Research Article

Adarsh Yadav*, Manoranjan Parida, and Brind Kumar

Statistical modeling of traffic noise at intersections in a mid-sized city, India

https://doi.org/10.1515/noise-2022-0164
received December 04, 2022; accepted March 13, 2023

Abstract: The modeling of traffic noise is more debated around intersections due to traffic flow and road geometry complexity. The available intersection-specific traffic noise models cannot be transferred to predict the traffic noise at intersections in the mid-sized Indian cities due to traffic heterogeneity, variety in driving conditions, and vehicle compositions. This article aims to develop an intersection-specific traffic noise model by collecting data at 19 intersections in Kanpur, India. The data include a wide range of traffic, road, and weather-related variables. Furthermore, significant input variables are determined and used in the statistical regression model to develop an intersection-specific traffic noise model for the mid-sized Indian cities. This study develops a separate entrance and exit arm model based on the corresponding influencing variables. The coefficient of determination ($R^2$) value is 0.74 and 0.69 for the developed model at the entrance and exit arms, respectively, whereas these models achieve $R^2$ values of 0.73 and 0.67 in the validation step. Also, the performance of developed models is evaluated on the standard and mean absolute errors as performance metrics. This study finds that traffic volume and receiver distance are relatively the most important variables in the entrance and exit arm noise models.

Keywords: traffic noise, influencing variables, traffic heterogeneity, intersection-specific traffic noise model

1 Introduction

Due to urbanization, mid-sized cities are experiencing significant vehicular growth in recent years. Mid-sized Indian cities mainly carry heterogeneous traffic and poorly organized public transport systems. In the mid-sized cities, personalized two wheelers and shared three wheelers play a major role in daily commuting. Vehicular growth is severely impacting the environmental climate of mid-sized cities due to the lack of proper infrastructure and poor city planning. Traffic noise has emerged as a major threat to environmental noise pollution, causing severe health issues and reducing the quality of life of urban dwellings, particularly those living in the vicinity of roads. Recent growth in the transportation sector has increased traffic noise contribution by 70–80% to the overall urban noise pollution [1,2]. Existing literature shows that traffic noise exposure may cause sleep disorders with awakenings [3], learning and hearing impairment [4,5], myocardial infarction [6], high blood pressure [7], hypertension [8], annoyance [9], ischemic heart disease [10], cardiovascular diseases [11], etc. Over the last few years, many studies have been conducted on the importance and methods of controlling noise pollution caused by road traffic, both indoors and outdoors [12–16]. Consequently, assessing noise pollution is essential at the design stage and for existing roads [17,18]. Thus, an efficient model needs to be developed with higher prediction accuracy and wider applicability to solve the traffic noise problem in the mid-sized cities.

1.1 Concerning issues of the traffic noise model

Traffic noise models embody different influencing variables to predict traffic noise alongside the road network. Primarily, these variables are traffic volume, speed, honking, heavy vehicle percentage, vehicle composition, acceleration/deceleration, carriageway width, lane number, road gradient, road surface texture, road curvature, receiver distance
from the source, temperature, wind velocity, and land use [1,19–26]. Traffic volume, speed, and road geometry are considered major influencing variables of traffic noise and are incorporated in studies for developing traffic noise models. In addition, context-specific variables are also included to improve the prediction accuracy of the model. For instance, some studies have involved context-specific variables, i.e., the percentage of motorcycles [27] and honking [28] for modeling traffic noise. Consequently, it is essential to develop traffic noise models considering variabilities present at road sections to predict traffic noise with greater accuracy.

Moreover, different models are developed for different road structures. Same model cannot be applied to predict traffic noise at intersections and midblock due to changes in geometrical structure and traffic flow pattern. Intersections are the most conflict-prone and geometrically complex part of the road network system with relatively higher adjacent populations [29,30]. Also, the traffic flow pattern is complicated and heterogeneous that causes congestion, acceleration/deceleration, idling, honking, etc., mainly in the mid-sized cities. Therefore, intersections are significantly exposed and sensitive to traffic noise pollution. The retrospective literature has also indicated that intersections are the most chronic and prone to traffic noise pollution [31,32]. Conversely, traffic flow patterns and geometrical structures are comparatively uniform at mid-blocks. In light of this, it is imperative to conduct a comprehensive and systematic investigation for developing the intersection-specific traffic noise model for the mid-sized cities. Due to the complexity of road intersections, limited studies have focused on developing the traffic noise model of road intersections. Nevertheless, some researchers have made efforts to predict traffic noise in the vicinity of intersections across the globe using different modeling approaches [2,31,33,34].

1.2 Modeling approaches to traffic noise at the intersection

Due to various control measures and complexity, the traffic flow pattern is interrupted at the intersections [35,36]. Therefore, intersection-specific traffic noise models are developed considering the effect of these variations. Traffic noise models have been mainly developed by conventional, mathematical, microsimulation, and machine learning approaches. Conventional models, i.e., Federal highway administrative (FHWA-USA) [37], Calculation of Road Traffic Noise (CORTN-UK) [38], Common Noise aSSessment method (CNOSSOS-EU) [39], Acoustical Society of Japan-Road Traffic Noise (ASJ-RTN) [40], and German-specific Richtlinien für den Lärmschutz an Straßen (RLS-90) [41], are commonly used for modeling of traffic noise. FHWA, RLS-90, ASJ-RTN, and CNOSSOS models include correction factors for intersection noise modeling [18,39]. However, the CORTN model has not included correction factors for predicting traffic noise at intersections. Nonetheless, studies have modified the CORTN model to estimate traffic noise at intersections [2,42] accurately.

The mathematical approach is based on segment-by-segment analysis to encounter the noise emission of different flow regimes at road intersections [36,43]. The result of each segment is superimposed to estimate traffic noise at intersections. Moreover, simulation approach requires the kinematic characteristics of each vehicle (i.e., position, speed, and acceleration) present in the traffic stream [44,45]. Further, these characteristics are incorporated in microscopic traffic flow model, vehicle noise emission, and propagation model for modeling traffic noise at intersections [46,47]. Empirical approach uses statistical regression equations to develop the traffic noise model [20,48]. The model envisages the effect of influencing variables on traffic noise. In addition, recent studies have used machine learning approaches for modeling traffic noise [49–51].

As discussed earlier, various approaches are employed for modeling traffic noise at road intersections. Conventional models are typically created by developed countries, where traffic flow is more regulated and homogeneous than in developing countries. Significant modification is needed in conventional models for applying them in developing countries. Thus, conventional model transferability is complex and may lead to inaccurate prediction. The mathematical models employ different methods, such as inverse square law and integral Fourier transforms, for the estimation of traffic noise. The mathematical approach has been developed by previous studies [33,36,52,53]. Nevertheless, recent studies have not been carried out to update the mathematical approach for modeling traffic noise at intersections. Simulation approach requires detailed and microscopic information for modeling traffic noise. Gathering microscopic information in heterogeneous traffic conditions in the mid-sized cities can be difficult. Therefore, simulation approach is time-consuming and uneconomical for modeling traffic noise in the mid-sized cities in India. In addition, machine learning-based models are time-consuming and require larger data samples for greater prediction accuracy. Conversely, empirical models are presented through simpler statistical regression equations and require lesser information for developing the traffic noise model. Recent literature proves that empirical models efficiently predict traffic noise at intersections [31].
1.3 Motivation for the study

These intersection-specific traffic noise models are mainly built for economically developed cities. However, the mid-sized Indian cities embrace different traffic conditions than economically developed cities in terms of heterogeneity and mixedness, resulting in honking and varied driving styles [54]. Vikram-rickshaw (para-transit type) and bicycles are prominent in traffic streams of the mid-sized cities in India. Vikram-rickshaw is one of the loudest vehicles due to excessive engine noise [55]. Thus, there is a need to build an intersection-specific traffic noise model for the mid-sized cities using context-sensitive variables. Bearing this in mind, the current study aims to develop an intersection-specific traffic noise model based on the data collected in Kanpur, India. The Indian government has identified 100 cities, mainly mid-sized, with the mission to develop them as smart cities, making them people-friendly and sustainable under the National Smart Cities Mission. Therefore, the model is developed with the aim of easy interpretation and larger implementation to predict traffic noise at intersections in the mid-sized cities.

Further, the objective of study is to identify the relevant input parameters and conduct a statistical regression analysis to develop the intersection-specific noise model for the mid-sized cities. Furthermore, the developed model is validated to check the suitability for predicting traffic noise at intersections in the mid-sized cities with heterogeneous traffic flow conditions. This study also investigates the relative importance and sensitivity analysis to explore the ranking of input variables based on their importance in the model and variation in noise level with the change of input variables in the vicinity of intersections. Thus, this study contributes to traffic noise literature by investigating the major influencing variables of traffic noise at intersections in the mid-sized cities. Moreover, it also contributes by estimating how these variables may affect traffic noise levels at intersections with the variation in the influencing variables.

2 Methodology

2.1 Study area

Kanpur is in the list of 100 smart cities identified by the Government of India. With an area of 403.7 km$^2$ and a population of 2.92 million (Census, 2011), Kanpur is ranked as the largest agglomeration of Uttar Pradesh state and the 11$^{th}$ most populous city in India. As per the Ministry of Road Transport and Highways report, the total number of registered vehicles was 1.9 million in Kanpur [56]. Based on the reconnaissance survey, 19 intersection locations are selected from the different land use zones. These locations are finalized based on the following conditions: 1) the locations should be away from other known noise sources, 2) traffic movements are significant at selected locations, and 3) the selected road stretches are straight with a planar terrain and gradient that did not exceed 3%. Figure 1 presents the selected locations for data collection in Kanpur, India.

The pavements at the selected study locations are constructed of bituminous material with good road surface conditions. Moreover, traffic flow is mixed and highly heterogeneous at the study locations. Traffic composition comprises cars, two wheelers, auto-rickshaws, e-rickshaw, Vikram-rickshaw, buses, trucks, trailer-trailers, pickups, bicycle-rickshaw, bicycles, and horse-driven vehicles. It is also visualized that traffic flow is heavy at these locations and sometimes leads to high congestion, even on the sidewalks. In Figure 1, Chauraha or chowk represents a four-leg intersection, and Tiraha indicates a three-leg intersection.

2.2 Data collection

For this study, data collection involves measuring the traffic noise level and corresponding variables related to traffic parameters, road characteristics, and weather conditions at each selected location. Data are collected at one leg of the intersection within 150 m from the intersection stop line to capture the effect of intersection influence zone, which lies up to 250 m from the intersection stop line [57]. The leg is finalized based on the involvement of significant traffic, pedestrian movements, and other human activities, as it represents typical intersection characteristics.

Each leg of the intersection constitutes two arms, known as the entrance and exit arms. The arm through which vehicle approaches and leaves the intersection is defined as entrance and exit, respectively. Figure 2 presents one of the study locations to get an insight into the traffic flow characteristics at intersections in the mid-sized cities.

Figure 3 demonstrates the instrumental setup for measuring the data at the study locations. Hourly data acquisition is carried out separately at both arms on the finalized leg of the intersection. Equivalent noise level on
Figure 1: Selected study locations in Kanpur, India.

Figure 2: Study location.
A-weighted sound pressure level ($L_{Aeq}$) is measured at 1.2 m height above the ground level [48] near the building line at least 1 m from the façade [58] at selected intersections. $L_{Aeq}$ represents the equivalent amount of acoustic energy the source generates at a particular time [59]. This article presents the equivalent noise level on A-weighted sound pressure level as the equivalent traffic noise level. In addition, the corresponding variables, such as traffic volume, speeds, vehicular composition, honking, road geometry and distances, temperature, wind, relative humidity, and pavement characteristics, are measured to calibrate the model [20,26,28,31,35].

This study uses a 01 dB sound-level meter (SLM) to measure equivalent traffic noise levels alongside the entrance and exit arms separately. Video camera is installed for continuous recording of traffic volume and honking. Surveyors manually operate the Falcon radar gun to record the vehicular speed. Road geometry and distances, such as carriageway width, median width, distance of SLM from near- and farside carriageway, distance of SLM from building line, and distance of SLM from intersection stop line, are measured using a measuring wheel and measuring tape. The Schelt technology weather station is employed to continuously record weather parameters such as temperature, wind speed, relative humidity, and atmospheric pressure at study locations.

This study considers SLM distance in terms of geometric mean in the modeling to encounter the effect of near and farside traffic on the $L_{Aeq}$ at the entrance and exit arms. It is assumed that the geometric mean distance indicates the position of effective source from the SLM at the entrance and exit arms. The geometric mean distance of SLM is calculated as Eqs. (1) and (2) for the entrance and exit arms, respectively [35].

$$D_{g(ent)} = \sqrt{d_{A(ent)} \times d_{B(exit)}}$$  \hspace{1cm} (1)

$$D_{g(exit)} = \sqrt{d_{B(ent)} \times d_{A(exit)}}$$  \hspace{1cm} (2)

where $D_{g(ent)}$ and $D_{g(exit)}$ are the geometric mean distances of SLM at the entrance and exit arms, respectively; $d_{A(ent)}$ and $d_{A(exit)}$ are the distance of SLM A from the centerline of the entrance and exit arm carriageway; and $d_{B(ent)}$ and $d_{B(exit)}$ indicate the distance of SLM B from the entrance and exit arm carriageway, respectively.

Data collection is carried out in six phases from February 2020 to September 2021 to incorporate the effect of seasonal, festival, and traffic variations. In each phase, continuous 3-h data are recorded at all 19 study locations to include peak and off-peak hours variations. Henceforth, a total of 342-h data collection is carried out in this study to develop the model.

### 2.3 Exploratory data analysis

Table 1 presents the statistical description of all 342-measured data samples from different intersections in Kanpur.
The equivalent traffic noise levels are extracted from the entrance and exit arm SLMs. $L_{Aeq}$ value ranges from 63 to 83 dBA at the entrance and exit arms. The lowest $L_{Aeq}$ value is reported at LML Chauraha, while the highest is measured at Tat Mill Chauraha. Noise level has crossed acceptable limits prescribed by Central Pollution Control Board, New Delhi, at all locations except LML Chauraha. Traffic volume, honking, and vehicle composition are extracted manually by playing recorded videos. In order to extract honk event, a continuous 1-s honking length is treated as a single honk event.

Traffic volume has a wide variability in the range of about 200–5,400 veh/h, with a lower value at LML Chauraha and a higher value at Tat Mill Chauraha. The average stream speed at the exit arm is higher than at the entrance arm, with minimum and maximum values of 15.5 and 30.2 km/h, respectively. The maximum percentage of Vikram-rickshaws is reported at Tat Mill Chauraha, with a value of almost 20% in the traffic. Moreover, the percentage of heavy vehicles is maximum at LML Chauraha. In this study, heavy vehicle refers to buses, trucks, and tractor-trailers in the traffic stream. In this study, SLM is installed near the building line at both arms to analyze traffic noise levels experienced by people residing near intersections. The geometric mean distance of SLM is approximately in the range of 7–25 m, with a mean value of 13 m at both arms.

### 2.4 Methodological approach

This study employs multiple linear regression (MLR) analysis approaches to develop a traffic noise model for urban road intersections. The regression analysis approach is widely used for prediction due to its flexibility and simple applicability. The method develops the relationship between independent and dependent variables. The method is defined as univariate regression for one independent variable and multivariate regression for more than one independent variable. Traffic noise depends on many independent variables; therefore, the current study employs multivariate regression to develop a traffic noise model. The assumptions of MLR methods are normal distribution, linearity, freedom from extreme values, and having no multiple ties between independent variables [60,61]. The MLR equation is formulated as presented in the following equation:

$$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \ldots + \beta_pX_p,$$

where $Y$ is the dependent variable and $X_i$ indicates the independent variables. The term $\beta_0$ represents the constant value, and $\beta_i$ describes the regression coefficient in the above equations. Of the 342 data samples, 70% (239 samples) data are used for developing the model and 30% (103 samples) data are used to validate the developed model. Reliability and validity of constructed models are checked regarding the statistical significance of independent variables by $t$-test, coefficient of multiple determination ($R^2$), and $F$-test. The $R^2$ values lie in the range of 0 to 1, and higher $R^2$ value indicates that independent variables can predict the dependent variables with higher prediction accuracy.

### 3 Analysis and results

This section demonstrates the analysis of a wide range of variables in order to develop an effective regression traffic

<table>
<thead>
<tr>
<th>Variables</th>
<th>Symbol</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entrance arm equivalent traffic noise level (dBA)</td>
<td>$L_{Aeq(entr)}$</td>
<td>63.6</td>
<td>83.6</td>
<td>73.81</td>
<td>3.44</td>
</tr>
<tr>
<td>Exit arm equivalent traffic noise level (dBA)</td>
<td>$L_{Aeq(ext)}$</td>
<td>65.7</td>
<td>82.1</td>
<td>73.43</td>
<td>3.21</td>
</tr>
<tr>
<td>Traffic volume at the entrance arm (veh/h)</td>
<td>$Q_{ent}$</td>
<td>235</td>
<td>5,438</td>
<td>1928.21</td>
<td>946.76</td>
</tr>
<tr>
<td>Traffic volume at the exit arm (veh/h)</td>
<td>$Q_{ext}$</td>
<td>205</td>
<td>5,112</td>
<td>1995.25</td>
<td>929.06</td>
</tr>
<tr>
<td>Number of honk events at the entrance arm (honk events/h)</td>
<td>$H_{ent}$</td>
<td>71</td>
<td>1,479</td>
<td>551.94</td>
<td>271.73</td>
</tr>
<tr>
<td>Number of honk events at the exit arm (honk events/h)</td>
<td>$H_{ext}$</td>
<td>46</td>
<td>1,479</td>
<td>568.80</td>
<td>270.52</td>
</tr>
<tr>
<td>Average stream speed at the entrance arm (km/h)</td>
<td>$V_{ent}$</td>
<td>14.7</td>
<td>33.1</td>
<td>22.06</td>
<td>3.28</td>
</tr>
<tr>
<td>Average stream speed at the exit arm (km/h)</td>
<td>$V_{ext}$</td>
<td>15.5</td>
<td>30.2</td>
<td>23.20</td>
<td>3.46</td>
</tr>
<tr>
<td>% Heavy vehicle at the entrance arm (%Heavy vehicle/h)</td>
<td>$%HV_{ent}$</td>
<td>0</td>
<td>13.5</td>
<td>0.93</td>
<td>1.65</td>
</tr>
<tr>
<td>% Heavy vehicle at the exit arm (%Heavy vehicle/h)</td>
<td>$%HV_{ext}$</td>
<td>0</td>
<td>11.3</td>
<td>0.96</td>
<td>1.61</td>
</tr>
<tr>
<td>% Vikram-rickshaw at the entrance arm (%Vikram-rickshaw/h)</td>
<td>$%V_{ent}$</td>
<td>0</td>
<td>18.8</td>
<td>5.10</td>
<td>4.08</td>
</tr>
<tr>
<td>% Vikram-rickshaw at the exit arm (%Vikram-rickshaw/h)</td>
<td>$%V_{ext}$</td>
<td>0</td>
<td>22.5</td>
<td>5.03</td>
<td>4.20</td>
</tr>
<tr>
<td>Geometric mean distance of the entrance arm SLM (m)</td>
<td>$D_{gent}$</td>
<td>7.4</td>
<td>24.5</td>
<td>12.79</td>
<td>3.81</td>
</tr>
<tr>
<td>Geometric mean distance of the exit arm SLM (m)</td>
<td>$D_{gext}$</td>
<td>7.38</td>
<td>25.11</td>
<td>13.28</td>
<td>4.27</td>
</tr>
</tbody>
</table>
noise model for intersections. The current study attempts a stepwise MLR approach to investigate the best predictors for equivalent traffic noise.

3.1 Development of traffic noise prediction model

Traffic noise model is developed by establishing a relationship between equivalent traffic noise levels and influencing variables. The literature illustrates that traffic parameters are the key influencing variables affecting traffic noise. Traffic characteristics are not the same at the entrance and exit arms of urban intersections. Consequently, the current research adopts a separate-arm analysis to develop a separate traffic noise model for the entrance and exit arms. Previous studies also prove that the separate-arm analysis provides a better result than the single-arm model at intersections [35,48]. However, these researchers have developed acceleration and deceleration models for the exit and entrance arms of intersections, respectively, although the models have not incorporated acceleration and deceleration parameters [35,48]. Also, these models are mainly developed for metropolitan cities where traffic flow is more homogeneous and regulated. Henceforth, acceleration and deceleration models are not feasible for modeling traffic noise at intersections in Indian scenarios.

In line with developing the traffic noise model, data sample is checked to meet the assumptions of MLR analysis in this study. Further, Pearson’s correlation coefficient (r) is carried out at 95% confidence interval to investigate the variables substantially affecting $L_{Aeq}$ at the entrance and exit arms, i.e., $L_{Aeq}$(ent) and $L_{Aeq}$(ext). All variables with a substantial correlation ($r \geq 0.2$) are checked for their linear relationship with $L_{Aeq}$. Collinearity is also assessed among the influencing variables for developing entrance and exit arm models. Concerning this, traffic volume at the entrance arm ($Q_{ent}$) and traffic volume at the exit arm ($Q_{ext}$) are transformed into a logarithmic form to achieve a linear relationship and avoid collinearity among the variables. Previous studies also consider traffic volume in the logarithmic form for modeling traffic noise [32,62].

Model is developed employing significantly correlated variables, i.e., $\log(Q_{ent})$, $\log(Q_{ext})$, $V_{ext}$, %HV_{ent}, %HV_{ext}, %Vik_{ent}, D_{gent}$, and $D_{gext}$. The correlation coefficient values greater than 0.5 indicate a good correlation, those ranging from 0.3 to 0.5 indicate a moderate correlation, and those less than 0.3 indicates a low correlation. Variables insignificant at 95% CI and less correlated with $L_{Aeq}$ are removed from the analysis. Accordingly, honk events, average speed at entrance arm, temperature, different vehicle compositions, and other variables are eliminated from the analysis. Figure 4 presents the Pearson correlation coefficient of employed variables with $L_{Aeq}$(ent) and $L_{Aeq}$(ext). The figure depicts that traffic volumes correlate well with $L_{Aeq}$ while geometric distance of SLMs and %heavy vehicles are moderately correlated with $L_{Aeq}$. Moreover, $V_{ext}$ shows a low correlation with $L_{Aeq}$, and %Vik_{ent} presents a moderate correlation with $L_{Aeq}$(ent). Further, a stepwise regression analysis is performed in the IBM SPSS Statistics 27 to develop the traffic noise model. The best-fitted model is presented in Table 2 separately for the entrance and exit arms.
Additionally, the coefficient values are examined for their logical meaning in traffic engineering and traffic noise analysis. Statistical values of these variables and their significance for the model of both arms are presented in Table 3. Data analysis is performed at 95% confidence interval, indicating that p-values should be less than 0.05. Table 3 demonstrates that the p-value is less than 0.05 for all variables, showing that incorporated variables are significant at a 95% confidence interval.

In addition, the maximum value of variance inflation factor (VIF) is obtained close to 4 for log(Q_entr) and log(Q_ext) in both models. A VIF lesser than 10 indicates no multicollinearity issue among the variables [61]. The model fitness is assessed using the adjusted coefficient of determination ($R^2_{adj}$), standard error (SE), and mean absolute error (MAE). Despite some expected causes, the measured noise level is also affected due to extreme heterogeneity and many other unanticipated reasons, such as noise from pedestrians, shopkeepers, customers, and sometimes bus, and rickshaw driver who park their vehicles on the road and shout to pick up the passengers at intersections in the mid-sized cities, India. Additionally, some taxi, rickshaw, and bus drivers play music loudly in their vehicles, which affects the measured noise level at the intersections. Even with these unexpected reasons, the developed models are adequately fit to predict traffic noise at intersections. The fitness criteria for developed models are presented in Table 4.

The $R^2_{adj}$ value is estimated to be 0.74 and 0.69 for the entrance and exit arms, respectively, indicating the higher prediction accuracy of the developed models. As a result, the entrance arm model predicts traffic noise with an accuracy of 74%, while the exit arm model predicts traffic noise with 69% accuracy at the entrance and exit arms of intersections, respectively. Moreover, the calculated errors are within acceptable range, which is in line with the values obtained in other literature [31]. Analysis of variance (ANOVA) test is also conducted to check the variance among independent and dependent variables. Table 5 describes the ANOVA test results for the entrance and exit arm models. The table also elaborates that the p-value is less than 0.001 (<0.05) for both the entrance and exit

---

### Table 2: Regression coefficients ($\beta$) of the developed model

<table>
<thead>
<tr>
<th>Variables</th>
<th>Entrance arm model</th>
<th>Exit arm model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>44.73</td>
<td>47.95</td>
</tr>
<tr>
<td>Logarithmic of the entrance arm volume ($\log(Q_{ent})$)</td>
<td>6.29</td>
<td>2.84</td>
</tr>
<tr>
<td>Logarithmic of the exit arm volume ($\log(Q_{ext})$)</td>
<td>5.19</td>
<td>7.83</td>
</tr>
<tr>
<td>Stream speed of the exit arm ($V_{ext}$)</td>
<td>-0.15</td>
<td>-0.18</td>
</tr>
<tr>
<td>Percentage of heavy vehicles at the entrance arm (%HV_{ent})</td>
<td>0.17</td>
<td>-</td>
</tr>
<tr>
<td>Percentage of heavy vehicles at the exit arm (%HV_{ext})</td>
<td>-</td>
<td>0.25</td>
</tr>
<tr>
<td>Percentage of Vikram-rickshaw at the entrance arm (%Vik_{ent})</td>
<td>0.06</td>
<td>-</td>
</tr>
<tr>
<td>Geometric mean of distance of the entrance arm SLM ($D_{ent}$)</td>
<td>-0.4</td>
<td>-</td>
</tr>
<tr>
<td>Geometric mean of distance of the exit arm SLM ($D_{ext}$)</td>
<td>-</td>
<td>-0.39</td>
</tr>
</tbody>
</table>

---

### Table 3: Statistical values of independent variables involved in the models

<table>
<thead>
<tr>
<th>Model</th>
<th>Standard error</th>
<th>t-statistics</th>
<th>Significance</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entrance arm model</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>2.092</td>
<td>21.385</td>
<td>&lt;0.001</td>
<td>4.781</td>
</tr>
<tr>
<td>$\log(Q_{ent})$</td>
<td>0.909</td>
<td>5.717</td>
<td>&lt;0.001</td>
<td>4.139</td>
</tr>
<tr>
<td>$\log(Q_{ext})$</td>
<td>1.000</td>
<td>6.287</td>
<td>&lt;0.001</td>
<td>1.310</td>
</tr>
<tr>
<td>$V_{ext}$</td>
<td>0.033</td>
<td>-4.414</td>
<td>&lt;0.001</td>
<td>1.200</td>
</tr>
<tr>
<td>%HV_{ent}</td>
<td>0.078</td>
<td>2.117</td>
<td>0.035</td>
<td>1.679</td>
</tr>
<tr>
<td>%Vik_{ent}</td>
<td>0.028</td>
<td>1.996</td>
<td>0.047</td>
<td>1.129</td>
</tr>
<tr>
<td>$D_{ent}$</td>
<td>0.034</td>
<td>-11.536</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Exit arm model</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>2.234</td>
<td>21.464</td>
<td>&lt;0.001</td>
<td>4.188</td>
</tr>
<tr>
<td>$\log(Q_{ent})$</td>
<td>0.930</td>
<td>8.423</td>
<td>&lt;0.001</td>
<td>3.909</td>
</tr>
<tr>
<td>$\log(Q_{ext})$</td>
<td>0.986</td>
<td>2.878</td>
<td>0.004</td>
<td>1.086</td>
</tr>
<tr>
<td>$V_{ext}$</td>
<td>0.034</td>
<td>-5.329</td>
<td>&lt;0.001</td>
<td>1.666</td>
</tr>
<tr>
<td>%HV_{ext}</td>
<td>0.096</td>
<td>2.617</td>
<td>0.009</td>
<td>1.154</td>
</tr>
<tr>
<td>$D_{ext}$</td>
<td>0.030</td>
<td>-13.000</td>
<td>&lt;0.001</td>
<td></td>
</tr>
</tbody>
</table>
Table 4: Reliability test of the developed models

<table>
<thead>
<tr>
<th>Model</th>
<th>Multiple correlation coefficient ($R$)</th>
<th>Adjusted coefficient of determination ($R^2_{adj}$)</th>
<th>SE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entrance arm model</td>
<td>0.865</td>
<td>0.74</td>
<td>1.73</td>
<td>1.34</td>
</tr>
<tr>
<td>Exit arm model</td>
<td>0.836</td>
<td>0.69</td>
<td>1.83</td>
<td>1.42</td>
</tr>
</tbody>
</table>

Table 5: ANOVA test of the variable for the entrance and exit arm model

<table>
<thead>
<tr>
<th>Model</th>
<th>Sum of squares</th>
<th>Degree of freedom</th>
<th>Mean square</th>
<th>$F$</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entrance arm model</td>
<td>Regression</td>
<td>2082.642</td>
<td>6</td>
<td>347.107</td>
<td>115.274</td>
</tr>
<tr>
<td></td>
<td>Residual</td>
<td>698.583</td>
<td>232</td>
<td>3.011</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>2781.225</td>
<td>238</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exit arm model</td>
<td>Regression</td>
<td>1811.528</td>
<td>5</td>
<td>362.306</td>
<td>108.418</td>
</tr>
<tr>
<td></td>
<td>Residual</td>
<td>778.629</td>
<td>233</td>
<td>3.342</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>2590.157</td>
<td>238</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

arm models, which indicates that both models are significant at 95% confidence interval.

### 3.2 Model validation

The performance of models is tested based on $R^2$ and other error terms, as shown in Table 4. Nonetheless, model validation is carried out to examine the model fitness on the field data. In order to perform model validation, testing data set is used to assess the prediction capability of the model. The goodness of fit of model is determined by comparing the predicted $L_{Aeq}$ with measured $L_{Aeq}$ at the entrance and exit arms of the intersection. The variations of predicted $L_{Aeq}$ with measured $L_{Aeq}$ for the testing data samples are presented through scatter plots along the 45° line for the entrance and exit arm models, as illustrated in Figures 5 and 6. Based on scatter plots, a relationship is formed between measured and predicted equivalent traffic noise levels, as shown in Table 6.

As per the entrance arm model, the absolute difference between predicted $L_{Aeq}$ and measured $L_{Aeq}$ is less than 3 dBA for 92% of training and testing data samples. Additionally, the mean absolute difference between predicted $L_{Aeq}$ and measured $L_{Aeq}$ is 1.3 and 1.4 dBA for training and testing data samples, respectively. However, the absolute difference between predicted $L_{Aeq}$ and measured $L_{Aeq}$ for the exit arm model is less than 3 dBA for 92% of training and 83% of testing data samples. In addition, the mean absolute difference between predicted and measured values for the exit arm model is 1.4 and 1.5 dBA for training and testing data samples, respectively.

Low error values are obtained in the training and testing conditions for the entrance and exit arm models. These results demonstrate that developed models have a high degree of prediction accuracy. Thus, the developed entrance and exit arm models can be used to predict traffic noise at intersections in the mid-sized Indian cities. These models include a variety of input variables to predict traffic noise levels at intersections. The next subsection discusses the relative importance of input variables involved in the models.

![Figure 5: Scatter plot between measured $L_{Aeq}$ and predicted $L_{Aeq}$ for the entrance arm.](image-url)
3.3 Relative importance

In regression analysis, the relative importance of input variables is described as a contribution to variability in $R^2$. According to $R^2$-based metric analysis, the input variables are ranked according to their effectiveness [21,63,64]. Thus, this study examines the relative importance of input variables by measuring the change in $R^2$ values after removing the variable. The previous step illustrates how the model reacts after removing input variables. Moreover, the technique evaluates the change in the error value when one input variable is removed from the model. The entrance and exit arm models have six and five input variables, respectively. The input variables are ranked separately for the entrance and exit arm models, as depicted in Figures 7 and 8.

As per Figure 7, the $R^2$ value reduces by 19% on removing $Q_{ent}$ and $D_{gent}$ from the entrance arm model. Thus, $Q_{ent}$ and $D_{gent}$ are ranked first in terms of their importance among the input variables in the entrance arm model. Moreover, the $R^2$ value reduces by 10% and 3% by removing $Q_{ext}$ and $V_{ext}$, respectively. Thus, they have ranked the second and third important variables in the entrance arm model. The lowest important variables are $\%HV_{ent}$ and $\%V_{ikent}$, as almost 1% variation is measured in $R^2$ by removing them from the entrance arm model. Consequently, $\%HV_{ent}$ and $\%V_{ikent}$ are ranked fourth and fifth in terms of their importance in the entrance arm model.

Figure 8 depicts that variation in $R^2$ is equal to 36% and 25% on removing $Q_{ext}$ and $D_{gext}$, respectively, from the exit arm model. Thus, $Q_{ext}$ and $D_{gext}$ are ranked first and second important input variables in the exit arm model. Moreover, $V_{ext}$ and $Q_{ent}$ cause 4% and 3.7% reductions in $R^2$ value, ranked third and fourth important input variables in the exit arm model. The 1% reduction in $R^2$ value is measured by removing $\%HV_{ext}$ variable from the

---

**Table 6: Relationship between measured and predicted $L_{Aeq}$ for the entrance and exit arm models**

<table>
<thead>
<tr>
<th>Model</th>
<th>Equation</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entrance arm</td>
<td>$L_{Aeq}(\text{measured}) = 0.98 \times L_{Aeq} + 1.44$</td>
<td>0.73</td>
</tr>
<tr>
<td>Exit arm model</td>
<td>$L_{Aeq}(\text{predicted}) = 1.05 \times L_{Aeq} - 3.56$</td>
<td>0.67</td>
</tr>
</tbody>
</table>
exit arm model. Consequently, the percentage of heavy vehicles at the exit arm is the least important input variable in the exit arm model.

Relative importance analysis indicates that traffic volume and distance of SLM substantially affect the equivalent traffic noise level in the vicinity of intersections. Therefore, sensitivity analysis is carried out to measure the variation of these variables on traffic noise in subsequent subsections.

### 3.4 Sensitivity analysis

Sensitivity analysis is carried out to investigate the $L_{Aeq}$ variation in different scenarios. In this case, variation of traffic noise is observed based on developed models for the following scenarios: 1) 10%, 25%, 50%, and double growth in traffic volume at the entrance and exit arms, separately and together, and 2) increasing distance of SLM from the effective source twice and four times. Sensitivity analysis provides insight into how traffic noise varies for each scenario by keeping the other scenarios constant. The result of sensitivity analysis is presented in Figures 9 and 10 for the entrance and exit arm models, respectively.

The result illustrates that $L_{Aeq(\text{ent})}$ reduces 5.1 dBA by doubling $D_{\text{gent}}$. However, $L_{Aeq(\text{ent})}$ decreases significantly, almost by 15.3 dBA, on increasing the distance between SLM and effective source four times. Likewise, $L_{Aeq(\text{ext})}$ reduces by 4.7 dBA and 15.0 dBA on increasing the $D_{\text{gext}}$ value twice and four times, respectively. A larger and almost similar reduction is observed in $L_{Aeq}$ at the entrance and exit arms by increasing the distance of SLM from the effective source.
Moreover, traffic volume growth at individual arms causes less change in traffic noise. For instance, a 10%, 25%, 50%, and double increment in $Q_{\text{ent}}$ with no change in $Q_{\text{ext}}$ cause a growth of 0.3, 0.6, 1.1, and 1.9 dBA in $L_{\text{Aeq(\text{ent})}}$, and of 0.1, 0.3, 0.5, and 0.9 dBA in $L_{\text{Aeq(\text{ext})}}$. A similar growth in $Q_{\text{ext}}$, with no change in $Q_{\text{ent}}$, raises the $L_{\text{Aeq(\text{ent})}}$ by 0.2, 0.5, 0.9, and 1.6 dBA and $L_{\text{Aeq(\text{ext})}}$ by 0.3, 0.8, 1.4, and 2.4 dBA, although concurrent growth in traffic volume at both arms substantially affects $L_{\text{Aeq}}$, compared to growth in traffic volume at a single arm. The concurrent growth in $Q_{\text{ent}}$ and $Q_{\text{ext}}$ by 10%, 25%, 50%, and double causes 0.5, 1.1, 2.0, and 3.5 dBA rise in $L_{\text{Aeq(\text{ent})}}$, and 0.4, 1.0, 1.9, and 3.2 dBA rise in $L_{\text{Aeq(\text{ext})}}$.

4 Discussion

The study uses 342 h of data samples collected in Kanpur to develop the intersection-specific traffic noise model. This study discusses how these variables affect traffic noise at intersections.

4.1 Effect of independent variables on the equivalent traffic noise level

This study finds that speed and distance of effective source from SLM negatively affect the $L_{\text{Aeq}}$ at both arms of the intersection. The speed effect is explained by the fact that other vehicular noise sources have a minimal impact on $L_{\text{Aeq}}$ except vehicular engine noise at lower speeds. Therefore, if vehicles move at lower speeds, engine noise contributes longer durations to overall noise and causes a growth in $L_{\text{Aeq}}$. Previous literature also shows that speed is negatively correlated with traffic noise at lower speeds [63]. Moreover, the negative effect of distance is explained as the intensity of sound energy diminishes as the distance increases due to absorption, reflection, and other attenuation. Consequently, a lesser noise level is measured by increasing $D_{\text{gent}}$ and $D_{\text{gext}}$. This result concords with Pamanikabud and Tharasawatpipat (1999) findings that noise level decreases with the increase in distance of effective source from SLM.

Surprisingly, $%HV_{\text{ent}}$, $%HV_{\text{ext}}$, and $%V\text{i}_{\text{ent}}$ have lower regression coefficient values, which indicate a lesser impact of vehicle composition on the $L_{\text{Aeq}}$. Relative importance and sensitivity analysis results also identify them as lower ranked input variables. It is believed that high traffic volume overshadows the effect of vehicular composition on $L_{\text{Aeq}}$. Nonetheless, the effect of percentage of heavy vehicles and Vikram-rickshaw is significant on $L_{\text{Aeq}}$ due to their higher engine noise. Additionally, tire pavement interaction noise is also significant for heavy vehicles. However, only Vikram-rickshaw at the entrance arm is considered significant for developing the entrance arm model. It can be attributed to the fact that Vikram-rickshaw...
halts for a longer duration at the entrance arm due to traffic flow patterns and to deboard passengers. Henceforth, % Vikent significantly affects $L_{A_{eq}}(ent)$ and is involved in developing the entrance arm model. Moreover, Vikram-rickshaw rarely halts and deboards the passengers at the exit arm since the traffic flow pattern prohibits their stopping, thereby contributing less to $L_{A_{eq}}$.

Nevertheless, both models indicate that traffic volume substantially affects traffic noise. The $L_{A_{eq}}(ent)$ and $L_{A_{eq}}$ (ext) are largely and positively affected by $Q_{ent}$ and $Q_{ext}$ at intersections with a higher impact of nearside arm traffic volume. The interrupted flow, heterogeneity, and mixed vehicle composition affect the traffic flow significantly, which further causes extra emission of vehicular source noise at intersections. This study finding is supported by the previous literature [35,48]. However, Abo-Qudais and Alihary (2005) investigated that farside traffic volume has a negligible impact on the equivalent traffic noise level at the intersection. But this study measures the effect of farside vehicles on the $L_{A_{eq}}$ by running a single vehicle in uninterrupted conditions at the farside arm, thus ignoring the effect of traffic flow conditions at the intersection.

4.2 Analysis of unexpected causes of traffic noise level at the intersection

According to the study, higher noise levels are sometimes measured even at lower traffic volumes and greater SLM distances. This unexpected fluctuation in traffic noise is revealed through a careful data observation. Different cases are identified for this dramatic variation in the measured equivalent traffic noise level at both arms of the intersections. Case 1 indicates that the absence of a median causes greater vehicular interaction in the traffic stream, which leads to an increase in heterogeneity, congestion, and honking. In Case 2, the commercial shops and other infrastructure buildings are in clusters and often built near intersections in the midsized cities of India. As a result, traffic loads increase on the roads, causing platoon formation and unnecessary honking. In Case 3, the drivers generally expect to drive faster at the exit arm of the intersection, but this may not have occurred due to unexpected obstacles. For example, pedestrians not following the cross rules or absence of footpaths, unauthorized parking, shop owners’ encroachments on the roads, and sometimes drivers of a local intermediate public transport (e.g., Vikram-rickshaw, minibuses etc.) stop their vehicles on the exit arms to board passengers. These interruptions cause the narrowness of road stretches, restricting continuous movement, and regulation of traffic flow at exit arms. In addition, pavement is deteriorated at some locations, resulting in potholes and cracks on the road surface. This causes hindrance, extreme start-stop conditions, honking, and sometimes queue formation at the exit arms. Eventually, drivers frequently accelerate and decelerate at lower gears on the exit arms, increasing engine labor and generating higher vehicular noise at similar speeds [19,31]. All these are the possible reasons for unprecedented growth in the noise level at the intersections in the mid-sized cities.

The noise scenarios are different in metropolitan cities than in the mid-sized cities due to a variability in traffic flow and other disturbances. Traffic flow is more regulated, and drivers tend to follow the rules more cautiously in metropolitan cities than in the mid-sized cities. Therefore, traffic flow is more homogenous and uniform in metropolitan cities. Also, vehicle maintenance is more regular in metropolitan cities than in the mid-sized cities. In addition, pedestrians, urban dwellers, and shopkeepers cause lesser interference on the roads in metropolitan cities and higher disturbances in the mid-sized cities. Therefore, traffic noise is more chaotic in the mid-sized cities due to higher hindrances and disturbances.

Overall, traffic noise is affected by various expected and unexpected causes at intersections in the mid-sized cities in countries like India. Traffic volume at the exit arms ranked as the most important input as it is disturbed by several expected and unexpected causes. These causes are carefully examined at the site and by playing the recorded videos.

5 Conclusion

This study develops an intersection-specific traffic noise model for the mid-sized cities based on the data collected in Kanpur, India. A wide range of influencing variables is collected from the 19 intersection locations to develop the traffic noise model. Further, the result identifies that traffic volume, speed, receiver position, percentage of heavy vehicles, and Vikram-rickshaw are the most significant variables affecting the equivalent traffic noise level. These findings answer the research question of the major influencing variables affecting traffic noise at intersections. The identified variables are inserted into the statistical regression model for developing the intersection-specific traffic noise model. This study opted a separate arm analysis approach for developing the model due to differences in traffic flow characteristics and control measures at the entrance and exit arms. In addition,
the relative importance and sensitivity analysis is carried out to check the importance of input variables and noise variation due to fluctuation in variables. This experimental study indicates that traffic volume and position of receivers are the most dominant influencing variables, while vehicle composition is the least influencing variable affecting the traffic noise at intersections. This study also reveals that traffic volume and vehicle composition positively affect traffic noise; conversely, the speed and distance of the effective source from the SLM negatively influence the traffic noise. Henceforth, these outcomes answer the research questions on how these variables affect the traffic noise level at the intersections.

The developed model is easy to interpret and can assist the agencies to predict the traffic noise experienced by urban dwellers in the vicinity of intersections in the mid-sized cities. Also, this study can be helpful for urban planners and authorities to mitigate and control traffic noise at intersections. Instead of the above-discussed variables, traffic noise is also influenced by honking, road surface conditions, queue length, etc. Unfortunately, these variables have not been employed in our study. Future studies may include these variables for developing intersection-specific traffic noise models. Besides this, the current study’s findings are also limited to the mid-sized urban cities.

**Acknowledgments:** For the present study, data collection is carried out under the IMPRINT project. The Ministry of Housing and Urban Affairs and the Ministry of Education support the IMPRINT project. The authors are also thankful to the Ministry of Education for awarding the PhD research assistantship to Mr. Adarsh Yadav.

**Funding information:** The author states no funding involved.

**Author contributions:** All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

**Conflict of interest:** The authors state no conflict of interest.

**References**


[55] Yadav A, Parida M, Kumar B. Determination of passenger car noise equivalent for mid-sized cities in India. INTER-NOISE and


