manage the high variability and nonlinearity of traffic within the city.

Table 3. Model Accuracy and Predictive Performance for Key Traffic Congestion Variables (GPR vs. RVM)

Congestion Variables	**Model**	**RMSE**	**SD**	**Normalized RMSE**	**R^2**
Queue Length (m)	GPR	24.99	133.28	0.19	0.96
	RVM	50.23	133.28	0.38	0.86
VCR (CV)	GPR	0.19	0.44	0.43	0.80
	RVM	0.22	0.44	0.50	0.74
Total Delay (s)	GPR	14457.22	40648.91	0.36	0.87
	RVM	18144.97	40648.91	0.45	0.80

Table 3 presents the model accuracy and predictive performance for key traffic congestion variables for GPR and RVM. The GPR model is more precise than RVM and better at predicting traffic congestion trends. In the Queue Length case, the GPR estimates were much closer to the actual queue lengths, with a natural deviation of approximately 19 percent, compared to RVM, which had a deviation of approximately 38 percent. This means that GPR can correctly identify the presence of long vehicle queues at crossroads,

whereas RVM cannot. Across the data, GPR explains 96 percent of the variation in queue length, compared to 86 percent for RVM, indicating that GPR is a more reliable predictor of the key determinants of vehicle build-up.

Among the measures, the Volume-to-Capacity Ratio (VCR) had the lowest relative error (43 percent) compared with RVM (50 percent), and it explained 80 percent of its variation, which is greater than the 74 percent for RVM. Although it is not as precise as the predictive procedures of VCR because the ratio can change radically when traffic is slightly altered, GPR would still offer a better sense of the overall picture of road congestion than capacity. The actual travel delays in the case of Total Delay were predicted with a relative error of 36%, which was lower than the 45% for RVM and GPR, and GPR explained the delay variation better than RVM, which had an 80% relative error. This means that GPR is better positioned to account for unexpected factors that cause delays, such as traffic lights, turnings, and random delays. As shown in the results analysis in Table 3, GPR is more accurate and valid across all congestion variables. It is the most precise at predicting queue lengths, followed by total delays, and VCR is the least predictable because it is affected by even slight changes in traffic. The strong explanatory power and low prediction error of GPR make it an effective tool. They can be used to study congestion and the remedies to be adopted in traffic, rather than RVM, which is slightly less accurate but can still provide a few meaningful estimates, at least for understanding the manner of using road capacity.

3.1. Performance Evaluation of GPR vs RVM Models After Predicting Traffic Congestion Data

Table 4 shows the prediction accuracy of the Gaussian Process Regression (GPR) and Relevance Vector Machine (RVM) models. Root Mean Square Error (RMSE), Standard Deviation (SD), Normalized RMSE, Mean Absolute Error (MAE), coefficient of determination (R2), and 95% coverage probability were used to measure performance. The set of these metrics was used to evaluate predictive accuracy, explanatory power, and uncertainty calibration of the three variables of congestion. As far as queue length is concerned, GPR has a much lower RMSE (22.09) than RVM (50.47), and the predicted queue lengths were far less. The normalized RMSE did the same, where GPR (0.17) performed better than RVM (0.38), meaning that the GPR predictions were close to the observed queue lengths. These results were in compliance with this pattern: GPR was characterized by a very low MAE of 4.96m, but RVM possessed a much bigger MAE of 37.89m. GPR was more explanatory (R2 of 0.97) with 97 percent of the variation in queue length being explained than RVM (R2 of 0.85). This demonstrates that GPR is able to capture a large portion of the factors influencing vehicle queues at crossroads. The probability of coverage of GPR (0.93) was very similar to the theoretical value (0.95), and thus, this implied that the GPR could be trusted when estimating uncertainty.

Table 4. Performance metrics for GPR and RVM

Congestion Variables	Model	RMSE	SD	Normalized RMSE	R^2	MAE	95 % Coverage
Queue Length (m)	GPR	22.09	133.28	0.17	0.97	4.96	0.93
	RVM	50.47	133.28	0.38	0.85	37.89	-
VCR (CV)	GPR	0.22	0.44	0.50	0.74	0.083	0.90
	RVM	0.25	0.44	0.57	0.66	0.154	-
Total Delay (s)	GPR	13691.14	40648.91	0.34	0.88	1562.69	0.94
	RVM	17731.94	40648.91	0.44	0.80	13889.52	-

On VCR prediction, GPR was also very good with a smaller RMSE (0.22) than RVM (0.25). GPR also had a smaller normalized RMSE (0.50) compared to RVM (0.57). This is supported by the values of the MAE: GPR generated an MAE of 0.083, and RVM generated an MAE of 0.154, which means that the average of the predicted VCR values is more similar to the actual ones in GPR. GPR and RVM explained 74 and 66 percent of the variability in VCR, respectively. The reduced predictive performance of VCR as compared to queue length is probably explained by the fact that VCR is sensitive to small changes in traffic demand and capacity of the roads. Nevertheless, GPR showed better modeling ability of this variable. The 95 coverage probabilities of 0.90 show that there is fairly good predictive uncertainty calibration. Finally, under the congestion variable (total delay), GPR performed better than RVM in all the measures of the assessments. GPR recorded an RMSE of 13,691.14, compared with 17,731.94 for RVM. The GPR-normalized RMSE (0.34) also showed a smaller value as compared to the RVM-normalized RMSE (0.44). The difference between the values of MAE was particularly high: GPR had an MAE of 1,562.69 seconds, and RVM had a much higher value of 13,889.52 seconds, meaning that GPR is more accurate in delay predictions. GPR explained 88% of the total delay variation compared to 80% of variation in RVM in terms of

explanatory power. It means that GPR is more efficient to approximate these non-linear and complicated factors of traffic delays, such as signal timing, turning movements and stochastic changes in traffic flow. The applicability of GPR, with a 95% coverage probability of 0.94, also confirms that its predictive uncertainty ranges are well-calibrated and relevant to theory.

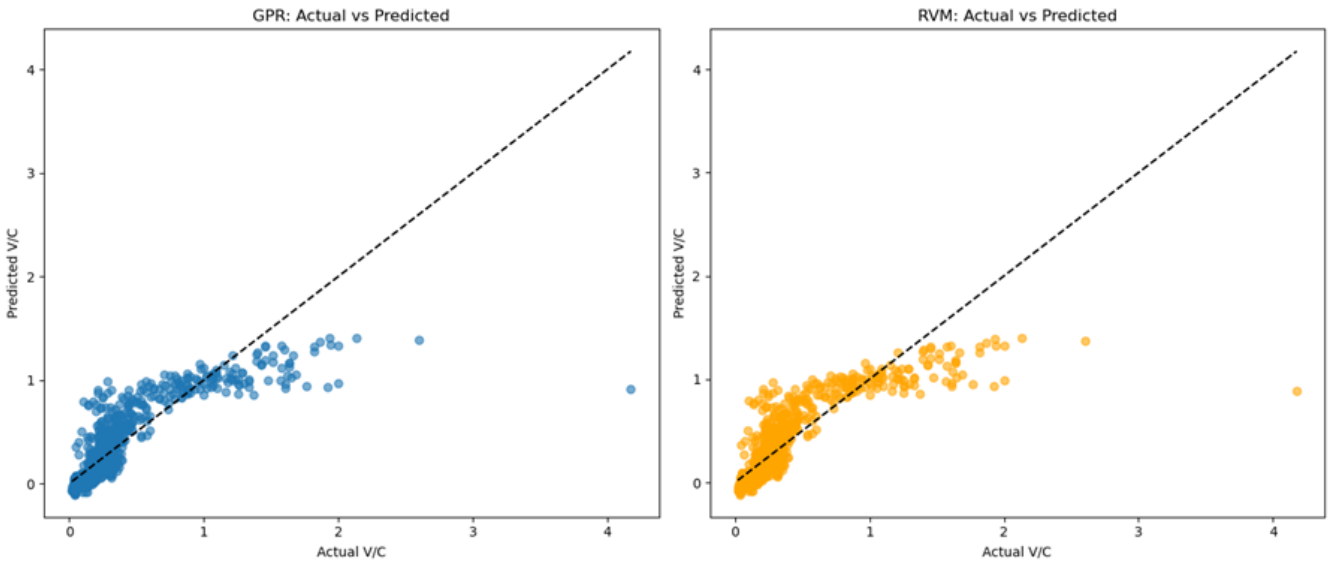

Figure 1. Assessment of GPR vs RVM fitness via scatter plots for VCR

Figure 1 shows two scatter plots of the actual and predicted volume-to-capacity (V/C) of two models: Gaussian Process Regression (GPR) and Relevance Vector Machine (RVM). The black dotted line in both plots is used to show the ideal situation where the predicted V/C is equal to the actual V/C, and the points along the line are accurately predicted. The points are generally clumped in the lower-left corner, suggesting that most actual and predicted V/C values are quite low. Both models underpredict when the actual V/C values are above approximately 0.5, as data points below the ideal line indicate the need to estimate high-traffic conditions even when they are low. The spread in the predicted values increases with the actual V/C, indicating that the models are less consistent at higher traffic levels. Specifically, the GPR model shows a more concentrated cluster of points around the ideal line at lower V/C values, but the predictions are more dispersed and underpredicted at larger V/C values, which are actually higher; for example, the point at (4, 1) is predicted to be at a significantly lower value. This pattern is also evident in the RVM model, which is closer to lower V/C values but overpredicts higher values; the difference is slightly smaller than with GPR, and large outlier high V/C values do not occur, indicating a slightly higher degree of robustness. In general, both models are quite useful in the low V/C conditions prediction, but are severely challenged with the high V/C conditions, presumably because of limited training data available to forecast high-congestion conditions, the weaknesses of the model under consideration, and the potential lack of predictive information in the chosen input features to predict high-congestion conditions.

Figure 2 presents the bias and central tendency analyses, showing that both the Gaussian Process Regression (GPR) and the Relevance Vector Machine (RVM) models tend to underestimate the V/C values, particularly at higher levels, as evidenced by the positive skewness of the error distributions (Actual minus Predicted). More precisely, the GPR error distribution is slightly skewed to the right, and the mean and median errors are also slightly positive, confirming a slight underprediction bias in the GPR error distribution compared with the RVM error distribution. Both distributions have their maximum and median much further to the right. The prediction distribution analysis shows that the GPR error distribution is relatively larger, with a range from -0.5 to more, including a few extreme errors of less than -3.0 and a few extreme errors to the right, which is consistent with the scatter plot revealing that the number of more widely spread points is higher with larger V/C. RVM, on the other hand, has a somewhat smaller error range, mostly between -0.5 and 0.5, with shorter tails; the longer tail is to the left, indicating fewer extreme errors below -3.0. It is also worth noting that the outlier in the GPR at the extreme of the scatter plot around (4,1) is associated with strong underprediction, i.e., the right tail of the error distribution, which is not true of RVM. Collectively, these findings corroborate the initial argument that both models are more systematic predictors of V/C at larger values. Nevertheless, the GPR is less predictable with its error distribution being wider, with both

extreme over- and underprediction. Conversely, the RVM, although also being biased towards prediction, has a narrower distribution of error and a narrower predictive distribution.

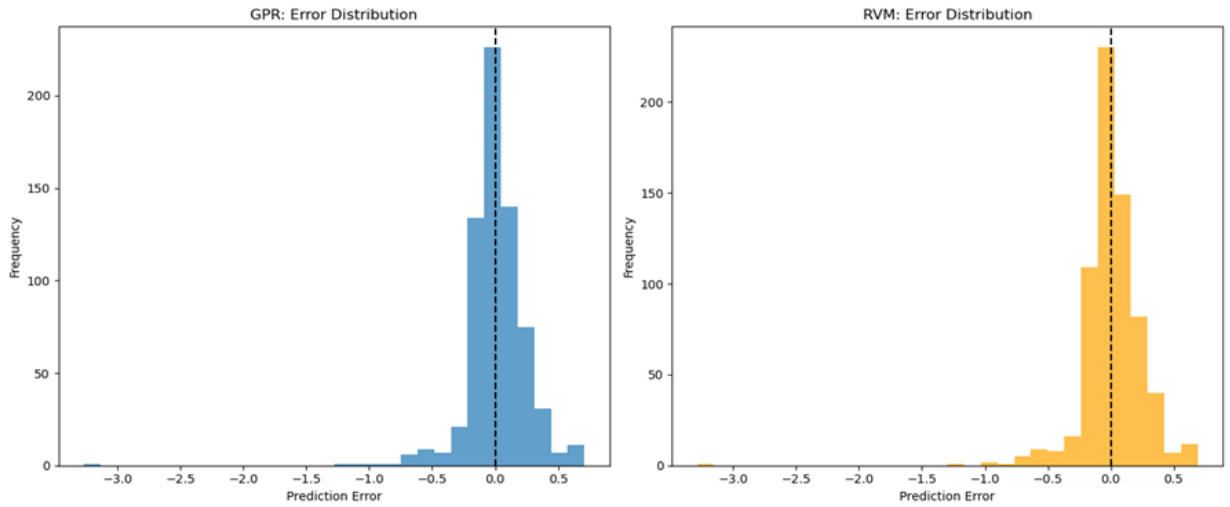

Figure 2. Assessment of GPR vs RVM fitness via error distribution for VCR

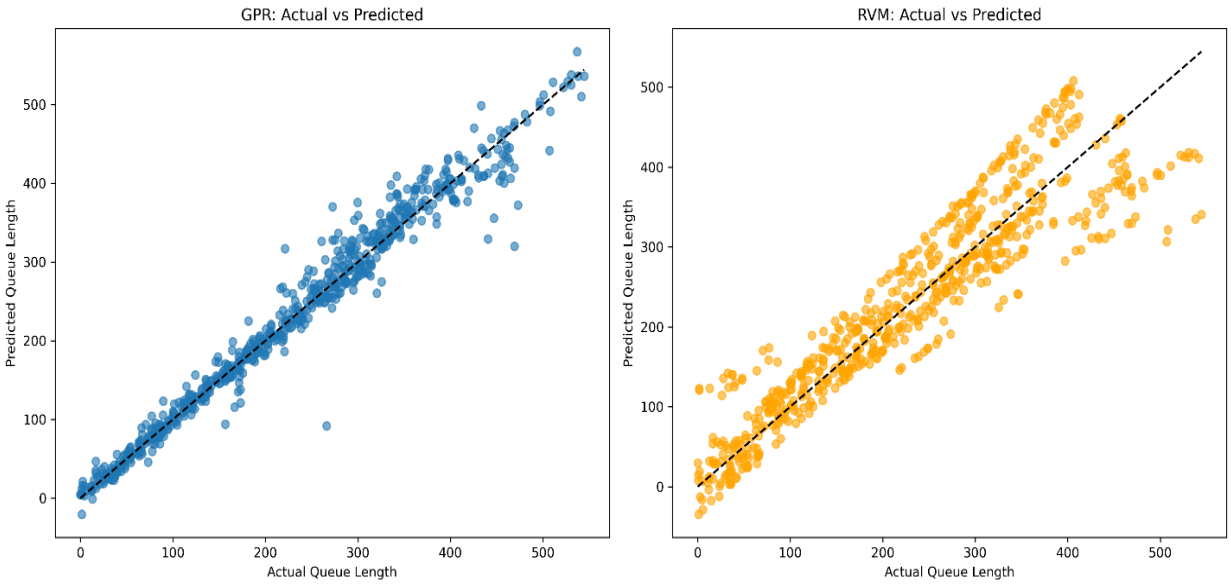

Figure 3. Assessment of GPR vs RVM fitness via scatter plots for queue length (m)

Figure 3 shows the GPR and RVM fitness for queue length. The scatter plots show the correlation between the actual and predicted queue lengths. The Gaussian Process Regression (GPR) and Relevance Vector Machine (RVM) models exhibit a strong positive correlation, with the predicted queue length closely matching the actual queue length. The points on the black curve indicate the ideal range, where the models achieve perfect prediction of queue length. The GPR model fits the entire range of queue lengths very well, with most points closely following the ideal line, indicating high consistency and minimal underprediction, as the predicted values are close to the actual values. However, there are a few exceptions, such as the point around (200,100) and a few points around (100,200) that occasionally show under- and overprediction. The RVM model also has a fairly good fit, with a larger variability value and points more widely spread around the ideal line. There is a slight underprediction at actual queue lengths of 200-250, and the bands of predicted values are wider. The RVM also shows a smattering of prediction bands at higher queue lengths, which could be due to the model's discrete output or to the RVM's implementation.

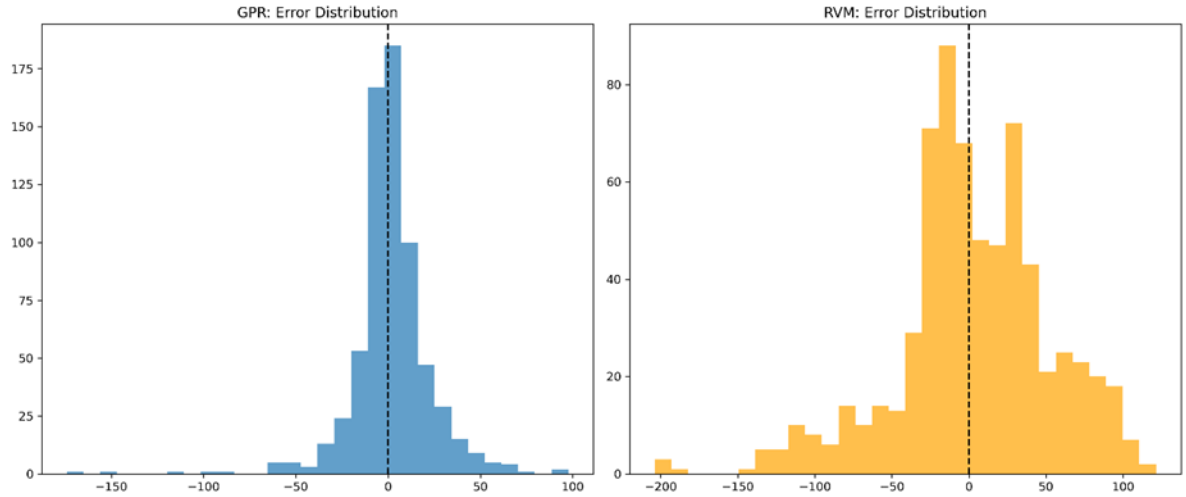

Figure 4. Assessment of GPR vs RVM fitness via error distribution plots for queue length (m)

Figure 4 shows the error distribution plot for queue length. Analysis of central tendency and bias indicates that the Gaussian Process Regression (GPR) model fits the data well, with small bias. The error distribution is tightly peaked at zero, and the histogram and the mean/median error are nearly equal at zero, revealing that the GPR does not consistently overfit or underfit queue lengths. Nevertheless, the Relevance Vector Machine (RVM) model shows a slightly different error distribution, with the highest values and the mean/median error located on the right-hand side (around +10 to +20), indicating a slight overall tendency to underpredict, as in the scatter plots. Moreover, in the distribution of predictions, the high consistency of GPR is characterized by errors concentrated within +/-25 units, with nearly all within +/-50 units, and only some irregularities at approximately -150 and +100 units, pointing to high precision. RVM, on the other hand, has a far more extended and flatter error distribution, with numerous errors outside +/-50 units and a wider dispersion of approximately -200 to +200 units, as anticipated, and is considerably less consistent and accurate than GPR. Taken together, these findings provide a full picture that complements the scatter plots, indicating that GPR is highly precise, reliable, and consistent in predicting queues, and that RVM. However, capable of making reliable predictions, slightly underpredicts, and has a much more scattered error distribution, as evidenced by discernible bands in its scatter plot.

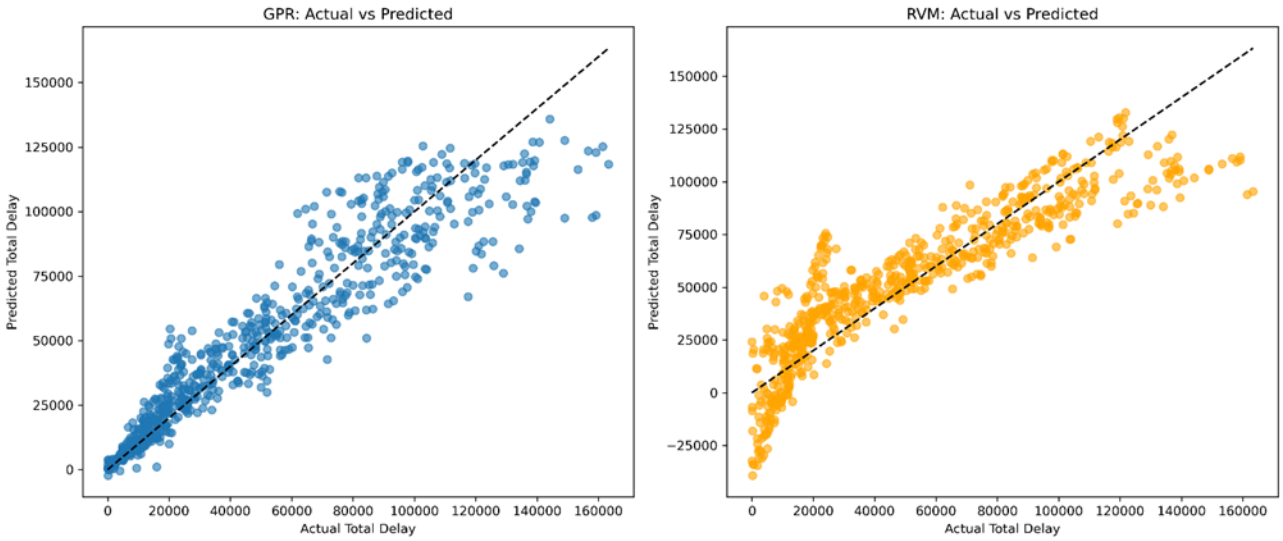

Figure 5. Assessment of GPR vs RVM fitness via scatter plots for total delay (s)

Figure 5 shows a scatter plot of the actual and predicted total delay, indicating that the two are highly correlated for both the Gaussian Process Regression (GPR) and Relevance Vector Machine (RVM) models. The dashed black line represents a perfect fit, where the predicted total delay equals the actual total delay. The total delay values range up to approximately 160,000, indicating a cumulative delay over an extended period or across multiple entities. A good fit is observed in the GPR model, where the points are near the ideal line, yielding good predictions at low total delays, particularly below 20,000, when the model is unbiased and very consistent. As the actual total delays increase, the distribution of the predicted values also increases, reflecting a loss of precision and a small tendency toward underprediction at higher levels of total delays, say above 80,000. On the other hand, the RVM model is heavily biased toward underprediction across nearly the whole range of possible actual delays, especially for larger delays (i.e., larger than 20,000). Its predictions are much lower than the actual delays, are more diffuse, and are less predictive than those of GPR. It also has cases of physically impossible negative predicted delays (e.g., -25,000, --25,000), which is a severe limitation of the model for this task. In addition, one can see slight banding or layered patterns in the RVM predictions, suggesting properties of the algorithm or the input feature set. In conclusion, GPR is more accurate and predictable at lower and middle delays, but very inaccurate at very high delays. Predictability is also lower, and negative predictions, low predictability, and the presence of negative-prediction issues are critical for revealing that GPR does not apply well to the specified prediction task.

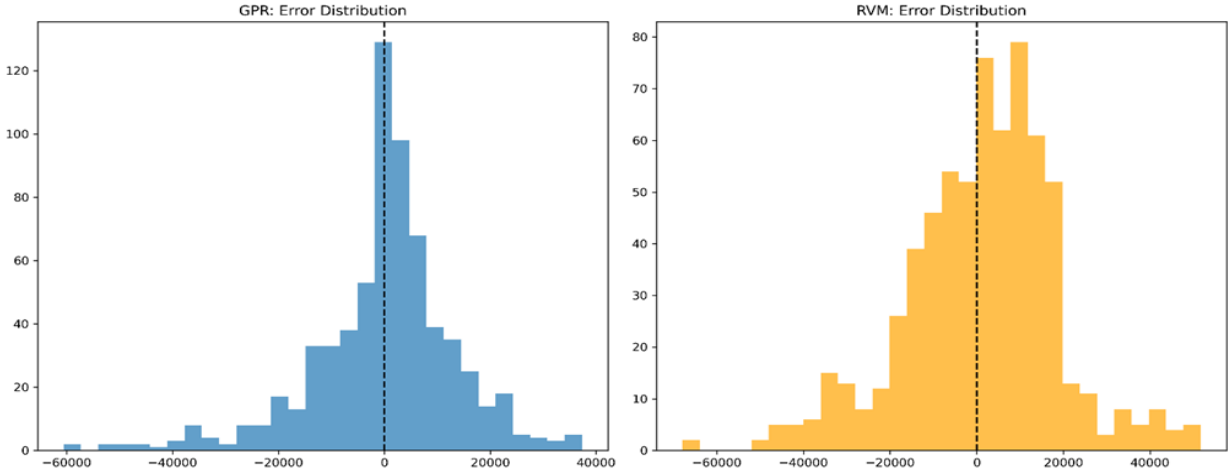

Figure 6. Assessment of GPR vs RVM fitness via error distribution plots for total delay (s)

Figure 6 shows the error distributions for total delay prediction. The analysis indicates that the error distributions of the Gaussian Process Regression (GPR) model are largely unbiased, with errors centered around zero. The histogram and the near-zero mean and median errors support this. Errors at higher delays are slightly underestimated and, as such, have no significant effect on the overall distribution. On the other hand, the Relevance Vector Machine (RVM) model is highly underestimated, as its error distribution is heavily skewed to the right, ranging from +5,000 to +10,000, with a very large positive mean/median error because underestimation always occurs in the total delays in RVM. The prediction consistency analysis shows that the GPR prediction errors are more consistent, varying around -60,000 to +40,000 (with more actual delays), than the RVM errors, which are more scattered around -60,000 to +50,000, with many errors not even close to zero. Furthermore, RVM exhibits physically implausible negative predicted delays, which, in combination with the actual delay, lead to large positive errors and contribute to the large range and high positive bias of RVM's error distribution. Collectively, these findings indicate that, despite being a fairly successful predictor of total delay, providing predictions of, on average, reasonable precision, with large variance at large delay times, RVM is a poor predictor, with large bias in under-prediction, poor predictive accuracy, and the phenomenon of negative predicted delay, rendering it an ineffective predictor of total delay.

3.2. Feature Relevance

Based on the correlation results (Appendix tables A5- A9), the number of vehicle arrivals was the strongest positive correlate for both total delay (r = 0.669) and queue length (r = 0.858). The demand of traffic, however, was positively correlated with the volume-capacity ratio (0.568). The Gaussian process regression that uses ARD length scales offered a marginally different view of feature significance. In the case of total delay (Appendix table A6), the most significant predictors were the number of potholes (length scale 0.55), then the

number of vehicle arrivals (0.95), and road class (1.96). In the queue length (Appendix table A8), the order was reversed: the most significant impact was on the number of vehicle arrivals (0.67), the number of potholes (0.85), and the number of lanes (1.00). For VCR (Appendix table A10), the key features were the time difference between measurements (0.055), the number of potholes (0.78), and traffic demand (0.81). These findings underscore the importance of using nonlinear methods to capture feature relevance in traffic congestion modeling, as linear correlations alone may overlook critical drivers of congestion.

4. CONCLUSION

The findings suggest that traffic congestion in Calabar Metropolis is typically moderate but highly skewed across time and space. Although the average volume-to-capacity ratio (VCR) of about 0.44 shows that most of the roads are underutilized during the low-traffic time of the day, high VCRs of up to 4.18 represent extreme but rare congestion at the busiest crossings, including Watt Market Roundabout, Uwanse-Mount Zion, Tinapa-Technical Roundabout This variability is also reflected in the queue lengths and the mean queue length is about 224 m and maximum length is 545 m, which means high network-wide disruption and spillback during the And likewise, total delay is in the range of about 0 to about 163,500 s with a median of 50,632 s indicating the time lost in congestion. Comparison results show that GPR was always superior to RVM on the measures and dataset, but the difference was insignificant when predicting VCR. GPR obtained lower normalized RMSE of about 0.17, 0.34, and 0.50 in queue length, total delay, and VCR, respectively, with respect to the variation in queue length up to 97 percent and total delay up to 88 percent. RVM, on the other hand, exhibited greater prediction errors and an underestimation bias with heavy congestion, especially when the delays were significant. In general, the results indicate that Calabar congestion is discontinuous and determined by the existence of local points of congestion and variability of demand, and the probabilistic models, e.g., GPR, are a more efficient and context-specific model to forecast congestion and traffic control in data-limited medium-sized cities.

APPENDICES

Table A1: Per-location distribution of observations into training (70%) and test (30%) sets

Location	Training	Test	Total
Abang Aseng Akim Qua Town	111	57	168
Atimbo Round About	120	48	168
Effio Ette Round About	126	42	168
Ekpo Abasi Round About	122	46	168
Eleven-Eleven – Total	117	51	168
Ettagbor – Akim	118	50	168
Ettagbor – Main Gate	98	70	168
Ettagbor Roundabout – IBB	123	45	168
Goldie by Mount Zion	118	50	168
Harbour Junction	122	46	168
Main Avenue-Ekpo Abasi	124	44	168
Mary-Slessor – Total Filling Station	115	53	168
Mobil by MCC	114	54	168
Ndidem Usang Iso, Atekong	124	44	168
Stadium – IBB Way	116	52	168
Tinapa Junction	118	50	168
Uwanse – Mount Zion	118	50	168
Watt Market Round About	125	43	168
White House – Chamberly	114	54	168
Yellow Duke	109	59	168

Table A2: Tukey HSD homogeneous subsets for total delay (α=0.05)

Location	N	Subset 1	Subset 2
Ekpo Abasi Round About	168	44961.07	
Uwanse – Mount Zion	168	45125.36	
Mary-Slessor – Total Filling Station	168	45568.21	
Goldie by Mount Zion	168	45612.86	
Ettagbor – Akim	168	46211.61	
Ettagbor – Main Gate	168	46324.82	
Main Avenue-Ekpo Abasi	168	47399.11	47399.11
Effio Ette Round About	168	47862.14	47862.14
Ndidem Usang Iso, Atekong	168	48790.18	48790.18
Watt Market Round About	168	49039.46	49039.46
Atimbo Round About	168	49075.36	49075.36
Abang Aseng Akim Qua Town	168	49899.64	49899.64
Ettagbor Roundabout – IBB	168	50116.61	50116.61
White House – Chamberly	168	50291.43	50291.43
Eleven-Eleven – Total	168	52326.07	52326.07
Yellow Duke	168	53644.46	53644.46
Stadium – IBB Way	168	57243.21	57243.21
Mobil by MCC	168	58657.32	58657.32
Harbour Junction	168		62148.75
Tinapa Junction	168		62345.71
Sig.		0.179	0.082

Table A3: Tukey HSD homogeneous subsets for queue length (α=0.05)

Subset	Locations (mean queue length in meters)
1 (shortest)	Ekpo Abasi (196.13), Uwanse (202.10), Goldie (202.17)
2	Mary-Slessor (204.97), Ettagbor-Akim (211.82), Watt Market (212.96), Ettagbor-Main Gate (214.83), Effio Ette (214.94), Main Avenue (215.18), Ndidem Usang (218.37), Abang Aseng (219.12), Atimbo (219.25)
3	Eleven-Eleven (220.98), White House (221.42), Ettagbor Roundabout-IBB (221.44)
4	Yellow Duke (232.85), Stadium (249.29)
5 (longest)	Mobil by MCC (253.57), Harbor (271.68), Tinapa (275.02)

Table A4: Tukey HSD homogeneous subsets for VCR (α=0.05)

Subset	Locations (mean VCR)
1 (lowest)	Uwanse (0.121)
2	Eleven-Eleven (0.159)
3	White House (0.286), Yellow Duke (0.318), Mobil by MCC (0.325), Atimbo (0.326), Effio Ette (0.366), Mary-Slessor (0.366), Ettagbor Roundabout-IBB (0.370), Ekpo Abasi (0.415), Harbor (0.428), Watt Market (0.432), Ndidem Usang (0.433)
4	Stadium (0.455)
5 (highest)	Abang Aseng (0.617), Ettagbor-Akim (0.630), Tinapa (0.643), Ettagbor-Main Gate (0.644), Goldie (0.651), Main Avenue (0.755)

Table A5: Pearson correlation with total delay

Feature	Correlation
Number of vehicle arrivals	0.669277
Time Difference (minutes)	0.593615
Class of Road_Trunk "A"	0.091549
Road width (m)	0.087816
Number of lanes	0.070242
Day_Friday	0.047599
Length of Segment (m)	0.031652
Day_Monday	0.030193
Time of Day_Afternoon	0.022577
Day_Thursday	0.015403
Day_Saturday	0.010436
Time of Day_Morning	0.008373
Presence of traffic warden_Yes	0.007453
presence of traffic light_No	0.000408
presence of traffic light_Yes	-0.000408
Number of potholes	-0.005318
Presence of traffic warden_No	-0.007453
Green time(sec)	-0.015316
Day_Wednesday	-0.027790
Time of Day_Evening	-0.030397
Day_Tuesday	-0.034557
Day_Sunday	-0.044287
Class of Road_Trunk "C"	-0.091549
Traffic demand	-0.383337

Table A6: ARD feature importance for total delay (smaller length scale = more relevant)

Feature	Length Scale	Relevance (1/Length Scale)
Number of potholes	0.549900	1.818512
Number of vehicle arrivals	0.953193	1.049105
Class of Road_Trunk "C"	1.956621	0.511085
Class of Road_Trunk "A"	1.956621	0.511085
presence of traffic light_Yes	2.557938	0.390940
presence of traffic light_No	2.557938	0.390940
Green time(sec)	2.840905	0.352001
Time Difference (minutes)	3.292547	0.303716
Presence of traffic warden_Yes	5.005862	0.199766
Presence of traffic warden_No	5.005862	0.199766
Length of Segment (m)	5.399128	0.185215
Road width (m)	7.169431	0.139481
Number of lanes	7.603643	0.131516
Day_Saturday	60.367228	0.016565
Time of Day_Morning	85.509696	0.011695
Time of Day_Evening	100.000000	0.010000
Time of Day_Afternoon	100.000000	0.010000
Day_Wednesday	100.000000	0.010000

Day_Tuesday	100.000000	0.010000
Traffic demand	100.000000	0.010000
Day_Thursday	100.000000	0.010000
Day_Sunday	100.000000	0.010000
Day_Monday	100.000000	0.010000
Day_Friday	100.000000	0.010000

Table A7: Pearson correlation with Queue Length

Feature	Correlation
Number of vehicle arrivals	0.857814
Class of Road_Trunk "A"	0.183928
Number of lanes	0.120585
Length of Segment (m)	0.099755
Road width (m)	0.091337
Presence of traffic warden_Yes	0.047186
Day_Thursday	0.040666
Day_Monday	0.038633
presence of traffic light_No	0.018200
Time Difference (minutes)	0.007133
Day_Wednesday	0.006997
Traffic demand	-0.003270
Number of potholes	-0.004339
Time of Day_Evening	-0.005221
Day_Sunday	-0.010947
Day_Tuesday	-0.011453
presence of traffic light_Yes	-0.018200
Time of Day_Afternoon	-0.018381
Day_Friday	-0.022951
Green time(sec)	-0.026437
Day_Saturday	-0.039610
Presence of traffic warden_No	-0.047186
Class of Road_Trunk "C"	-0.183928

Table A8: ARD feature importance for Queue Length (smaller length scale = more relevant)

Feature	Length Scale	Relevance (1/Length Scale)
Number of vehicle arrivals	0.668359	1.496201
Number of potholes	0.849012	1.177839
Number of lanes	1.000046	0.999954
Class of Road_Trunk "C"	1.034086	0.967038
Class of Road_Trunk "A"	1.034086	0.967038
Green time(sec)	2.066782	0.483844
presence of traffic light_Yes	3.146752	0.317788
presence of traffic light_No	3.146752	0.317788
Length of Segment (m)	3.276618	0.305193
Road width (m)	3.348127	0.298674
Presence of traffic warden_No	3.686793	0.271238

Presence of traffic warden_Yes	3.686793	0.271238
Day_Monday	24.822172	0.040287
Day_Sunday	28.561244	0.035012
Day_Wednesday	74.268869	0.013465
Time of Day_Afternoon	100.000000	0.010000
Time Difference (minutes)	100.000000	0.010000
Day_Tuesday	100.000000	0.010000
Day_Thursday	100.000000	0.010000
Traffic demand	100.000000	0.010000
Day_Saturday	100.000000	0.010000
Day_Friday	100.000000	0.010000
Time of Day_Evening	100.000000	0.010000

Table A9: Pearson correlation with VCR

Feature	Correlation
Traffic demand	0.567642
Number of potholes	0.239612
Time of Day_Morning	0.118797
Green time(sec)	0.089622
Number of vehicle arrivals	0.071649
Day_Thursday	0.045073
Day_Monday	0.033744
Class of Road_Trunk "A"	0.021763
Time of Day_Evening	0.021117
presence of traffic light_Yes	0.014282
Length of Segment (m)	0.012643
Presence of traffic warden_No	0.003217
Day_Friday	-0.000006
Presence of traffic warden_Yes	-0.003217
Day_Tuesday	-0.007751
Day_Sunday	-0.010333
presence of traffic light_No	-0.014282
Number of lanes	-0.016621
Day_Wednesday	-0.018270
Class of Road_Trunk "C"	-0.021763
Day_Saturday	-0.040338
Road width (m)	-0.041911
Time of Day_Afternoon	-0.143528
Time Difference (minutes)	-0.759946

Table A10: ARD feature importance for VCR (smaller length scale = more relevant)

Feature	Length Scale	Relevance (1/Length Scale)
Time Difference (minutes)	0.054726	18.272917
Number of potholes	0.776617	1.287636
Traffic demand	0.814231	1.228153
Green time(sec)	1.214924	0.823097
Time of Day_Morning	3.168352	0.315622

Presence of traffic warden_Yes	5.202228	0.192225
Presence of traffic warden_No	5.202228	0.192225
Time of Day_Afternoon	5.535634	0.180648
Day_Thursday	5.812379	0.172047
Day_Wednesday	6.201189	0.161259
presence of traffic light_No	6.262232	0.159687
presence of traffic light_Yes	6.262232	0.159687
Road width (m)	7.099722	0.140851
Day_Monday	7.245357	0.138019
Day_Tuesday	12.078426	0.082792
Length of Segment (m)	14.708998	0.067986
Day_Saturday	26.336100	0.037971
Number of lanes	100.000000	0.010000
Time of Day_Evening	100.000000	0.010000
Day_Sunday	100.000000	0.010000
Class of Road_Trunk "A"	100.000000	0.010000
Class of Road_Trunk "C"	100.000000	0.010000
Number of vehicle arrivals	100.000000	0.010000
Day_Friday	100.000000	0.010000

Acknowledgement
We would like to express my sincere gratitude to all those who supported me throughout the course of this research.

Author Contributions
This manuscript has been read and approved by all the authors. Each of the authors believes that this manuscript represents honest work done by us.

Ethical issues
Not applicable. This study does not involve any experiments on humans or animals. Hence, ethical approval was not required.

Informed consent
Not applicable.

Funding
This research did not receive any external funding like specific grant from funding agencies in the public, commercial, or nonprofit sectors.

Conflict of Interest
The authors declare that they have no conflicts of interest, competing financial interests or personal relationships that could have influenced the work reported in this paper.

Data and materials availability
All data associated with this study will be available based on the reasonable request to corresponding author.

REFERENCES

1. Akinlana DM, Fadikar A, Wild SM, Zuniga-Garcia N, Auld J. O'Hare Airport roadway traffic prediction via data fusion and Gaussian process regression. J Traffic Transp Eng (Engl Ed) 2024;11(4):721-732.

2. Alghamdi M, Akhtar M, Moridpour S, Nazem M. Traffic forecasting in a freeway corridor using a seasonal ARIMA model. In: Proc 42nd Australas Transp Res Forum (ATRF) 2021;10.
3. Attioui M, Lahby M. A Systematic Literature Review of Traffic Congestion Forecasting: From Machine Learning Techniques to Large Language Models. Vehicles 2025; 7(4):142.
4. Bhattacharya S, Kulkarni A, Gupta P. Bayesian Methods in Machine Learning Applications and Challenges. J Artif Intell Res 2025;82:115-40.
5. Cui Z, Zhang Y, Chen X. Uncertainty quantification in traffic prediction with Gaussian processes. Transp Res Part C Emerg Technol 2020;118:102678.
6. Diker AC, Nasibov E. Estimation of traffic congestion level via FN-DBSCAN algorithm by using GPS data. In: Proceedings of the 2012 International Conference on Machine Learning and Applications; Boca Raton, FL, USA. 2012;315-8.
7. Emmanuel AA, Mohammed H. Time headway as indices of traffic congestion. Uniosun J Agric Renew Res 2025;2(1).
8. Essien A, Giannetti C. A deep-learning model for urban traffic flow prediction with traffic events mined from Twitter. World Wide Web 2020;23(2):1173–1193. doi: 10.1007/s11280-020-00800-3
9. Haghighat E, Afandizadeh S, Sharifi A. Deep learning algorithms for traffic forecasting: A comprehensive review and comparison with classical ones. Appl Sci 2024;14(18):8143.
10. Kashyap AA, Raviraj S, Devarakonda A, Nayak K SR, Santhosh KV, Bhat SJ. Traffic flow prediction models: A review of deep learning techniques. Cogent Engineering 2022;9(1):2010510.
11. Kazaura WG. Evaluating the usability and variability of data on transport for modeling: A case study of Dar Es Salaam, Tanzania. Afr J Land Policy Geospatial Sci 2025; 8(4).
12. Khin YL. Understanding the limitations of ARIMA forecasting. Towards Data Sci 2024.
13. Kumar SV. Traffic Flow Prediction using Kalman Filtering Technique. Procedia Engineering 2017;187:582-587.
14. Lahijanian M. Empirical analysis of traffic patterns: Leveraging machine learning for traffic anomaly detection [Internet]. Transp Res Interdiscip Perspect 2024;25:101097.
15. Li Y, Wu X, Gong W, Xu M, Wang Y, Gu Q. Evolutionary competitive multiobjective multitasking: One-pass optimization of heterogeneous Pareto solutions. IEEE Trans Evol Comput 2025;29:2757-2770.
16. Liu X, Chen X, Wang Y. Gaussian process regression for traffic forecasting: A review and empirical evaluation. Transp Res Part C Emerg Technol 2020;116:102635.
17. Liu Y, Wu H, Kang L, Gong Z. Deep learning for short-term traffic flow prediction. Transp Res Part C Emerg Technol 2017;77:103–115.
18. Mihaita AS, Li Z, Singh H, Sharma N, Tuo M, Ou Y. A Comprehensive Review of Traffic Congestion Prediction Models: Machine Learning and Statistical Approaches. 2024.
19. Momin KA, Hamim OF. Short-duration traffic flow prediction using Kalman filtering. In: 6th International Conference on Civil Engineering for Sustainable Development (ICCESD 2022). Khulna, Bangladesh; 2022.
20. Pawel D, Miroslaw M, Maksymilian J. Optimizing traffic volume prediction: Linear regression vs. random forest. Adv Sci Technol Res J 2025;19(11):382-399.
21. Rahmani MKI, Khan S, Ahmed ME, Jawad K. An Intelligent Transport System for Prediction of Urban Traffic Congestion Level. Int J Adv Comput Sci Appl 2024;15(11).
22. Rasmussen CE, Williams CKI. Gaussian processes for machine learning. MIT Press 2006. www.gaussianprocess.org
23. Rodrigues F, Pereira FC. Heteroscedastic Gaussian processes for uncertainty modeling in large-scale crowdsourced traffic data. Transp Res Part C Emerg Technol 2018;95:636-51.
24. Satria S. Traffic congestion prediction using multi-layer perceptrons. In: Proceedings of the International Conference on Informatics and Computing (ICIC). Jakarta, Indonesia 2021;1-6.
25. Sulaiman OS, Biliyamin AI, Abdulrauf TK. Evaluation of traffic congestion in an urban road: A review. Abacus J Eng Appl Sci 2024;2(2):1-15.
26. Tipping ME. Sparse Bayesian learning and the relevance vector machine. J Mach Learn Res 2001;1:211–244.
27. Toba A-L, Kulkarni S, Khallouli W, Pennington T. Long-Term Traffic Prediction Using Deep Learning Long Short-Term Memory. Smart Cities 2025;8(4):126.
28. Ullah S, Chen S, Wu G. Analysis of route-way dynamics in urban traffic congestion using space syntax and remote sensing. Sustainability 2025;5(2):71.
29. Wang H, Liu X, Zhang Y. Probabilistic traffic forecasting with Gaussian processes: A review and case study. IEEE Trans Intell Transp Syst 2021;22(7):4234-4248.
30. Yang C, Yin W, Shi X. Short-term forecasting of transit travel demand: A customized relevance vector machine model. ASCE ASME J. Risk Uncertain. Eng Syst A Civ Eng 2024;10(4): 04024095.