



Traffic-Induced Pollution Modeling at Signalized Intersections in Medium-Sized Urban Areas

Name of the authors: Rama Kanth Angatha¹, Sai Sahitya Kurre^{2*}

Affiliation of the authors:

¹Rama Kanth Angatha

Associate Professor, Department of Civil Engineering,

Marri Laxman Reddy Institute of Technology and Management, Hyderabad-500043, India

Email: ramakanthangatha300@gmail.com

ORCID:0009-0006-3187-954X

^{2*} *Sai Sahitya Kurre (Corresponding Author)*

Assistant Professor, Department of Civil Engineering,

VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, Telangana, India.

E-mail: saisahitya_k@vnrvjiet.in

ORCID: 0000-0002-3684-6855

Received: 15/10/2025

Revised: 22/07/2026

Accepted: 26/08/2026

Abstract

Emissions from vehicles at signalized intersections are one of the major sources of urban pollution, more prevalent in mid-size cities of India due to heterogeneous traffic conditions. The current study

analyzes concentration of six pollutants (CO₂, TVOC, PM_{2.5}, CO, PM₁₀ and HCHO) emitted at intersections of Warangal, Tirupati and Vijayawada. Information was collected on traffic volume, signalization duration, traffic queues and concentration of the mentioned pollutants in several intersections of those three cities. MLR, SVR and ANN models were built based on the collected information as a function of traffic volume and signal parameters. High correlation was detected between the concentration of all the pollutants and the studied parameters, particularly with approach volume and average length of queue. The ANN model provided the highest prediction accuracy among the tested models, with R² higher than 0.85 in all cases. Diurnal patterns indicate maximum emissions occurring in evenings corresponding to maximum traffic flows. Validation of the models proved statistical reliability of the models that can be applied in practice.

Keywords: Traffic emissions, air pollution, signalized intersection, modeling of pollutants, urban transportation, ANN

Introduction

Air pollution is a serious environmental problem in growing cities due to increasing vehicle ownership, rising traffic volume, and congestion which lead to worsening air quality conditions locally. Signalized intersections play an important role in contributing to emissions due to vehicle deceleration, idling, acceleration, and stop-and-go operations. These operating conditions can substantially increase pollutant emissions and localized exposure. Ahn et al. (2002) demonstrated the importance of instantaneous vehicle speed and acceleration in estimating vehicle fuel consumption and emissions, providing an important basis for microscopic traffic-emission modelling. Recent research has increasingly combined traffic simulation, emission models, and dispersion approaches to assess pollution at signalized intersections. Chen et al. (2022) examined the sensitivity of VISSIM parameters in project-level traffic-emission estimation and demonstrated that traffic-simulation and driving-behavior parameters can substantially affect emission estimates. Sun et al. (2021) combined CAL3QHC, ENVI-met, and ANSYS Fluent to assess PM_{2.5} dispersion at an urban signalized intersection and found that pollutant

concentrations were strongly affected by vehicle idling, intersection characteristics, and meteorological conditions. Zhao et al. (2021) developed a CO₂-emission model and signal-timing optimization approach for mixed traffic consisting of fuel and electric vehicles, showing that signal timing can influence both vehicle delay and stop rates and consequently emissions. The effects of signal control and traffic operating conditions on emissions have also been examined using different modelling and optimization techniques. Deschle et al. (2022) quantified CO₂ and NO_x emissions of heavy-duty vehicles at signalized intersections and showed that avoiding unnecessary stops can substantially reduce emissions. Fazzini et al. (2022) evaluated smart traffic-signal control based on multi-agent reinforcement learning and reported reductions in traffic emissions when signalized intersections were coordinated. Marzoug et al. (2022) proposed a cellular automata approach to examine the CO₂, PM, VOC, and NO_x emissions from signalized intersections. Fan et al. (2023a) studied the effects on emission levels from lane configurations for left turn under mixed heavy and light duty traffic, whereas Fan et al. (2023b)

optimized the effects on emissions of signal control strategy for an intersection under mixed traffic conditions. Such studies suggest that signal coordination, lane configuration, traffic interaction, and vehicle types are significant factors affecting emissions from intersections. Real-world measurement studies have shown that pollutants can differ greatly based on the signal phase, traffic level, type of intersections, and weather. Significant fine-scale variations in traffic-related air pollutants at two-phase and four-phase intersections were noted by Zheng et al. (2023), along with distinct variations during the red signal phase and green signal phase. Li et al. (2023) compared PM characteristics at adjacent signalized intersection and roundabout and noted variations in PM_{2.5} and particle number concentrations as well as pedestrian exposure. This shows that the evaluation of pollutants at urban intersections needs to take into account variations not just in time but in space as well.

A recent paper published in *Civil Engineering Infrastructures Journal* by Urmia and Saman (2024) investigated CO, PM_{2.5}, PM₁₀, and noise levels before and after the construction of the Lal Shahbaz Qalandar Underpass in Lahore, Pakistan. Their monitoring covered four seasons and showed elevated pollutant concentrations under high traffic and congestion conditions, with particularly high CO and particulate-matter levels during winter. The study highlights the importance of field-based environmental monitoring for assessing the pollution implications of urban transport infrastructure. The present study builds on this broader intersection and transport-infrastructure pollution context by focusing specifically on traffic-induced pollutant modelling at signalized intersections. Urmia and Saman (2024) is

also included to incorporate a recent related publication from the target journal.

Recent studies have also improved the resolution and predictive capability of intersection-emission assessment. Wang et al. (2024) used roadside LiDAR-derived vehicle trajectories to estimate emissions and investigate their spatial and temporal characteristics at signalized intersections. Šarić (2024) compared several instantaneous emission models for different urban intersection configurations and showed that model selection can influence estimated CO₂, CO, NO_x, and PM₁₀ emissions. Moges and Alemu (2024) investigated the contribution of congested traffic-flow conditions to air pollution at intersections in Addis Ababa and highlighted the effects of traffic characteristics, congestion, road geometry, and meteorological conditions. Roy et al. (2025) developed a multi-task deep-learning framework for jointly predicting traffic emissions and travel delay at signalized intersections, illustrating the potential of machine-learning approaches for capturing complex relationships between traffic operations and environmental outcomes.

However, despite these advancements, a knowledge gap still exists regarding the evaluation of pollutant concentration levels at signalized intersections from the perspective of heterogeneous urban traffic environment. Many of the past research works have considered either a single pollutant, a certain type of vehicle, a certain approach to the intersection, or a certain type of signal control technique. Furthermore, spatial variations among the approaches to the intersection and temporal variations caused by the signal control, queuing, and idling have not been simultaneously taken into account. This research paper is concerned with traffic-related pollution at signalized intersections.

The study aims to (i) determine the concentrations of major traffic-related pollutants, including CO, NO_x, PM_{2.5}, and VOCs, in relation to traffic volumes approaching a signalized intersection from different approaches; (ii) evaluate the influence of signal parameters and queuing characteristics under heterogeneous traffic conditions; and (iii) develop predictive models using machine-learning and dispersion-based approaches to support intersection-level pollution assessment and mitigation.

2. Methodology

A thorough technique is created to investigate the impact of traffic volume on urban air quality at various phases. To determine the study's goal, a comprehensive literature review is the first step in the technique. The second step of the technique will involve choosing study

regions and signalized junctions in particular cities. Data collection at intersections, traffic volume and signal data (red light duration, cycle time duration, and number of queued cars) at various approaches to intersections, and pollution concentration data are all included in the third step of the process. The fourth stage of the methodology includes analysis of pollution data, signal data, and volume data. In the fifth stage of the methodology, models to predict the concentration of major pollutants based on various factors will be developed at the signalized intersections. Steps involved in conducting the current study are illustrated in Figure 1.

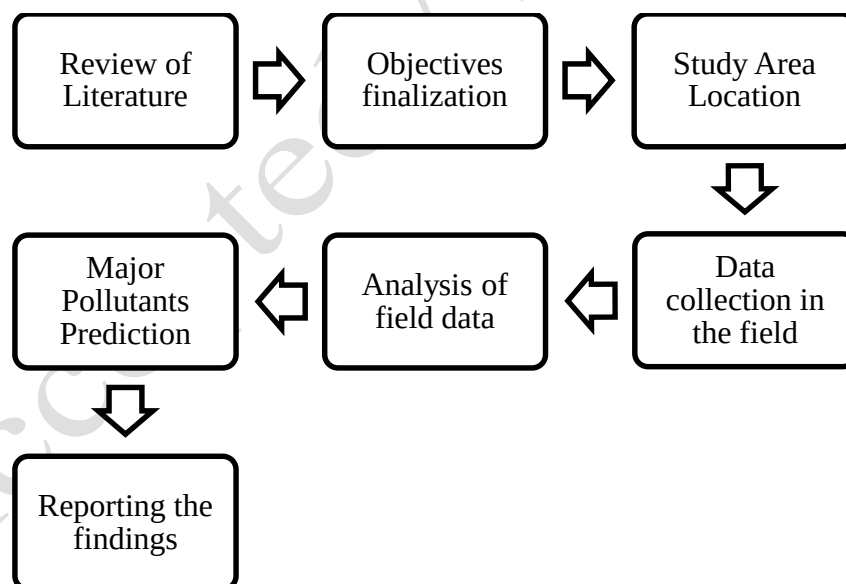


Fig. 1. Step wise procedure for the methodology used in the study

In the current study, various portable devices like CO meter, laser particle multi-functional detector and air quality detector were used. All air quality measuring

devices that were employed in the current research were factory-calibrated, and their accuracy was verified using the span test and the zero test prior to conducting field

experiments. The measurements were performed at a height of 1.5 meters above the ground level to reflect human exposure

levels. The list of hand-held devices that were used to measure pollutants concentrations is presented in Table 1.

Table 1. Equipment details used to measure pollutant concentrations

Parameter	Laser Particle Multi-Functional Detector	CO Meter	Air Quality Detector
Pollutants Measured	HCHO, TVOC	CO	CO ₂ , PM _{2.5} , PM ₁₀
Sensor Type / Principle	Electrochemical (HCHO) + Laser sensor	Electrochemical sensor	NDIR (CO ₂) + Laser sensor
Measurement Range	0–5 mg/m ³	0–1000 ppm	CO ₂ : 0–5000 ppm; PM: 0–999 µg/m ³
Resolution	0.01 mg/m ³	1 ppm	CO ₂ : 1 ppm; PM: 0.1 µg/m ³
Accuracy	±5% ±10%	±5% ±10 ppm	±5% ±10%
Operating Temperature	0°C–50°C	0°C–50°C	0°C–50°C
Key Features	Measures HCHO & TVOC simultaneously	Portable, real-time LCD display	Multi-parameter display with AQI

The data was collected on three successive days during the weekdays at each intersection from 7:00 AM until 9:00 PM, involving both the morning and evening peak hours as well as midday periods. Meteorological factors like temperature, wind speed, and humidity were used systematically within the analysis. The data

3. Site Selection

Intersections with three legs and four legs are chosen as case studies. Field data will be gathered in three Indian cities – namely, Warangal, Vijayawada and Tirupathi.

collected during extreme climatic conditions like rainfall and wind speeds beyond 15 km/h were not taken into account because of the possible errors associated with dispersion. Among the above meteorological factors, temperature was one of the selected explanatory variables within the model.

Three signalized intersections are chosen for field data gathering from each chosen city. Map or area of the chosen cities for the study is illustrated in Figure 2 below:

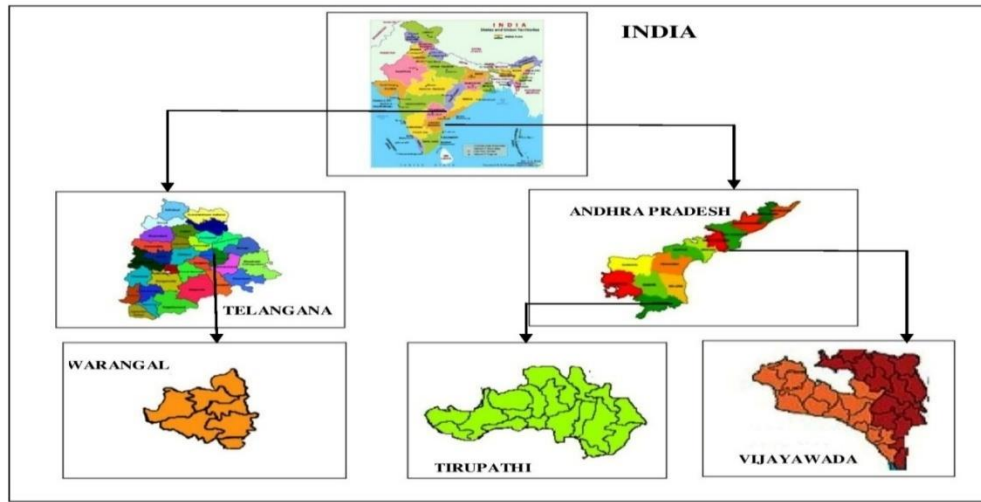


Fig. 2. Cities chosen for the research

The data related to various signalized intersections selected are provided in Table

Table 2. Signalized intersections chosen for the study

Name of the city	Intersection ID	Name of the intersection
Warangal	I	Kazipet railway station junction
	II	Spencors circle
	III	GeethaBhavan Circle
Tirupathi	IV	Central Park intersection
	V	Leela Mahal circle
	VI	Annamayya circle
Vijayawada	VII	BRTS junction
	VIII	Kothavanthena junction
	IX	Ramesh hospital intersection

2. The images of signalized intersections are provided in Figures 3 to 5.



Fig. 3. Spencors intersection in Warangal city (Intersection II)



Fig. 4. Annamayya circle intersection in Tirupathi city (Intersection VI)



Fig. 5. Kothavanthena junction in Vijayawada city (Intersection VIII)

4. Data collection

For the current study, a variety of data types, including traffic volume and signal data, pollutant data, and air quality index data from signalized crossings, have been collected. Using video cameras, traffic volume (V) data was gathered during an eight-hour period for various approaches to signalized crossings. The recorded movies were played on a wide-screen display in order to gather the classified traffic volume data. In field conditions, the red time and cycle duration have been noted at various intersections. From the recorded footage, the queued cars at different approaches at several junctions have been carefully gathered. Table 3 contains information on the roads at each chosen signalized intersection. At signalized intersections, six pollutants have been detected, including

The selected intersections represent a diverse range of traffic and geometric conditions, including three-legged and four-legged configurations, varying traffic volumes, and different surrounding land-use characteristics (commercial and mixed-use areas). The cities of Warangal, Vijayawada, and Tirupati were chosen as representative medium-sized Indian urban environments characterized by heterogeneous traffic conditions. This selection ensures that the developed models are generalizable and applicable to similar urban contexts.

CO₂, TVOC, PM_{2.5}, CO, PM₁₀, and HCHO. The concentration of the pollutants was monitored through hand held instruments. Data collected from all the pollutants' concentrations were recorded every 5 minutes and then calculated per 1 hour interval to determine average and total concentration. The calculation of 5-minutes pollution concentrations into 1 hour average was done to minimize short term variations and adhere to standard practice in the field (CPCB and WHO). The process of averaging is also very important in machine learning-based predictions. In the current study, various hand-held instruments like CO meter, Laser Particle Multi-Functional Detector, and Air Quality Detector were utilized.

Table 3. Details of the road geometry at signalized intersections

Intersection ID	Intersection name	No. of approaches	No. of lanes
I	Kazipet railway station junction	4	6
II	Spencors circle	4	6
III	GeethaBhavan Circle	4	6
IV	Central Park intersection	3	4
V	Leela Mahal circle	4	6
VI	Annamayya circle	4	6
VII	BRTS junction	4	6
VIII	Kothavanthena junction	4	6
IX	Ramesh hospital intersection	4	6

5. Field data analysis

Data on the pollutant's concentration, traffic flow, signals, queues of vehicles, and the Air Quality Index are collected from the field for particular signalized intersections and analyzed.

5.1 Analysis of traffic volume data

The videography technique was adopted to obtain data on the traffic volume for various approaches at signalized intersections. The classified data of the traffic volume was obtained from the video using the large screen. In this case, the five types of vehicles that were taken into consideration

in the study were 2W, Car, 3W, LCV, and HV. In order to convert the heterogeneous traffic volume into the homogeneous traffic volume, multiplication of each 5 minutes data with the corresponding PCU of the vehicle was done. The PCU of the vehicles used in this case is based on Indo-HCM (2017). The descriptive analysis of approach traffic flow is conducted, and the information like lower value, higher value, and mean value of approach traffic flows are presented in Table 4 at each signalized intersection. Due to the scale of the city's signalized intersections, Vijayawada's mean approach traffic flow is observed to be high. The major National Highways are connected to the signalized intersections in Vijayawada city that were taken into consideration for this study.

Table 4. Descriptive statics of the approach traffic volume

Signalized intersection	Mean approach Traffic volume (PCU/hr)	Range of approach Traffic Volume (PCU/hr)	
		Lower	Higher
I	4205	3780	4597
II	3180	2906	3512
III	5117	4718	5599
IV	1435	1088	1904
V	4032	3284	4753
VI	2173	1778	2608
VII	4535	2298	5817
VIII	4982	2698	6081
IX	3085	1813	3581

The time of the day and its effect on traffic volumes have also been studied to determine how the volume of traffic changes with time of the day. Figure 6 shows how traffic volume varies throughout the day at several signalized

crossings. According to this figure, junction III experiences the highest volume of traffic of 5921 PCU/hr between 6:00 and 7:00 p.m. There is a lot of business activity at this specific crossroads in the evening, which is why there is a lot of traffic.

Between 7:00 and 8:00 AM, intersection IV has a low volume of traffic of 1088 PCU/hr.

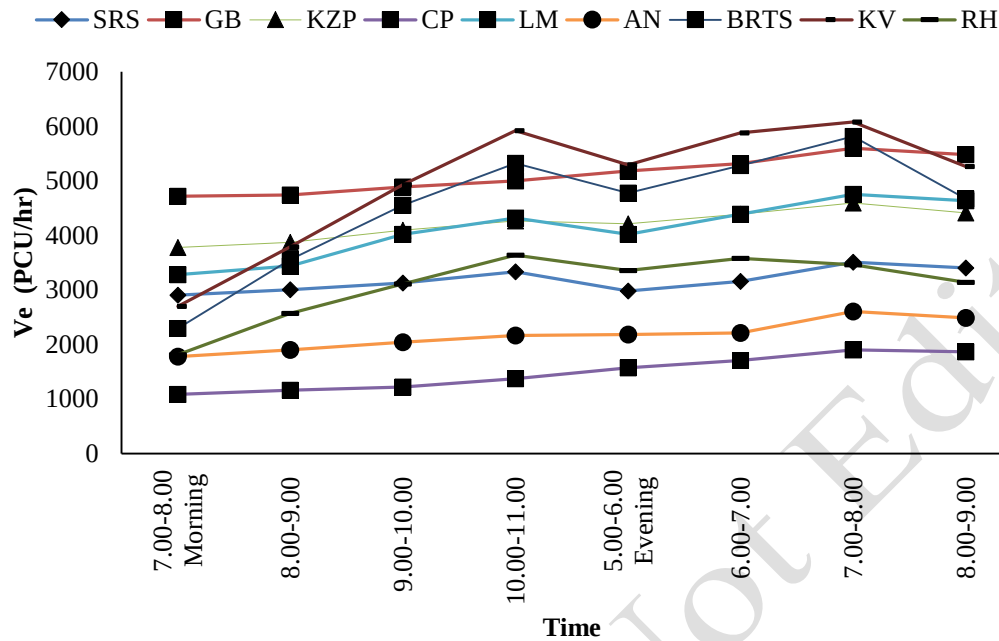


Fig. 6. Time-of-Day Variation in Traffic Volume Across Approaches

5.2 Analysis of pollution concentration

Variation in the concentration of different air pollutants over time, calculation of the concentration of pollutants, and analysis of the concentration of pollutants over time are all included in the pollution data analysis. Six significant pollutants have been discovered in the study sites: CO₂, TVOC, PM_{2.5}, CO, PM₁₀, and HCHO. Handheld devices are used in the field to gather data on pollution concentrations. At the

signalized junctions, the concentration of each pollutant was recorded every hour and monitored every five minutes. Figure 7 shows the fluctuation in CO concentrations at various times of the day at particular mid-block areas. Similarly, research has been done on how the concentrations of CO₂, HCHO, TVOC, PM_{2.5}, and PM₁₀ change over time.

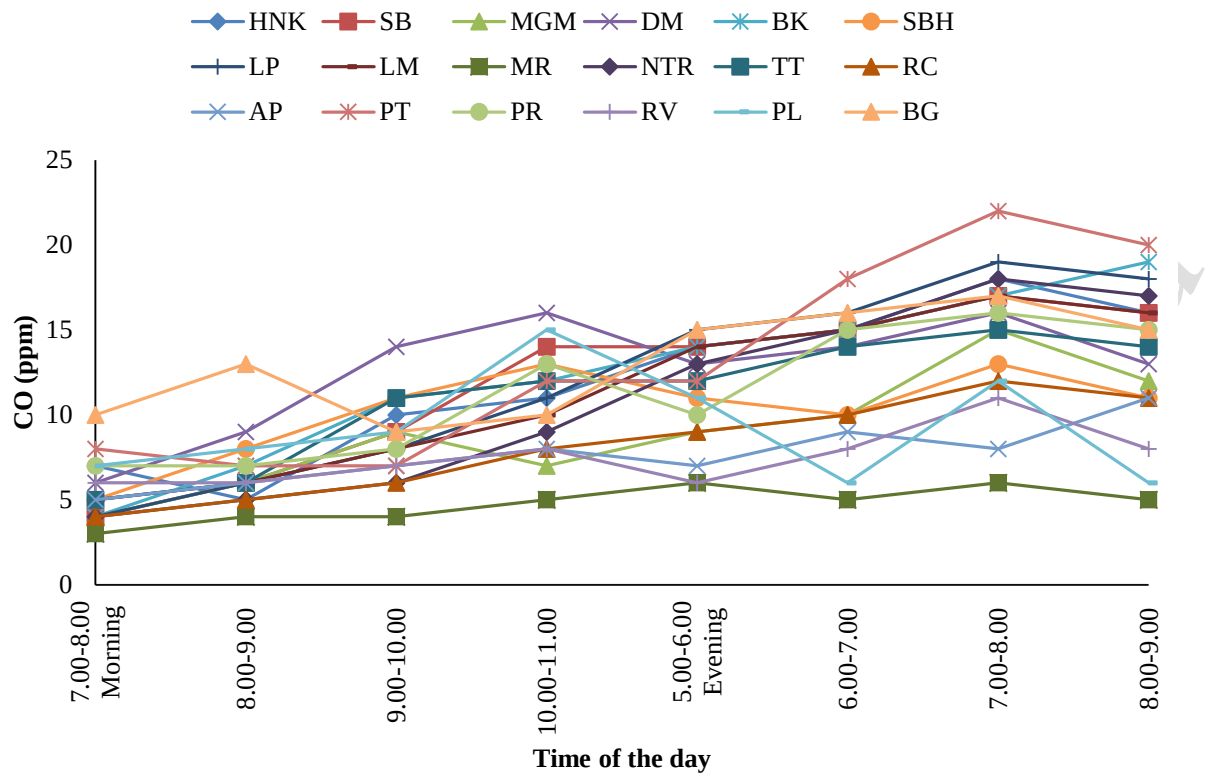


Fig. 7. Temporal Changes in CO Concentrations Throughout the Day

6. Pollution prediction models

The variation of concentration for each pollutant with respect to the volume of traffic flow is obtained through correlation analysis. Correlation analysis has been done to study the relationship between different pollutants concentrations and traffic volume at various junctions. Various models are established in order to predict concentration for various pollutants using MLR, SVR and ANN models. Concentration for each pollutant is the dependent variable while red time length (RT), cycle time length (CT), average

queue length (N_q) and traffic volume (V) are the independent variables in the model.

6.1 Development of MLR models

The correlation between pollution concentration and signal parameter strength is evaluated by means of correlation analysis. Table 5 below shows the r-values derived from correlation analysis.

Table 5. Correlation Coefficients Between Air Pollutant Concentrations and Traffic Signal Characteristics

	V_e	T	RT	CL	N_q
CO	0.71	0.43	0.39	0.42	0.65
CO ₂	0.51	0.54	0.52	0.41	0.58
HCHO	0.75	0.41	0.35	-0.51	0.36
TVOC	0.77	0.43	0.41	0.52	0.39
PM _{2.5}	0.84	0.35	0.32	0.35	0.41
PM ₁₀	0.82	0.39	0.41	0.42	0.42

Pollutant concentrations of CO₂, TVOC, PM_{2.5}, CO, PM₁₀, and HCHO are predicted using the MLR approach. Six crossings with Intersection IDs I, II, IV, V, VII, and VIII provided field data from which the models were developed. Warangal is home to intersections I and II, whereas Tirupathi is home to intersections IV and V. Similarly, Vijayawada is located at crossroads VII and VIII.

Each pollutant's concentration level serves as the model's dependent variable. The model takes into account the following input variables: temperature (T), approach

traffic volume (V_e), average queued vehicles (N_q), cycle time length (CL), and red time duration (RT). The mean number of cars that are stopped during each red time period in various intersection approaches is known as average queuing vehicles (N_q). The total number of cars that approach the intersection in a certain amount of time is known as approach traffic volume (V_e).

Equations (1) through (6) display the MLR models built for each pollutant. Table 6 displays the regression results of the models that were created.

$$CO = 4.64 + 0.05 \times N_q + 0.02 \times V_e \quad (1)$$

$$CO_2 = 853.3 - 0.56 \times RT + 1.53 \times N_q - 10.13 \times T + 0.06 \times V_e \quad (2)$$

$$HCHO = -0.027 - 0.0001 \times CL + 0.0005 \times V_e \quad (3)$$

$$TVOC = 0.01 + 0.0003 \times CL + 0.0003 \times V_e \quad (4)$$

$$PM_{2.5} = 8.97 + 0.01 \times V_e \quad (5)$$

$$PM_{10} = 13.77 + 0.02 \times V_e \quad (6)$$

Table 6. Performance and Regression Coefficients of the MLR Models

Dependent variable	Independent variable	p-value	t-stat value	R ²	Adjusted R ² value
CO (ppm)	Intercept	0.000	5.95	0.87	0.87
	Average queued vehicles	0.000	6.99		

CO ₂ (ppm)	Approach traffic volume	0.000	14.90		
	Intercept	0.000	8.69	0.83	0.81
	Red time length	0.001	-3.50		
	Average queued vehicles	0.006	2.85		
	Temperature	0.023	-2.35		
HCHO (mg/m ³)	Approach traffic volume	0.000	8.09		
	Intercept	0.000	-4.81	0.87	0.88
	Cycle length	0.007	-2.79		
	Approach traffic volume	0.000	14.38		
TVOC (mg/m ³)	Intercept	0.001	2.98	0.78	0.77
	Cycle length	0.000	3.48		
	Approach traffic volume	0.000	5.99		
PM _{2.5} (µg/m ³)	Intercept	0.018	2.44	0.79	0.78
	Approach traffic volume	0.000	4.07		
PM ₁₀ (µg/m ³)	Intercept	0.036	2.15	0.80	0.80
	Approach traffic volume	0.000	13.92		

The combined impact of the predictor factors on the pollution concentration is examined. The suggested model for estimating the concentration of contaminants in signalized junctions produced R² values ranging from 0.77 to 0.87, indicating satisfactory predictive ability, according to MLR. The approach traffic volume, average vehicle queue length, and cycle length are significant factors influencing the pollution concentration, according to the results. The findings show that temperature has an impact on CO₂ concentration.

6.1.1 Models validation

The Multiple Linear Regression (MLR) models developed were tested on independent field data that were collected at three intersections equipped with traffic signals. The performance of the MLR models was assessed by calculating the coefficient of determination (R²), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The validation results of each pollutant are summarized in Table 7 below.

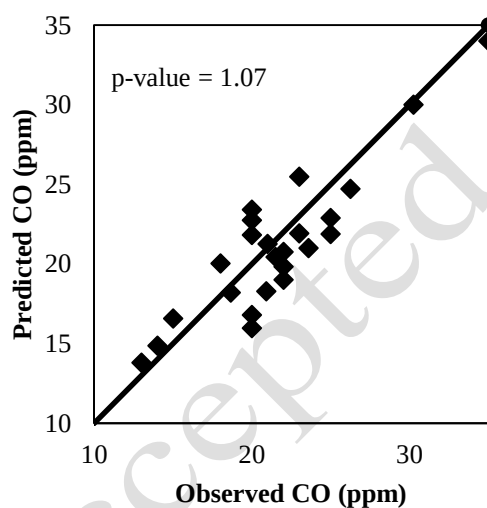
Table 7. Model Validation Metrics for the Multiple Linear Regression (MLR) Models.

Pollutant	R ²	MAE	MAPE (%)
CO	0.82	0.21	2.5

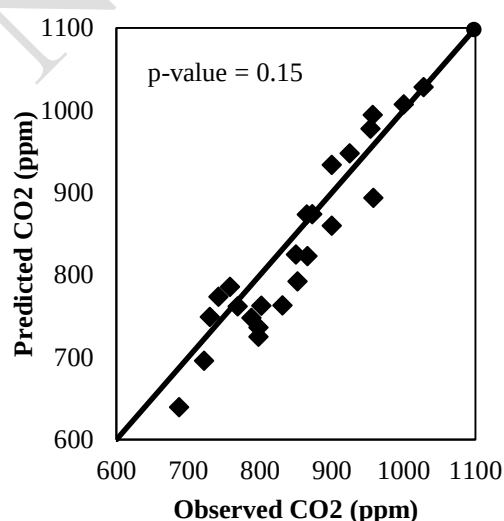
CO ₂	0.80	0.18	2.2
HCHO	0.85	0.24	3.0
TVOC	0.78	0.16	1.9
PM _{2.5}	0.77	0.13	1.5
PM ₁₀	0.79	0.18	2.2

The obtained MLR models yielded R^2 value from 0.77 to 0.85, which demonstrates good agreement between the experimental and predicted concentrations of pollutants. At the same time, MAE values were from 0.13 to 0.24 and MAPE values varied from 1.5% to 3.0%, which indicates the satisfactory accuracy of predictions of concentrations of all investigated

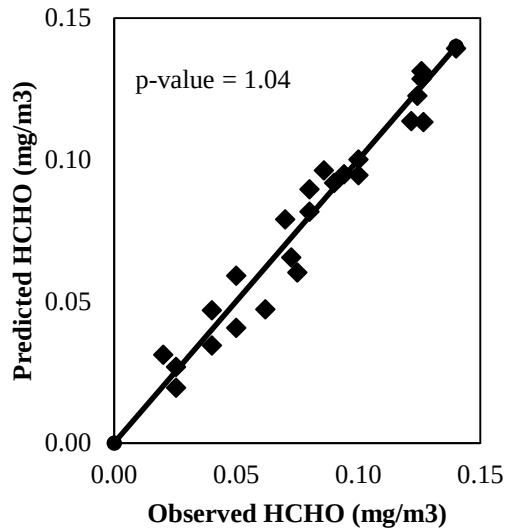
pollutants. Moreover, the graphs in Figure 8 compare the observed and predicted concentrations of CO, CO₂, HCHO, TVOC, PM_{2.5}, and PM₁₀. The good agreement of experimental and predicted values of concentrations, indicated by the proximity of the data points to the line of agreement, once again proves the effectiveness of the obtained MLR models.



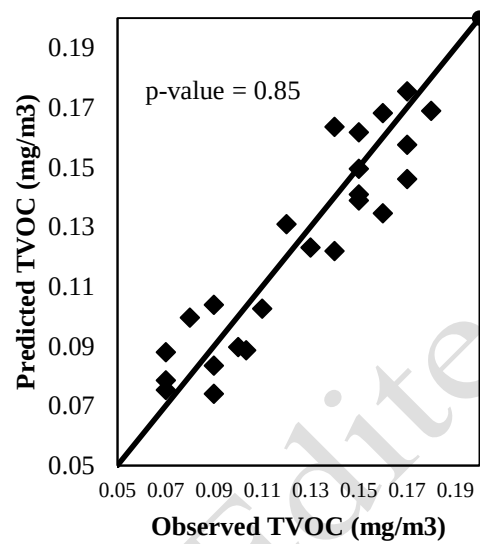
(a) CO



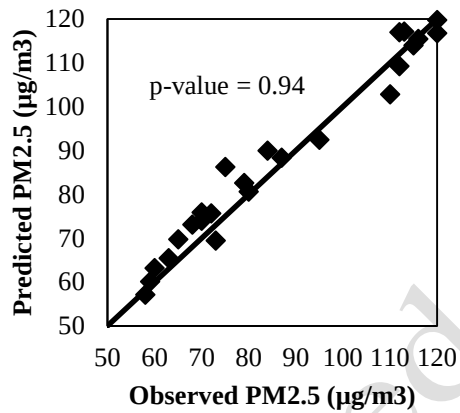
(b) CO₂



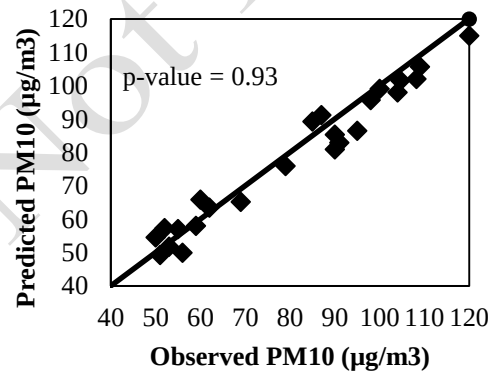
(c) HCHO



(d) TVOC



(e) PM_{2.5}



(f) PM₁₀

Fig. 8. Model Validation Plots for different Pollutants.

6.2 Development of SVR Models

The current study has used the Support Vector Regression (SVR) model. The modeling technique known as Vector Regression (SVR) offers a strong framework for carrying out the regression. When handling datasets with several dimensions and non-linear relationships between variables, the SVM method has shown itself to be quite beneficial. In order to forecast the amounts of different contaminants in the middle block portions

of the road, the current study used the Support Vector Regression (SVR) model in the R software.

Thirty percent of the data is utilized for testing and seventy percent is used for training in the SVR modeling approach. The eps-regression SVM technique is used to perform the regression. The radial SVM kernel function is used in this study's regression. The 95% confidence interval was used to perform the regression.

```

1 d=CO_SI
2 summary(d)
3 y=d$CO
4 y=
5 attach(d)
6 r=data.frame(y,AVCR,ATV)
7 r=
8 #data partitioning
9 n=nrow(r)
10 trainIndex = sample(1:n, size = round(0.7*n), replace=FALSE)
11 train = r[trainIndex,]
12 test = r[-trainIndex,]
13 #linear regression
14 md<-lm(y~.,train)
15 py<-predict(md,test)
16 summary(md)
17 #svm regression
18 library(e1071)
19 library(caret)
20 model=svm(y~., data=train)
21 summary(model)
22 predicted<- predict(model, test)
23 x=1:length(test$y)
24 plot(x,test$y,pch=18, col="red")
25 lines(x,predicted,lwd="1",col="blue")
26 predictedY1<-predict(model, train)
27 library(caret)
28 mae=MAE(test$y,predicted)
29 r2=R2(test$y,predicted)
30

```

Fig. 9. Sample Input Data for the SVM Models at Signalized Intersections

Red light duration (RT), cycle time (CL), number of queues (Nq), and approach traffic (Ve) are the predictors in each model. The concentration of contaminants

is one of the dependent variables. Table 8 displays the outcomes of the statistical tests carried out during the creation of the support vector regression models.

Table 8. Regression Results of the Developed Support Vector Regression (SVR) Models for Signalized Intersections

Dependent variable	Independent variables	Co-efficient	R ²	Adjusted R ²
CO (ppm)	Intercept	3.98	0.90	0.89
	Average queued vehicles	0.05		
	Approach traffic volume	0.02		
CO ₂ (ppm)	Intercept	754.3	0.85	0.86
	Red time length	-1.22		
	Average queued vehicles	1.20		
	Temperature	-9.03		
HCHO (mg/m ³)	Approach traffic volume	0.07	0.91	0.92
	Intercept	0.001		
	Cycle length	-2.62		

TVOC (mg/m ³)	Approach traffic volume	3.94		
	Intercept	0.001	0.81	0.81
	Cycle length	0.0003		
PM2.5 (µg/m ³)	Approach traffic volume	0.0002		
	Intercept	7.59	0.84	0.85
	Approach traffic volume	0.01		
PM10 (µg/m ³)	Intercept	12.54	0.85	0.84
	Approach traffic volume	0.02		

For every Support Vector Machine (SVM) model, graphs displaying the anticipated pollutant values were created. The number of test samples is displayed on the x-axis of the graphs, while the predicted value of the pollutant concentration is displayed on the y-axis. The Support Vector Regression (SVR) model predicts the expected values of various pollutant concentrations, which

are represented by each point on the graph. The expected values' trend is depicted by the trend line. The model output makes it evident that the expected values closely resemble the trend line. Figure 10 below provides an example of expected HCHO values.

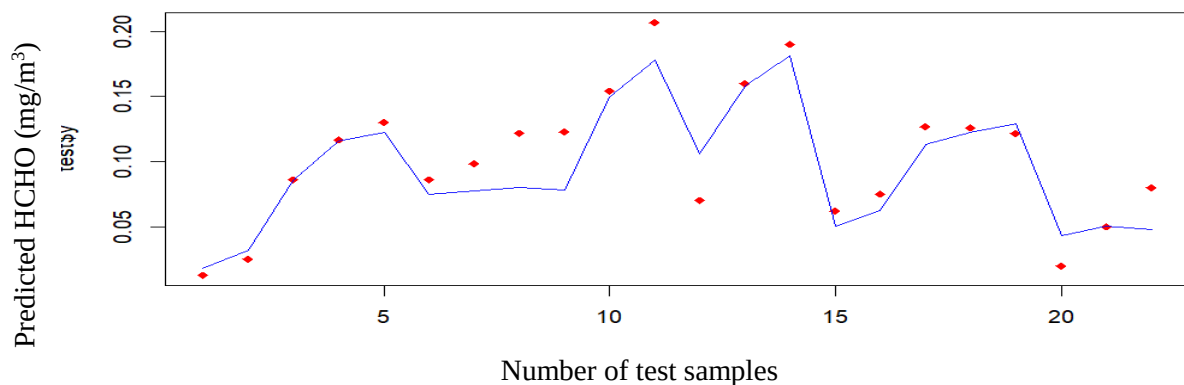


Fig. 10. HCHO Concentrations Predicted by the SVR Model

6.3 Development of ANN Models

Additionally, the current work employed a variety of signal parameters to determine the concentration of pollutants using the ANN technique. Approach traffic volume

(Ve), temperature (T), pollution concentration (RT), cycle time length (CL), average queued vehicles (Nq), and other signal parameters are used in the current study to use ANN.

Ten neurons (input layer and hidden layer) and an output layer serve as input layers that decide the neural network architecture

for the models built from signalized intersections. Figure 11 below provides an example of the neural network configuration for the CO model. For other contaminants, the same procedure is used.

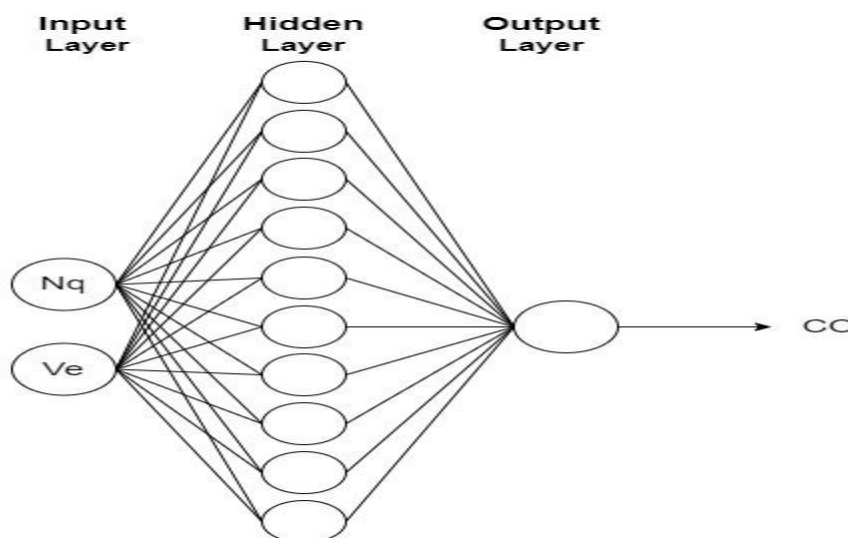
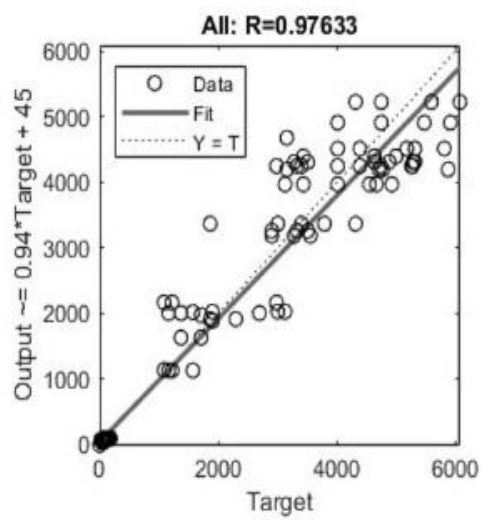


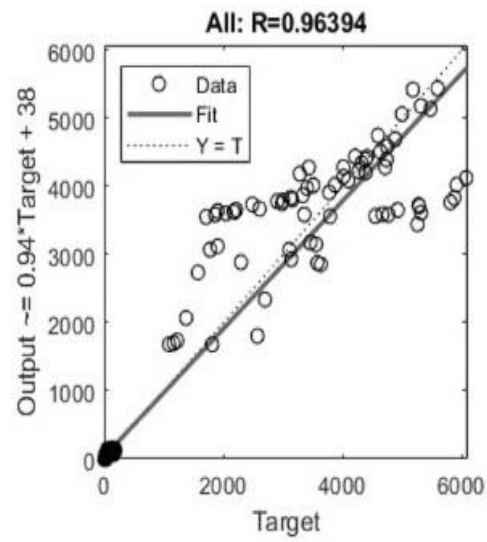
Fig. 11. Neural Network Architecture of the CO Prediction Model

ANN models have been created using a feedforward backpropagation architecture, having one hidden layer with 10 neurons. A sigmoid function was chosen as an activation function for the hidden layer while a linear function for the output layer. Training of the models has been done using the Levenberg–Marquardt algorithm with 70% training, 15% validation, and 15% testing data sets. Training of the models has been done using the software MATLAB,

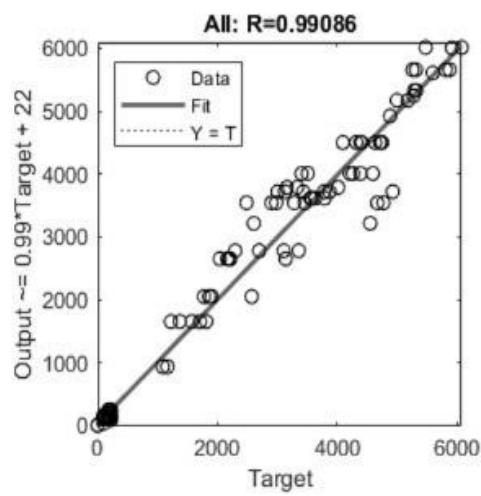
where the performance has been considered according to minimization of mean squared error. Various changes in the number of hidden layers and neurons have been made in multiple iterations in order to find the best configuration of the network. The network that minimizes the discrepancy between the real and ANN models is considered ideal. Figure 12 displays the results of the ANN models.



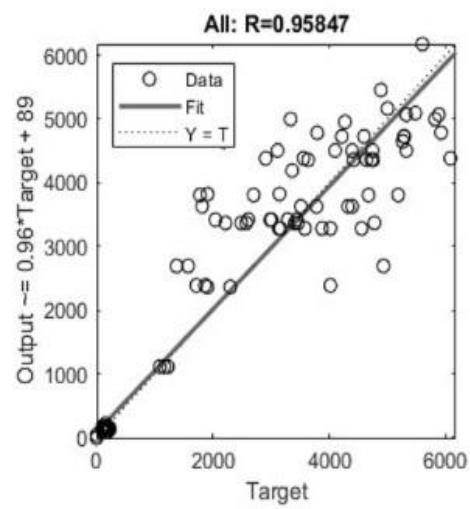
(a)



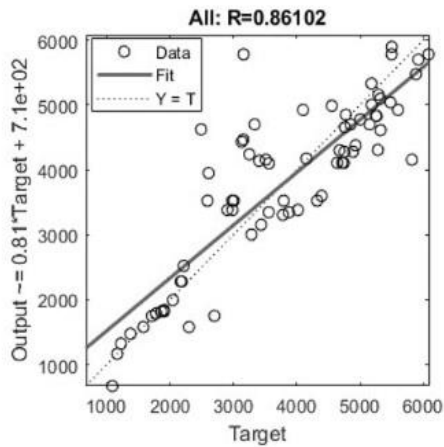
(b)



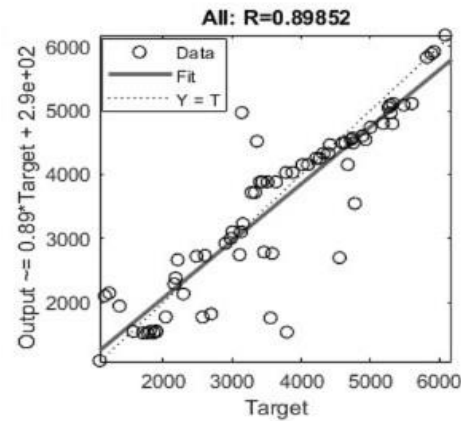
(c)



(d)



(e)



(f)

Fig. 12. ANN Model Outputs for (a) CO, (b) CO₂, (c) HCHO, (d) TVOC, (e) PM_{2.5}, and (f) PM₁₀

The correlated values between the estimated and measured pollutants are displayed in the ANN graphs. The measured values are displayed on the x-axis. The neural network model's estimated values are shown on the y-axis. The r value indicates the relationship between the actual and ANN calculated values. The likelihood that the discrepancy between the measured and estimated values is not significant rises with the r value. When the acquired points are near the Y=T line, there are no significant variations. The estimated and measured values will never differ significantly.

6.4 Comparison of MLR, SVM and ANN models developed for signalized intersections

In this section, a comparative analysis of the models created using MLR, SVM, and ANN is given to determine how accurately these models predict the pollution concentrations in regard to the traffic flow. The difference between the estimated and actual values gives an insight into the ability of the different approaches in predicting. Statistical estimates such as R² value and Mean Absolute Percentage Error (MAPE) and Mean Absolute Error (MAE) are used to determine precision and estimate errors. The model with high R² value and low MAPE, MAE values provide accurate predictions. R², MAPE, and MAE values calculated for these models are presented in Table 9.

Table 9. Comparative Performance Evaluation of MLR, SVR, and ANN Models Based on R², MAPE, and MAE

Pollutant	Model	R ²	MAPE (%)	MAE
CO	MLR	0.87	2.5	0.21
	SVR	0.90	2.0	0.17
	ANN	0.94	1.2	0.10
CO ₂	MLR	0.83	2.2	0.18
	SVR	0.85	2.0	0.16
	ANN	0.92	0.5	0.05
HCHO	MLR	0.87	3.0	0.24

TVOC	SVR	0.91	2.5	0.20
	ANN	0.98	1.2	0.09
	MLR	0.78	1.9	0.16
PM _{2.5}	SVR	0.81	1.5	0.12
	ANN	0.90	1.1	0.09
	MLR	0.79	1.5	0.13
PM ₁₀	SVR	0.84	1.2	0.10
	ANN	0.86	1.0	0.08
	MLR	0.80	2.2	0.18
	SVR	0.85	1.5	0.13
	ANN	0.87	0.3	0.03

It is found that high values of R^2 and low values of MAPE, MAE have been achieved by using the ANN approach to estimate the concentration of pollutant values at road mid-block segments. This implies that the ANN approach can estimate the concentration of pollutant values more efficiently than other models such as MLR and SVR. The performance of ANN models was superior because the ANN network architecture has hidden layers. The hidden layers can establish complex interaction between signal parameters and traffic flow as well as concentration of pollutants.

7. Results and Discussion

The field observation study conducted on nine signalized intersections in the three locations – Warangal, Tirupati, and Vijayawada – demonstrated obvious temporal and spatial differences in traffic and pollutant concentration. Traffic flow was between 1,088 and 6,081 PCU/hr, with the greatest traffic and the maximum concentration of pollution being observed at 6:00–8:00 PM, specifically at the Intersection III. The correlation analysis has shown that all the six pollutants, including CO₂, TVOC, PM_{2.5}, CO, PM₁₀, and HCHO, were positively correlated with traffic volume and queue lengths, thus, indicating the effect of traffic build-up and idling on the pollutants' concentration. As for model development, it should be noted that the predictive power increased in the sequence of MLR, SVR, and ANN models,

where ANN produced the best results ($R^2 > 0.85$), followed by SVR and MLR.

The results provide confirmation of the fact that the pollution at the Indian signalized intersections is highly dependent on the heterogeneity of the traffic behavior where such phenomena as idling, stop-go traffic flow, and long queues lead to high emissions of the pollutants in the peak periods. The domination of such factors as traffic volume, queue length, and cycle time in predictions reflects well the results obtained in previous research and proves once again the significance of operation variables in the near road air quality evaluation. The superiority of ANN models in their prediction ability is determined by their capability of capturing the nonlinear and complex interrelations between the traffic flow, signalization, and pollutant generation that cannot be captured by such models as MLR.

This study's strong correlation between the traffic volume (approach) and pollution concentration is similar to the results of Pandian et al. (2009) in their analysis of Indian intersections. The effect of queue length and signal cycle time is further corroborated by Sahoo et al. (2021), who identified that idling and stop-and-go operation significantly increase particulate matter and carbon monoxide emissions. The better performance of ANN models is further justified by Mehta et al. (2023) and Zito (2009) in their finding that ANN

outperforms the linear model in capturing the nonlinearity in the emissions process.

8. Conclusions

These research findings indicate the statistically significant relation between the traffic volume and pollutant concentrations in the signalized urban intersections. Prediction models that used approach traffic volume, red signal time, cycle time and average number of queued vehicles were constructed based on Multiple Linear Regression (MLR), Support Vector regression (SVR) and Artificial Neural Network (ANN). The prediction was made for concentrations of CO₂, TVOC, PM_{2.5}, CO, PM₁₀ and HCHO. The output variable was a pollutant concentration. The independent variables which were used in predictive models showed their significance. The validation was performed via Chi-square test and indicated good accuracy of the models. The results showed the existence of diurnal variation in pollutant concentrations from 6 PM to 9 PM when there was an increase in concentrations of CO, HCHO, TVOCs, and particulate matter. The highest pollutant concentrations were observed during the

evening rush hour when the traffic volume varied from 6000 PCU/hr and the pollutant concentration from 1301 ppm from 7 PM to 8 PM. At this time, there was an increase of about 46% of pollution level while the increase in traffic volume from morning to evening was 80%. Among the key influence factors of pollutants, approach traffic flow, cycle time, and the average number of vehicles in queue at the signalized intersection were considered. It should be noted that the analysis demonstrated better results obtained by ANN models, since they had higher R² values as opposed to MLR and SVR models. The results of this study serve as a valuable source of information for transportation planners and policy makers, stressing the importance of managing the rising traffic volumes in urban locations. The model that has been created is an effective means of pollution prediction and may be improved by adding some other variables such as traffic signal aids, traffic countdowns, etc. The results obtained provide a basis for developing advanced traffic management strategies that are intended to reduce pollution in urban environment.

Acknowledgments

Acknowledgment: The authors would like to confirm the utilization of AI-based language editing and clarity enhancement tool for their manuscript. Authors have thoroughly gone through the content and are wholly responsible for the accuracy, novelty, and scientific integrity of their work.

References

- Ahn, K., Rakha, H., Trani, A. and Van Aerde, M. (2002). "Estimating vehicle fuel consumption and emissions based on instantaneous speed and acceleration levels", *Journal of Transportation Engineering*, 128(2), 182-190, [https://doi.org/10.1061/\(ASCE\)0733-947X\(2002\)128:2\(182\)](https://doi.org/10.1061/(ASCE)0733-947X(2002)128:2(182)).
- Chen, Y., Wen, C., Jiang, C. and Jiang, X. (2022). "Global sensitivity analysis of VISSIM parameters for project-level traffic emissions: a case study at a signalized intersection", *Environmental Technology*, 43(24), 3801-3820, <https://doi.org/10.1080/09593330.2021.1934737>.

- Deschle, N., van Ark, E.J., van Gijlswijk, R. and Janssen, R. (2022). "Impact of signalized intersections on CO₂ and NO_x emissions of heavy-duty vehicles", *Energies*, 15(3), 1242, <https://doi.org/10.3390/en15031242>.
- Fan, J., Li, A., Ilahi, A., Gao, K., Najafi, A., Sarang, J. and Li, T. (2023a). "Emission impacts of left-turn lane on light-heavy-duty mixed traffic in signalized intersections: optimization and empirical analysis", *Heliyon*, 9(5), e16260, <https://doi.org/10.1016/j.heliyon.2023.e16260>.
- Fan, J., Najafi, A., Sarang, J. and Li, T. (2023b). "Analyzing and optimizing the emission impact of intersection signal control in mixed traffic", *Sustainability*, 15(22), 16118, <https://doi.org/10.3390/su152216118>.
- Fazzini, P., Torre, M., Rizza, V. and Petracchini, F. (2022). "Effects of smart traffic signal control on air quality", *Frontiers in Sustainable Cities*, 4, 756539, <https://doi.org/10.3389/frsc.2022.756539>.
- Li, W., Qiu, Z. and Wang, X. (2023). "Comparison of PM spatiotemporal variations and exposure at adjacent signalized intersection and roundabout", *Urban Climate*, 50, 101590, <https://doi.org/10.1016/j.uclim.2023.101590>.
- Marzoug, R., Lakouari, N., Pérez Cruz, J.R. and Vega Gómez, C.J. (2022). "Cellular automata model for analysis and optimization of traffic emission at signalized intersection", *Sustainability*, 14(21), 14048, <https://doi.org/10.3390/su142114048>.
- Moges, G.T. and Alemu, G.T. (2024). "Contribution of congested traffic flow condition to air pollution at intersections in Addis Ababa, Ethiopia", *Heliyon*, 10(17), e36477, <https://doi.org/10.1016/j.heliyon.2024.e36477>.
- Roy, K., Chan, L.S., Zhang, X. and Nassir, N. (2025). "Multi-task deep learning for joint prediction of traffic emissions and travel delay", *Transportation Research Part D: Transport and Environment*, 146, 104846, <https://doi.org/10.1016/j.trd.2025.104846>.
- Šarić, A. (2024). "Consistency of different instantaneous emission models for assessing the environmental impact of urban intersections", *Urban Climate*, 57, 102115, <https://doi.org/10.1016/j.uclim.2024.102115>.
- Sun, D.J., Wu, S., Shen, S. and Xu, T. (2021). "Simulation and assessment of traffic pollutant dispersion at an urban signalized intersection using multiple platforms", *Atmospheric Pollution Research*, 12(7), 101087, <https://doi.org/10.1016/j.apr.2021.101087>.
- Urmia, A. and Saman, S. (2024). "Construction pre and post analyses of air quality parameters (PM_{2.5}, PM₁₀ and CO) and noise levels: a case study of Lal Shahbaz Qalandar Underpass, Lahore, Pakistan", *Civil Engineering Infrastructures Journal*, 57(2), 357-370, <https://doi.org/10.22059/cej.2023.358714.1921>.
- Wang, Y., Lin, C., Zhao, B., Gong, B. and Liu, H. (2024). "Trajectory-based vehicle emission evaluation for signalized intersection using roadside LiDAR data", *Journal of Cleaner Production*, 440, 140971, <https://doi.org/10.1016/j.jclepro.2024.140971>.

Zhao, H.-X., He, R.-C. and Yin, N. (2021). “Modeling of vehicle CO₂ emissions and signal timing analysis at a signalized intersection considering fuel vehicles and electric vehicles”, *European Transport Research Review*, 13, 5, <https://doi.org/10.1186/s12544-020-00466-y>.

Zheng, T., Xiang, S., Zhang, S. and Wu, Y. (2023). “Variability of traffic-related air pollutants at two- and four-phase intersections”, *Atmospheric Pollution Research*, 14(12), 101936, <https://doi.org/10.1016/j.apr.2023.101936>.

Accepted / Not Edited