International Journal of Mathematical, Engineering and Management Sciences
Vol. 9, No. 3, 411-434, 2024
https://doi.org/10.33889/IJMEMS.2024.9.3.022

Ram Arti Publishers

Traffic Noise Modeling under Mixed Traffic Condition in Mid-Sized Indian City: A Linear Regression and Neural Network-Based Approach

Rohit Patel
Department of Environmental Science and Engineering,
Indian Institute of Technology (Indian School of Mines), Dhanbad, 826004, Jharkhand, India.
E-mail: rohitpatelp7@gmail.com

Prasoon Kumar Singh
Department of Environmental Science and Engineering,
Indian Institute of Technology (Indian School of Mines), Dhanbad, 826004, Jharkhand, India.
Corresponding author: pks0506@iitism.ac.in

Shivam Saw
Department of Environmental Science and Engineering,
Indian Institute of Technology (Indian School of Mines), Dhanbad, 826004, Jharkhand, India.
E-mail: shivamsaw@gmail.com

(Received on November 5, 2023; Revised on January 9, 2024 & February 12, 2024; Accepted on February 20, 2024)

Abstract
Noise pollution is a significant concern in urban settings, caused by traffic increases, urban expansion, and industrial activity. The transportation sector is a crucial contributor to overall noise pollution, particularly in India, where different vehicles ply the roads, resulting in highly fluctuating noise levels. Consequently, traffic noise modeling is essential for addressing this severe issue. The present study employs the MLR (Multiple Linear Regression) and Artificial Neural Network (ANN) approach to model and predict traffic-induced noise levels. The ANN approach outperforms the MLR technique. The architecture of the ANN model integrates different vehicle categories and average speeds as input, resulting in precise predictions. Evaluation of the model's performance reveals an average Root Mean Squared Error (RMSE) of 0.204 and a high Coefficient of determination (R^2) value of 0.93, emphasizing its accuracy. Similarly, in the case of MLR model the RMSE for the training and testing dataset are 1.55 and 1.69 dBA, respectively with R^2 value of 0.84. Subsequently, sensitivity analysis highlights the substantial impact of 2-wheelers, tractors/trailers, and 3-wheelers on noise predictions. This study contributes valuable insights into noise management, urban planning, and sustainable development. It demonstrates the efficiency of the ANN approach in addressing complex noise pollution challenges, offering a path toward quieter and healthier urban environments.

Keywords- Road traffic, ANN, Traffic noise prediction, Noise pollution, Multiple linear regression.

1. Introduction

Noise pollution is an emerging concern in urban areas, primarily due to the consistent growth in road traffic, urban spread, industrial activities, and infrastructure expansion in and around cities. In urban settings, the combined impact of the industrial, commercial, and transportation sectors accounts for 75% of total noise pollution (Mishra et al., 2021). Within the transportation domain, road traffic noise accounts for the majority, accounting for nearly 70% of the total, as depicted in Figure 1. Noise pollution can have several adverse consequences for both human health and the environment. Long-term noise exposure can be a reason for various health issues, including stress, annoyance, sleep disruptions, hearing loss, and an increased risk of cardiovascular disease (Babisch & Van Kamp, 2009; Gupta et al., 2018; Lipowicz & Lopuszanska, 2005; Tiwari et al., 2023; Yankoty et al., 2021). Apart from these issues, it can also impair communication and concentration, resulting in lower productivity and quality of life (Münzel & Daiber, 2018). Moreover, noise pollution adds to overall environmental degradation and can have long-term adverse effects on human physical and mental well-being.

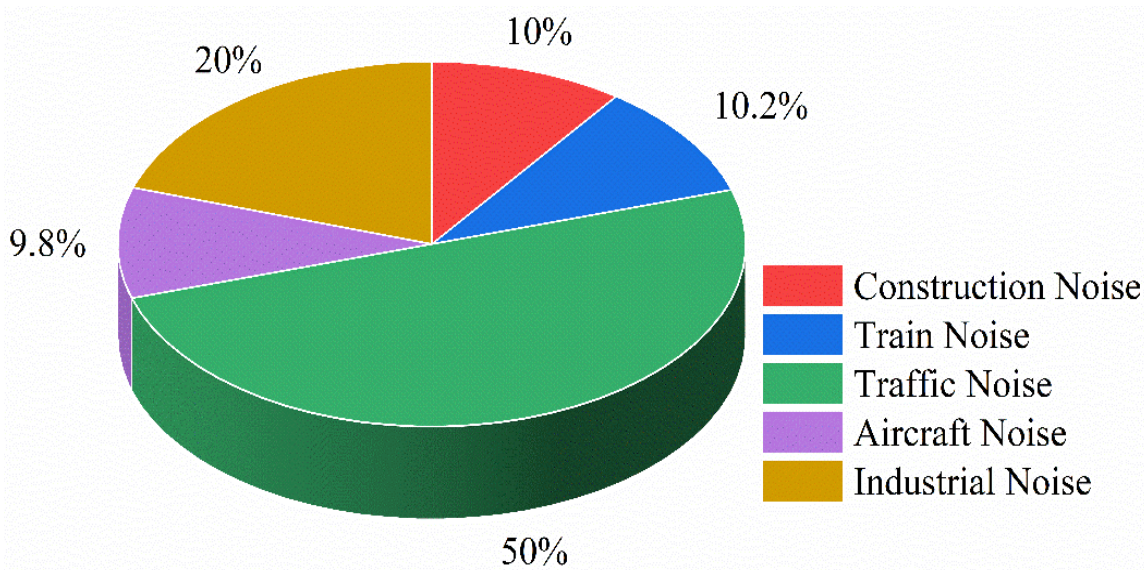

Figure 1. The distribution of noise pollution sources in urban setting.

Among all the sources of noise pollution, the transportation sector is the primary contributing source (Ahac et al., 2021). Transportation systems are crucial in progressive societies, providing the necessary infrastructure for convenient and efficient movement to meet societal demands (Ahmed et al., 2021). Traffic noise is a worldwide issue that affects both urban and rural locations. Traffic noise has become a primary environmental concern as cities have grown and vehicles on the road have increased (Nedic et al., 2014). The widespread usage of internal combustion engine vehicles for transportation is one of the significant causes of traffic noise worldwide. The sheer number of vehicles on the road, ranging from cars to buses and trucks, contributes to an endless flow of noise (Pascale et al., 2021). Additionally, variables including acceleration and braking produce loud, sudden sounds. Road surfaces and tire noise add to the total noise generation, particularly on highways. Congestion in urban locations magnifies traffic noise as vehicles idle, hoot, and jockey for a place (Ahac et al., 2023; Chevallier et al., 2009). Moreover, when all these causes are combined, the widespread and rising issue of traffic noise impacts communities globally, necessitating comprehensive noise reduction measures and environmentally friendly transportation solutions.

In the Indian context, the problem of noise pollution originating from road traffic is a significant and severe concern in medium-sized cities across India. Traffic noise primarily results from excessive horn usage, the contact between the vehicle’s tires and road surfaces, and the noise generated by the engine, among other factors (Konbattulwar et al., 2016). The continuous honking of vehicle horns on the road network is primarily caused by inadequate traffic management, a large number of vehicles, a lack of driver discipline, and heavy congestion during commutes (Thakre et al., 2020). Indian roads permit a wide range of vehicle categories, which include 2-wheelers, 3-wheelers, 4-wheelers, and larger vehicles such as trucks, lorries, trailers, and tractors. The presence of these vehicles on the road, each with unique speeds and characteristics, makes enforcing lane discipline extremely difficult. These components regularly influence noise characteristics in the Indian traffic landscape. It is characterized by its variability due to varied vehicle types, road dimensions, tire-pavement interactions, and operational aspects like vehicle speed and driver behavior (Vijay et al., 2018).

Traffic noise modeling plays an essential role in noise pollution management. It offers a systematic and scientific approach to investigating the sources, patterns, and amounts of traffic-generated noise. Moreover, it also assists in identifying noisy locations and the specific variables contributing to noise pollution by precisely modeling noise emissions from various types of vehicles, road designs, and traffic situations. Therefore, traffic noise modeling is crucial for implementing noise reduction measures and promoting noise-conscious development in urban areas to reduce the negative impact on the community's health and quality of life. There are several conventional traffic noise models which are widely used across the globe. Among these models, the United Kingdom developed a model called Calculation of Road Traffic Noise (CRTN), which was later implemented in other countries, including India (Manojkumar et al., 2019; Peng et al., 2019). The United States established its traffic noise prediction model, Federal Highway Administration (FHWA), validated in various other countries (Golmohammadi et al., 2009). Other countries like Germany followed the RLS-90 model, Switzerland developed the STL-86 model, Japan has the ASJ-1993 model, Thailand established the ERTC approach, France and Belgium developed the MITHRA model, and the most recent and advanced CNOSSOS-EU model was developed for European countries (Patel et al., 2022). Though different predictive noise models are employed, the contribution factors considered in such models differ based on the specific weather conditions, traffic volume, types of vehicles, vehicle speeds, road geometry, and specific traffic conditions (Hamad et al., 2017).

Furthermore, these prediction models have several limitations, including inability to capture non-linear relationships, dependency on fixed parameters, and limited adaptability (Tiwari et al., 2022; Zhang et al., 2021). Therefore, it resulted in the utilization of various AI-driven models in traffic noise modeling. A study showed that compared to empirical and statistical models, an ANN provided superior and more accurate predictions of traffic noise (Kumar et al., 2011). A study comparing ANN with seven acoustic models revealed that the mean error of the acoustic models was ten times greater than that of the ANN (Genaro et al., 2010). Another study compared the effectiveness of an ANN to six acoustic models and found that the ANN outperformed the other acoustic models (Nedic et al., 2014). Apart from conventional prediction models, many advanced techniques have evolved in the field of traffic noise modeling. Numerous researchers have also created noise prediction models employing the concept of the machine learning approach (Ahmed et al., 2021; Kumar et al., 2014; Singh et al., 2016; Umar et al., 2022), emotional ANN method (Nourani et al., 2020), the idea of time series like ARIMA (Garg et al., 2017; Ma et al., 2020), various hybrid deep learning methods such as CNN and auto-encoder based LSTM (Chen et al., 2021; Tiwari et al., 2022; Umar et al., 2023), regression model (Chang et al., 2019; Yadav et al., 2023), Bayesian approaches (Osborne et al., 2020; Zhu et al., 2019), and the graph theoretic methods (Gilani & Mir, 2021). These models are tailored to specific road conditions in different countries and establish relationships between variables, yielding reasonably satisfactory results.

Due to their distinctive features and advantages, Multiple Linear Regression and Artificial Neural Networks have been chosen as modeling techniques for handling the complexity of traffic-induced noise. Because of its ease of use and accessibility in describing linear correlations between noise levels and many contributing factors such as traffic volume, speed, and time of day, multiple linear regression is the preferred method. This approach works incredibly well in scenarios where effective stakeholder communication depends on having a clear and understandable grasp of how individual variables affect traffic noise. Conversely, Artificial Neural Networks are used to address the complex and nonlinear patterns present in noise caused by traffic. Artificial neural networks (ANNs) are a good choice in situations where the noise production process is dynamic and may not follow linear assumptions because of their flexibility and ability to capture complicated correlations. The selection of MLR and ANN is based on the particular modeling objectives, the characteristics of the data, and the balance between the degree of complexity and simplicity needed to capture the complicated interactions between traffic factors and noise levels. ANNs find widespread application across diverse domains, including finance, medicine, engineering, and science. This advantage becomes particularly evident when traditional techniques involving correlations and group differences prove inadequate due to the complexity of the underlying relationships (Givargis & Karimi, 2010). Several previous studies were conducted in different traffic environments, revealing the ANN approach's effectiveness and robustness over the regression-based approach (Garg et al., 2015; Steinbach & Altinsoy, 2019). The culmination of previous research provides strong evidence supporting the effectiveness of the ANN approach for traffic noise prediction.

1.1 Motivation for the Present Study

The majority of noise prediction models are designed for economically developed cities, with a primary focus on predicting noise levels at specific intersections. There is limited attention given to noise prediction models for the mid-block sections of roadways. Nevertheless, in mid-sized Indian cities, there is a noticeable variation in traffic flow conditions, with a diverse mix of vehicle categories plying on the roads. Notably, Bikes, Cars, and three-wheelers (diesel-driven and e-rickshaws) are the predominant vehicle category in these Indian mid-sized cities. As a result, there is a pressing need for the development of a novel and robust model that considers the context-specific factors influencing traffic noise generation on mid-block road segments. The objective of this study is to create a traffic noise prediction model using data collected from Dhanbad City, India.

With each passing day, there is a steady and significant increase in vehicular traffic on Indian roadways. Figure 2 depicts the trend of vehicular growth found in Dhanbad city, India, which is the study area in the current research. Additionally, from Figure 2, it is evident that between 2003 and 2020, the number of vehicles has exceeded threefold. Therefore, considering vehicle growth and awareness of the detrimental impact of noise pollution, it is crucial to develop a robust prediction model for accurate forecasting.

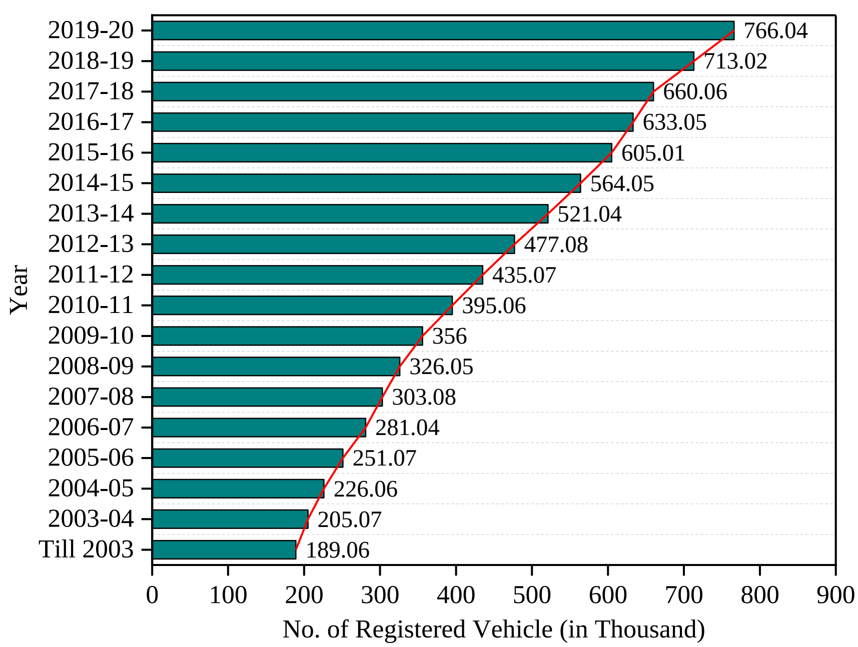

Figure 2. Trend of vehicle growth in Dhanbad city (Road Transport Year Book, 2019).

In contrast to modeling studies in other pollution-related fields like water and air pollution, there has been a notably limited amount of research conducted on prediction of noise pollution in the Indian context. While already established methods serve similar purposes, it is essential to stress the ongoing need to examine and model roadside noise in various areas due to its dependence on local characteristics. Further investigation is still crucial, particularly in developing countries, to comprehensively understand and model roadside noise.

Therefore, the present research aims to employ a Multiple Linear Regression and Multilayer Artificial Neural Network approach to model traffic noise in Indian traffic situations. The study's findings have the potential to contribute to noise management strategies, urban planning, and sustainable development not just in Dhanbad but also in similar urban contexts around the world.

2. Study Area

Dhanbad, also called the "Coal capital of India," is located in the eastern part of India, with geographic coordinates ranging from 23.7957°N latitude to 86.4304°E longitude. It is an essential urban center in the region, distinguished by its complex transportation network, diverse land uses, and high population density. The city's layout comprises residential, commercial, industrial, and recreational zones, resulting in complex traffic flow and noise dynamics. Therefore, it is necessary to investigate the complicated nature of traffic-induced noise pollution. A total of 22 monitoring sites were selected based on diverse land use patterns, the presence of intersections, and homogeneous coverage of the whole study area. All the selected sampling sites lie adjacent to the major highways (NH-19, NH-18, and NH-419) passing across Dhanbad city, as shown in Figure 3.

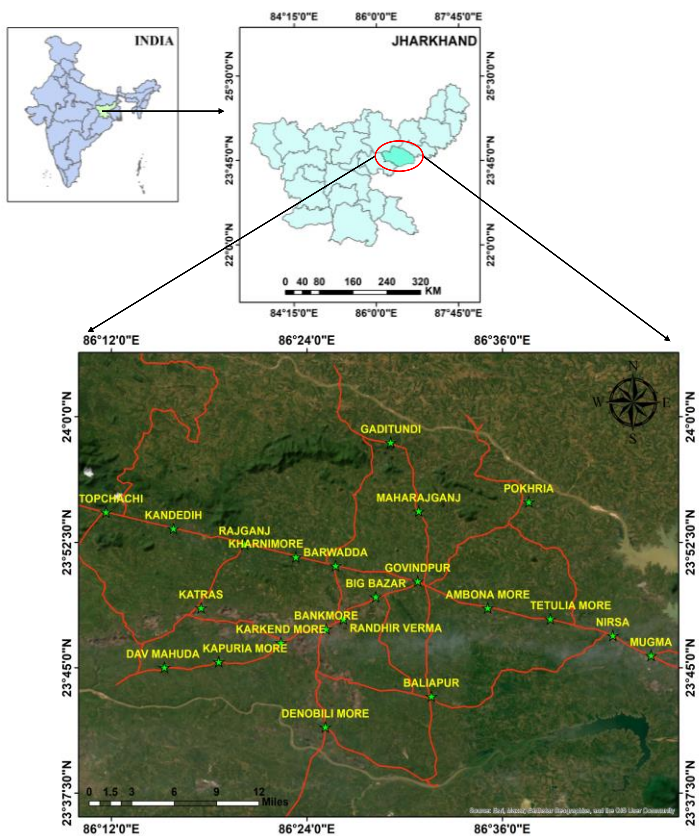

Figure 3. Study area shows all the monitoring sites.

3. Methodology

3.1 Noise Monitoring and Data Collection

In the current research, continuous monitoring of A-weighted noise levels (L_{eq}) was conducted in 1/3 octave band with Bruel & Kjaer 2238 mediator sound level meter (SLM) from July to September 2022. L_{eq} signifies the energy-equivalent representation of noise emissions from a source within a predefined temporal interval. One B&K 2238 SLM, categorized as a Class 1 SLM, was utilized and calibrated prior to its deployment. Additionally, a windshield was employed to cover the microphone during the entire monitoring process. Noise monitoring was conducted for 5-time slots, each of 3 hours duration in the morning, morning peak hours, afternoon, evening, and night time. As a result, 15 datasets were collected from each sampling location, leading to 330 datasets gathered from all the monitoring locations. At each monitoring site, the SLM was positioned a minimum of 1 meter away from any reflectors, excluding those placed on the ground (Thakre et al., 2020). The SLM was set at a height of 1.2 meters above the ground and 7.5 meters away from the center of the nearest lane, as shown in Figure 4 (Chen et al., 2021; Debnath & Singh, 2018).

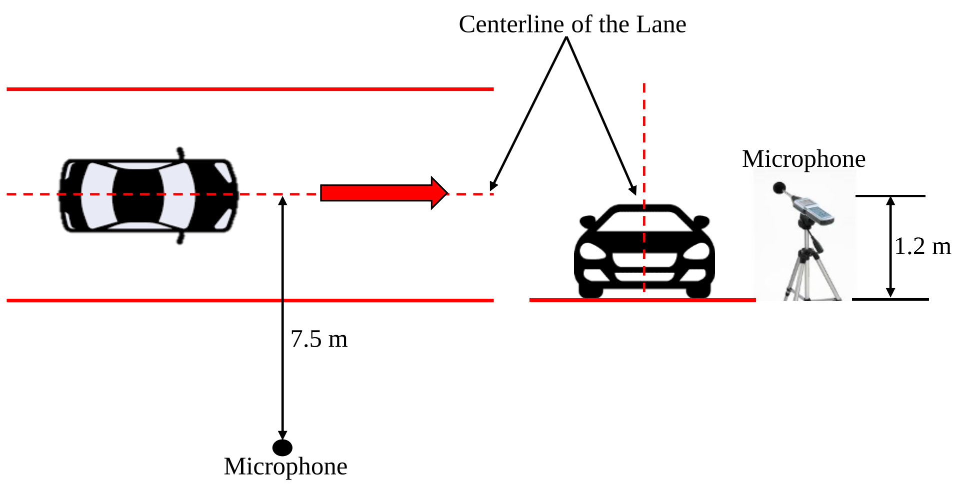

Figure 4. Schematic of the monitoring setup in the present study.

A description of all sampling locations with their geographical coordinates and types of land use is shown in Table 1. During the data collection, the temperature, relative humidity, and wind speed at all the monitoring sites were below 35^{0}C, 80%, and 5 m/s, respectively. Additionally, the sampled locations had bituminous pavement with a completely dry and smooth surface texture with a generally level terrain and gradient of about 3%. A lack of homogeneity characterizes Dhanbad's road infrastructure due to uncontrolled expansion and variation in the traffic mix.

Furthermore, the study also examined several relevant factors exerting influence on traffic noise generation. These factors include the composition of vehicular traffic, traffic volume, and vehicular speed at distinct locations. A recording camera was used to monitor vehicle movement while determining the hourly traffic volume at selected sampling sites. At the same time, individual vehicle speed was measured using a Radar Gun. Subsequently, the mean velocity for the whole traffic stream was calculated by averaging the individual speeds of all identifiable vehicle classifications. The classification includes the categories of the vehicles prominently observed on the roads of mid-sized Indian cities. Seven different classes of vehicles are categorized in the present study, which includes 2-wheelers, 3-wheelers, cars, light commercial vehicles (LCV), buses, trucks, and tractors/trailers.

Table 1. Description of sampling locations selected in the present study.

S. No.	Location	Latitude	Longitude	Land Use
1.	Topchachi	23.8096	86.5851	Commercial area
2.	Kandedih	23.7986	86.6491	Residential area
3.	Kharni More	23.7821	86.7132	Residential area
4.	Ambona More	23.7621	86.7524	Residential area
5.	Rajganj	23.8516	86.4293	Commercial area
6.	Barwadda	23.8603	86.3886	Commercial area
7.	Govindpur	23.8730	86.3362	Commercial area
8.	Mugma	23.8888	86.2632	Industrial area
9.	Nirsa	23.9051	86.1942	Commercial area
10.	Tetulia More	23.8361	86.5135	Residential area
11.	Big Bazar	23.8208	86.4704	Commercial area
12.	Randhir Verma Chowk	23.7983	86.4373	Silence area
13.	Bank More	23.7886	86.4196	Commercial area
14.	Kapuria More	23.7751	86.3734	Commercial area
15.	Karkend More	23.7504	86.2542	Commercial area
16.	DAV Mahuda	23.7558	86.3097	Silence area
17.	Denobli More	23.6909	86.4189	Silence area
18.	Katras	23.8095	86.2915	Industrial area
19.	Baliapur	23.7215	86.5274	Residential area
20.	Gaditundi	23.9745	86.4858	Residential area
21.	Maharajganj	23.9060	86.5144	Commercial area
22.	Pokhriya	23.9153	86.6271	Residential area

3.2 Modeling Approaches

3.2.1 Overview of MLR

Multiple linear regression is a statistical method for modeling the association across two or more independent variables and a dependent variable. It is an intricate version of simple linear regression, which only considers one independent variable. Multiple linear regression aims to identify the best-fitting linear equation that explains how the dependent variable varies when the values of all independent variables change. The MLR model enables researchers to investigate the combined effect of several independent variables on the dependent variable and quantify their separate contributions to the outcome. The general form of multiple linear regression is presented in Equation (1):

$$Y = \beta_0 + \beta_1 A_1 + \beta_2 A_2 \ldots\ldots \beta_k A_k + \varepsilon \tag{1}$$

where, y indicates the dependent or response variable, β_0 denotes the constant term or intercept, it is the value of the dependent variable when all the independent variables are zero, β_1, β_2,….. β_k denotes the regression coefficient of independent variables A_1, A_2, ….., A_k and ε represent the error term. In the present study, the selected independent variables encompass traffic volume, % of Heavy vehicles, % of 2-wheelers, % of 3-wheeler, % of car, % of LCV, and average speed of stream. However, equivalent traffic noise was regarded as the dependent variable during the modeling process.

The fundamental objective of multiple linear regression is to determine the regression coefficients (β_1, β_2,….., β_k) that reduce the sum of squared errors between the actual and anticipated values. These coefficients effectively describe the slope of the relationship between each independent variable and the dependent variable. It enables researchers and analysts to detect links between numerous variables and use the data to generate predictions or gain insights. It also provides a basis for more advanced regression and machine learning algorithms. In the current investigation, out of 330 data samples, 80 % of the dataset was used for training, while the remaining 20 % were employed to validate the developed model. The reliability and validity of the developed model were estimated using various statistical approaches, including the t-test, R^2 value, and F-test.

3.2.2 Artificial Neural Network Approach

Artificial neural networks are a machine learning approach inspired by the design and functioning of biological neural networks seen in the human brain. ANNs have grown in popularity because of their capacity to handle complicated, non-linear data relationships and their outstanding efficiency in various applications. ANNs comprise layers of interconnected nodes or neurons. The primary elements of construction are input, hidden, and output layers. Weighted connections connect each neuron in a layer to all neurons in the next layer. The architecture's depth (number of hidden layers) and width (number of neurons per layer) affect the network's ability to store complicated patterns. The present study utilizes a Multilayer Perceptron (MLP) neural network, which employs IBM SPSS software version 27. A Multilayer Perceptron (MLP) network's architecture typically comprises various layers: an input layer, one or more hidden layers, and a layer for the output. The flowchart of the proposed methodology for ANN modeling in the present study is shown in Figure 5.

Nevertheless, ANNs comprise several fundamental techniques for their applicability. Activation functions are one of these techniques that add nonlinearity to the network, allowing it to model complex data interactions. Common activation functions include the sigmoid, hyperbolic tangent, and rectified linear unit (ReLU). These functions convert a neuron's input signal into its output signal. Another vital technique is backpropagation, a vigorous training method for ANNs. In the backpropagation process, the error between expected and actual outputs is propagated backward through the network to modify the weights of connections. Gradient descent methods, which iteratively update the weights, are frequently employed to minimize error. Moreover, optimization techniques, such as stochastic gradient descent (SGD) and its derivatives (Adam, RMSprop, and others), define how weights are changed during training. These algorithms establish a balance between rapid convergence and avoiding local minima.

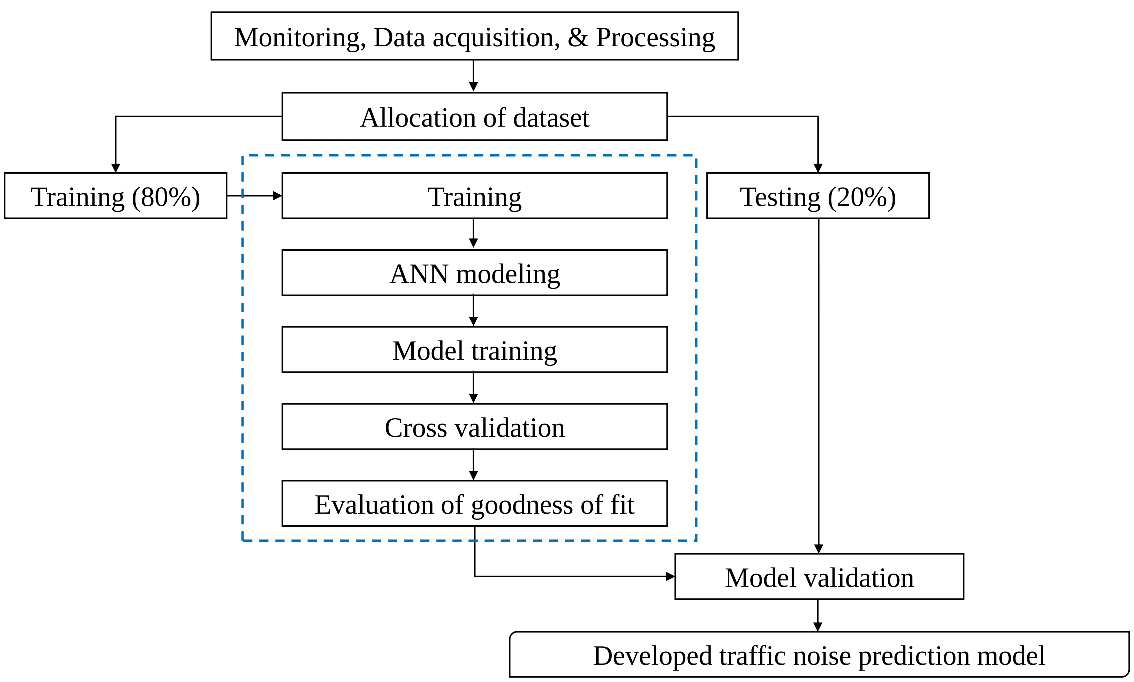

Figure 5. Flow diagram for ANN approach.

4. Performance Metrics

When comparing predicted and observed values in a regression analysis, the average magnitude of the error is measured by the Root Mean Square Error (RMSE). It computes the mean squared deviations between the expected and actual values or the square root of that mean. How well a model's predictions match the observed data may be clearly seen by looking at its root mean square error. Lower RMSE values show less prediction mistakes and, hence, higher model performance. An additional frequently used metric to assess the accuracy of regression models is the Mean Absolute Error (MAE). Each data point's average absolute difference between predicted and observed values is calculated. The coefficient of determination (R^2) provides information about the model's goodness of fit; higher values suggest that the model captures a larger percentage of the variation in the dependent variable. R^2 is a scale that ranges from 0 to 1, where 1 denotes perfect prediction. Mathematical expression for all these metrics is presented in Equations (2), (3), and (4).

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (O_i - P_i)^2} \quad (2)$$

$$\mathrm{MAE} = \frac{1}{n}\Sigma|O_i - P_I| \quad (3)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (O_i - P_i)^2}{\sum_{i=1}^{n} (O_i - O'_i)^2} \quad (4)$$

where, O_i represents the observed noise levels; P_i indicates the predicted values of noise levels; O'_i shows the mean of the observed value, and n is the number of samples.

5. Results and Discussion

5.1 Data Analysis

The present study's observed traffic noise levels range from 62 dBA to 80.5 dBA with mean value of 73.7 dBA. However, Topchachi exhibits the least amount of noise, while Mugma experiences the most significant level of traffic noise. Therefore, it is evident from the data analysis that all the sampling locations exceed the noise standards set by the Central Pollution Control Board (CPCB). Table 2 summarizes the data's vital descriptive statistics, including minimum, maximum, mean, and standard deviation. The hourly traffic volumes exhibit substantial variability, ranging from 414 to 4978 vehicles per hour across the chosen sampling locations. Notably, Big Bazar experiences the lowest traffic volume, while DAV Mahuda records the highest. Moreover, the average traffic speed of the stream ranges from 22 to 28.1 km/h, displaying variation within this speed range.

Furthermore, the data analysis shows that more than half of the total traffic comprises 2-wheeler vehicles, as depicted in Figure 6. Bank More has the maximum percentage of 2-wheelers, and Mugma has the highest proportion of heavy vehicles. Among all the seven vehicle categories, the percentage of Tractors/Trailers was minimal, with a value of 0.6 %.

Table 2. Descriptive statistics of the data.

Parameters	Notation	Minimum	Maximum	Mean	Std. Deviation
Traffic Volume (vehicle/h)	Q	414.00	4978.00	2332.88	869.08
% of Heavy Vehicles	% HV	1.10	33.30	10.20	5.86
% of 2-Wheelers	% 2-W	21.41	67.55	50.68	7.44
% of 3-Wheeler	% 3-W	6.03	30.78	19.53	4.99
% of Car	% Car	8.17	28.16	16.43	3.83
% of LCV	% LCV	1.20	7.54	3.13	1.08
Avg. Speed of Stream (km/h)	V	22.00	28.10	24.59	1.60
Equivalent Traffic Noise (dBA)	L_{eq}	62.00	80.50	73.75	3.61

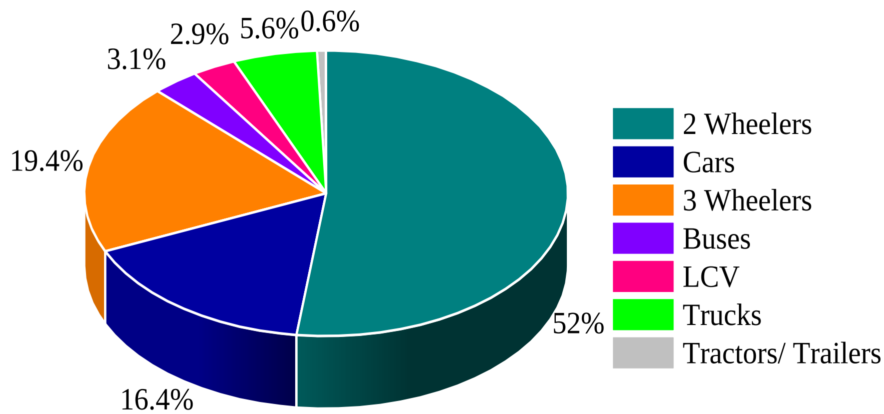

Figure 6. Percentage distribution of composition of all the vehicle categories.

5.2 Model Built using MLR

To develop a traffic noise model, a mathematical correlation is established between traffic-related parameters that influence the generation of traffic noise and the resulting L_{eq}. Previous literature in this field indicates that traffic parameters are critical in controlling traffic noise levels (Garg et al., 2015; To et al., 2002; Yadav et al., 2023). The literature demonstrates that traffic-related variables significantly impact the overall noise produced by traffic. Additionally, to ensure the validity of Multiple Linear Regression (MLR), it is essential to verify several assumptions, including linearity, multicollinearity, homoscedasticity, and independence of observation. All the statistical analyses in the current study were carried out using IBM SPSS Statistics 27 software.

Traffic volume and vehicular speed were reformed to logarithmic form to ensure the linearity in the dataset as done by previous studies (Calixto et al., 2003; Garg et al., 2015). Pearson correlation analysis is performed to assess the relationship between the independent variables and traffic noise level. The Pearson correlation coefficient (r) is used to quantify the strength of the linear connection between two variables. A correlation coefficient greater than 0.5 indicates a good correlation, suggesting a robust linear relationship between the variables. When the correlation coefficient falls within the range of 0.3 to 0.5, it means a moderate correlation, indicating a reasonably decent linear relationship between the variables. In the present research, the chosen independent variables (Log Q, % HV, % 2W, % 3W, % Car, % LCV, and Log V) demonstrate a significant correlation with the response variable (L_{eq}), with $r > 0.3$, as depicted in Figure 7. The figure illustrates a good correlation between traffic volumes, % of heavy vehicles, % of 2 wheelers, and % of 3 wheelers with the L_{eq}. Additionally, there is a moderate correlation between the % of cars, the proportion of light commercial vehicles, and the average vehicular speed with the L_{eq}. Therefore, all the chosen independent variables were checked for their linearity with L_{eq} and confirmed a linear relationship between them.

A collinearity diagnostic was conducted using the Tolerance and Variance Inflation Factor (VIF) to assess multicollinearity among the independent variables. The results indicated that the VIF values were considerably below 10, and the Tolerance values were greater than 0.1 for each independent variable. This analysis confirms the absence of significant multicollinearity among the independent variables (Uyanık & Güler, 2013). In addition, homoscedasticity and independence of variables were assessed using a residual plot and the Durbin-Watson statistic, respectively. In Figure 8, the plot between the standardized z-scores of the predicted variable and the residuals indicates a strong relationship, as the scores are cloud-clustered on both sides of the fitted regression line. This suggests that the assumption of homoscedasticity was met, as the distribution of the residuals remains consistent across the range of predicted values. Regarding the independence of observations, the Durbin-Watson statistic resulted in a value of 0.807. Since this value is less than the specified threshold of 4, it confirms the independence of observations (Ul-Saufie et al., 2013).

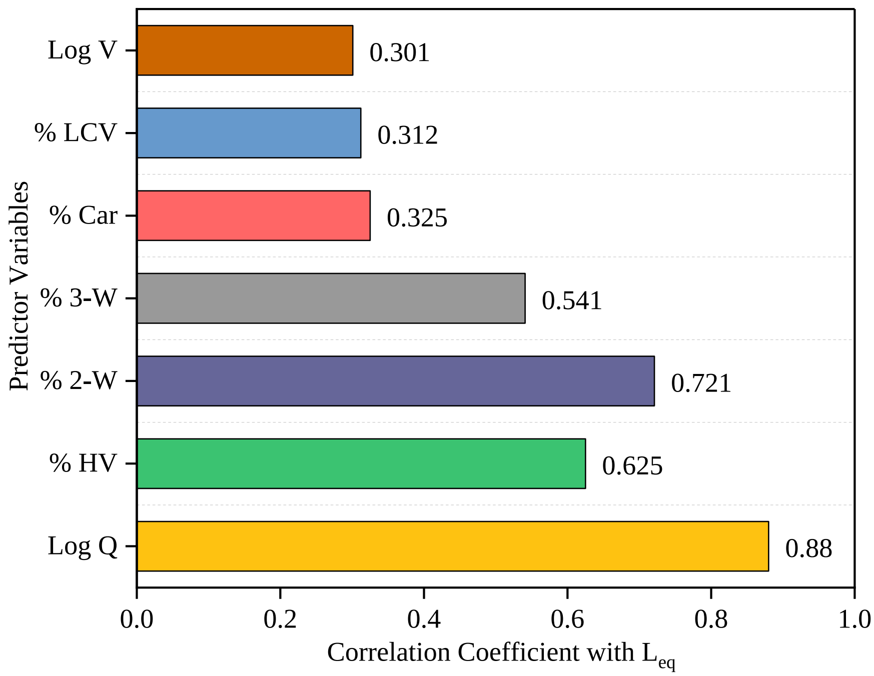

Figure 7. Pearson's correlation coefficient of dependent variable with L_{eq}.

A collinearity diagnostic was conducted using the Tolerance and Variance Inflation Factor (VIF) to assess multicollinearity among the independent variables. The results indicated that the VIF values were considerably below 10, and the Tolerance values were greater than 0.1 for each independent variable. This analysis confirms the absence of significant multicollinearity among the independent variables (Uyanık & Güler, 2013). In addition, homoscedasticity and independence of variables were assessed using a residual plot and the Durbin-Watson statistic, respectively. In Figure 8, the plot between the standardized z-scores of the predicted variable and the residuals indicates a strong relationship, as the scores are cloud-clustered on both sides of the fitted regression line. This suggests that the assumption of homoscedasticity was met, as the distribution of the residuals remains consistent across the range of predicted values. Regarding the independence of observations, the Durbin-Watson statistic resulted in a value of 0.807. Since this value is less than the specified threshold of 4, it confirms the independence of observations (Ul-Saufie et al., 2013).

Furthermore, multiple linear regression was conducted to develop a predictive model for road traffic noise. The statistical analyses are performed at a 95% confidence interval, establishing the threshold for statistical significance at p-values below 0.05. The regression model, shown in Table 3, includes the unstandardized regression coefficients (β) and their corresponding p-values for all variables considered in the model. Notably, the p-values for the model's dependent variables are less than 0.05, confirming their significance at the 95% confidence interval. The examination of β values for all the variables revealed that traffic volume is the most influential factor in the developed model, exhibiting a significant β value of 17.794 with a corresponding t-value of 25.571. The VIF and Tolerance values provide evidence that there are no concerns regarding multicollinearity in the model, as discussed earlier.

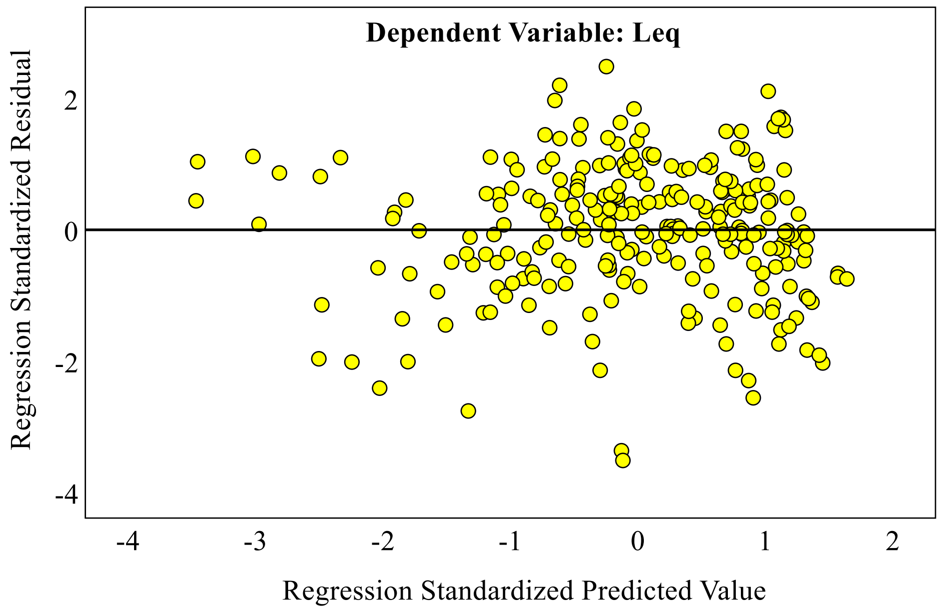

Figure 8. Scatterplot of z-residual vs z-predicted.

Table 3. Regression coefficient and statistical constraints for the developed model.

Variables	β value	Standard Error	t	Significance	Tolerance	VIF
Constant	26.374	4.911	12.298	0.195		
Log Q	17.794	0.696	25.571	0.000	0.591	1.691
% HV	1.229	0.023	5.588	0.000	0.677	1.478
% 2-W	2.886	0.118	4.741	0.000	0.630	1.587
% 3-W	0.358	0.054	3.521	0.000	0.591	1.512
% Car	0.162	0.043	4.012	0.000	0.610	1.385
% LCV	0.183	0.018	4.005	0.000	0.715	1.581
Log V	-10.445	3.364	-3.105	0.002	0.912	1.096

The accuracy and reliability of the developed model were assessed using various performance metrics shown in Table 4. These metrics include coefficient of determination (R^2), Standard error (SE), Mean absolute error (MAE), and Root mean square error (RMSE). The R^2 value for the fitted model was found to be 0.81, suggesting the higher prediction accuracy of the developed model. Therefore, this model can predict traffic noise with an accuracy of 81%, meaning that model fitness is adequate. Moreover, Standard error, MAE, and RMSE values are very low, supporting the model's accuracy and reliability. The results of calculated errors are consistent with the findings of other previous studies (Khajehvand et al., 2021; Yadav et al., 2023). Analysis of the variance (ANOVA) test was also performed, and test results are presented in Table 5. Based on the outcome of ANOVA, the model appears to be highly significant with a p-value less than 0.001 (<0.05), indicating that it effectively explains the variability in the dependent variable using four predictor variables. As a result, the traffic noise level is significantly associated with the predictor variable, implying that the regression model is a good fit for the data.

Apart from the expected causes, various unsuspected factors also influenced the actual noise level, including noise generated by pedestrians and shopkeeper-customer conversation. In addition, rickshaw drivers park their vehicles on the corner of the roadway and call out to passengers. Likewise, some car, rickshaw, and truck drivers play loud music inside the vehicles, further affecting the measured noise level. Despite all these unexpected causes of variation in measured noise levels, the developed model in the present study can accurately predict traffic noise levels.

Table 4. Accuracy parameters for the model.

Model	R-Square	Adjusted R-Square	Std. Error	MAE	RMSE	Durbin-Watson
1	0.819	0.816	1.60	1.20	1.55	0.807

Table 5. Result of the ANOVA test.

Model		Sum of Squares	Degree of freedom	Mean Square	F	Significance
1	Regression	3025.710	7	756.428	292.812	< 0.001
	Residual	699.081	259	2.583		
	Total	3694.791	263			

5.3 Model Validation

The model's performance is evaluated using various metrics, such as R^2 and MAE, as indicated in Table 4. However, model validation is conducted using a testing dataset to ensure model suitability for real-world scenarios. This dataset assesses how well the model predicts the L_{eq} by comparing it with the observed L_{eq}. Therefore, scatter plots are created to visualize the comparison between predicted and observed L_{eq} values for the testing data, as depicted in Figure 9. The scatter plot shows the relationship between the predicted and observed traffic noise levels along the fitted line. The data points distribution on this line indicates how closely the model predictions align with the actual measurements. Based on scatter plots, an equation mentioned in Table 6 is derived consisting of observed and predicted equivalent traffic noise levels as two variables. Analyzing Table 6, it becomes evident that goodness of fit is appropriate, with an R^2 value of 0.84, indicating higher prediction accuracy.

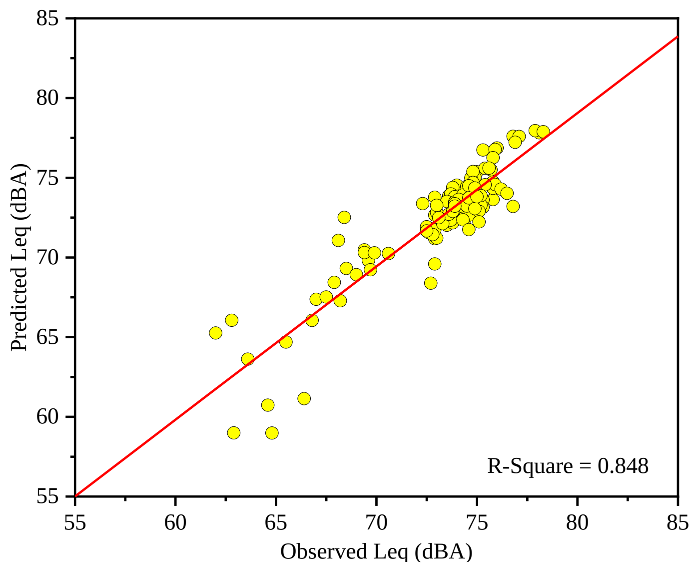

Figure 9. Observed vs Predicted noise level for MLR approach.

As per the developed model, the MAE for the training and testing dataset are 1.20 and 1.21 dBA, respectively. Likewise, the RMSE for the training and testing dataset are 1.55 and 1.69 dBA, respectively. The models exhibit minor error values in both the training and testing phases, indicating high accuracy in their predictions. These favorable results suggest that the developed models can accurately forecast traffic noise in Tier 2 cities. As a result, these models have the potential to be helpful in anticipating and minimizing traffic-related noise levels in such urban settings.

Table 6. Relation between predicted and observed L_{eq} with associated errors.

Model	Equation	R^2	MAE	RMSE
	y = 1.0343 x – 3.0406	R^2 = 0.848	1.21	1.69

5.4 Relative Importance Analysis

The relative importance of predictor variables in a multiple linear regression model is crucial in analyzing the relationships between predictors and the dependent variable. In the current research, R-squared-based metric analysis is utilized to establish a ranking of the relative importance of the seven chosen input variables that impact roadway noise levels (Hamad et al., 2017). According to R-squared-based metric analysis, change in R^2 value is measured after removing a particular input variable from the dataset. The input variable that contributes the most to the model's performance is the one whose removal results in the most significant decrease in the R^2 value. Hence, following the previous step, when traffic volume was excluded, there was a substantial 65% reduction in the R^2 value, reduced from 0.819 to 0.279 (Figure 10a). Therefore, Traffic volume ranked first as the most important contributing variable in the developed traffic noise prediction model.

Moreover, excluding average traffic speed led to a significant reduction of 24% in the R^2 value, changing it from 0.819 to 0.618. Thus, the average traffic speed ranked as the second most crucial input variable in the developed traffic noise prediction model. Similarly, the minimum R^2 change (from 0.819 to 0.811) was reported when the %LCV was excluded, suggesting the minimum contribution in the prediction of L_{eq}. The reduction in the R^2 value for %LCV is minimal among all the considered variables; therefore, it was ranked last in terms of importance in the developed model. The ranking of the seven selected input variables as per their relative importance in the developed regression model is illustrated in Figure 10b. Therefore, this section provides significant insights into which input variables have a greater impact on the outcome and assists researchers and decision-makers in making enlightened decisions.

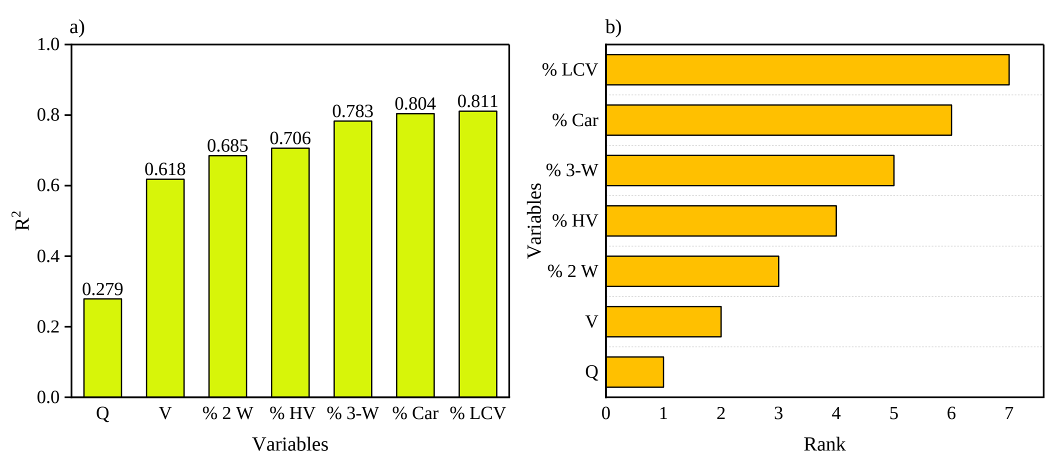

Figure 10. Relative importance analysis (a) Changed R^2 (b) Ranking.

5.5 ANN Modeling

All the analyses involved in the ANN modeling, including training, testing, and validation, were carried out in IBM SPSS software version 27. For training the artificial neural networks dataset, the process called Feed-Forward-Backward Propagation (FFBP) has been used (Chen et al., 2020). It comprises two major phases: forward propagation and backward propagation (also known as backpropagation). Feed-forward and backpropagation are critical components of artificial neural network training. During the feed-forward phase, input data is transferred through the network layer by layer, with each neuron producing an output using weighted calculation and activation functions. This procedure makes a prediction, which is then compared to the expected outcome with the loss function. Gradients of the loss function concerning the network parameters are computed in reverse order during the backpropagation phase, beginning with the output layer and progressing backward to the hidden layers. These gradients direct the optimization of the network's weights and biases, intending to minimize the prediction error. In this work, the gradient descent optimization technique has been utilized.

The present study employs an ANN characterized by a distinct architectural configuration, as depicted in Figure 11. The input layer encompasses eight input nodes, showing all the predictor variables: the count of 2-wheelers, cars, 3-wheelers, buses, LCVs, trucks, tractor/trailers, and average traffic speed. Simultaneously, the single hidden layer comprises six neurons, excluding the bias unit. The output layer includes one neuron that represents the response variable; in this study, it is the equivalent noise level generated by traffic.

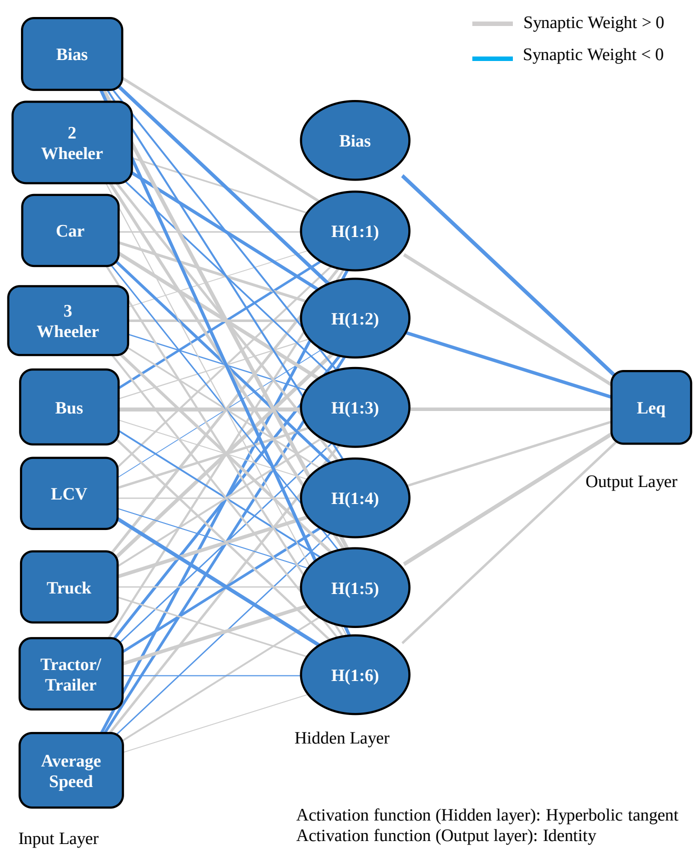

Figure 11. Architecture of the neural network.

The activation functions applied in this context involve using the hyperbolic tangent function within the hidden layer and the identity function within the output layer (Debnath et al., 2022). A comprehensive overview of the proposed ANN model's network specifications is provided in Table 7. Before progressing further, dividing the dataset into distinct training and testing subsets is essential. The primary objective behind this division lies in assessing the model's generalization to previously unseen data. Accordingly, the model was trained with the help of a training dataset, and its performance was subsequently evaluated using the testing dataset. To standardize the features, the training and testing dataset was preprocessed using Min-Max scaling. This approach allows us to obtain a more realistic estimate of the model's effectiveness under real-world conditions by subjecting it to data it has not encountered during training.

Moreover, 80% of the total dataset was allocated for training the model, and the remaining 20% was assigned for testing purposes (Singh et al., 2016). Following numerous iterations of the learning process, reducing errors is possible, thus substantially enhancing predictive accuracy. To counteract potential overfitting in this research, the authors employed a ten-fold cross-validation methodology to evaluate the performance of the ANN model (Singh et al., 2016).

Table 7. Network specification of the model.

1st Layer (Input)	Covariates	1	2-Wheeler
		2	Car
		3	3-Wheeler
		4	Buse
		5	LCV
		6	Truck
		7	Tractor/Trailer
		8	Average speed
	Number of Units[a]		8
	Rescaling Method for Covariates		Standardized
2nd Layer (Hidden)	Total Number of Hidden Layers		1
	Total Number of Units in Hidden Layer 1[a]		6
	Used Activation Function		Hyperbolic tangent
3rd Layer (Output)	Predicted Variables	1	L_{eq}
	Total Number of Units		1
	Method of Rescaling for Scale Dependents		Standardized
	Used Activation Function		Identity
	Used Error Function		Sum of Squares

Note: 'a' denotes the exclusion of the bias unit.

5.5.1 Performance Evaluation of the ANN Model

A comprehensive overview of the performance metrics of the ANN model across ten distinct runs is illustrated in Table 8, separating the training and testing stages. This information is critical for determining the ability of the model to learn from training data and generalize to unseen testing data. During the training stage, the Sum of Squared Error (SSE), and Mean Absolute Error (MAE) displays the degree of fit between the model's predictions and the trained data. Moreover, the Root Mean Squared Error (RMSE) shows the normalized representation of the error, giving insights into the model's accuracy on the training dataset.

During training, the model attains an MAE of 0.011 and RMSE of about 0.195, and these values imitate a typically favorable performance with minimal errors. The low standard deviation (SD) values for MAE (0.003) and RMSE (0.016) designate firmness in the model's performance over the ten runs. Subsequently, the testing phase inspects the model's capability to generalize to new and previously unseen data. During the testing phase, the model attains an average MAE of 0.153 and RMSE of 0.204, signifying a reasonably excellent performance. However, as demonstrated by the lower RMSE value in the testing phase, the model's performance on the testing data looks weak compared to its efficacy on the training data.

Table 8. Error in the training and testing phase.

Training				Testing				Overall Sample size
N	**SSE (Training)**	**MAE (Training)**	**RMSE (Training)**	**N**	**SSE (Testing)**	**MAE (Testing)**	**RMSE (Testing)**	
270	2.587	0.010	0.182	60	8.919	0.149	0.208	330
267	3.259	0.012	0.186	63	9.223	0.146	0.227	330
262	1.849	0.007	0.172	68	7.785	0.114	0.165	330
260	3.192	0.012	0.219	70	12.430	0.178	0.214	330
265	2.539	0.010	0.207	65	11.377	0.175	0.198	330
258	4.438	0.017	0.212	72	11.560	0.161	0.248	330
262	3.099	0.012	0.200	68	10.498	0.154	0.213	330
258	3.025	0.012	0.181	72	8.473	0.118	0.205	330
269	1.926	0.007	0.216	61	12.514	0.205	0.178	330
264	2.211	0.008	0.179	66	8.492	0.129	0.183	330
Mean	2.813	0.011	0.195	**Mean**	10.127	0.153	0.204	
SD	0.729	0.003	0.016	**SD**	1.671	0.028	0.023	

Note: N = Sample size, SD = Standard Deviation

In addition to the SSE, MAE, and RMSE, another performance metric, the Coefficient of determination (R^2), was also analyzed to check the prediction accuracy. The scatter plot in Figure 12 illustrates the relationship between observed and predicted noise levels. The high R^2 value of 0.93 indicates a significant connection between the two. In statistical terms, the ANN model can explain approximately 93% of the variance in traffic noise levels. Therefore, such a high R^2 value strongly indicates the model's accuracy and efficiency in predicting noise levels. This finding is particularly promising for applications where precise noise level predictions are crucial.

Furthermore, Table 9 provides a mathematical expression representing the relationship between observed and predicted noise levels. The best-fit line equation, $y = 4.05 + 0.94x$, signifies this relationship, where 'x' stands for the observed noise level, and 'y' stands for the predicted noise level. This equation underscores the strong linear relationship between the two variables. It means that for every unit increase in the observed noise level ('x'), the predicted noise level ('y') increases by 0.94 units, and there is a baseline noise level of 4.05 units. This demonstrates the model's capability to estimate noise levels reliably. Hence, the developed ANN model excels in predicting traffic noise levels, as evidenced by its high R^2 value and the strong linear relationship between predicted and observed values.

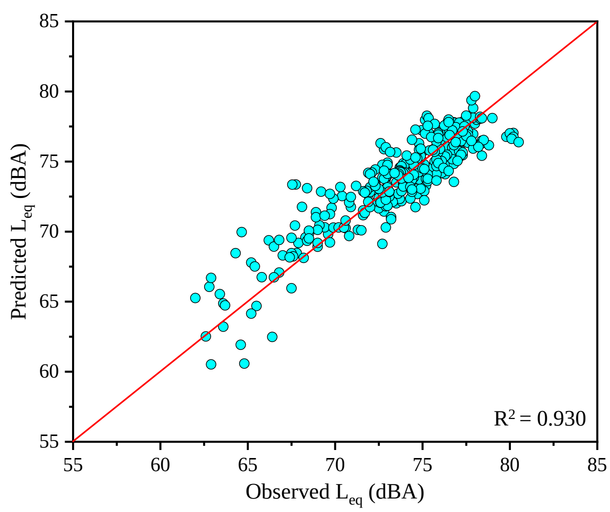

Figure 12. Observed vs Predicted noise level.

Table 9. Relationship among the observed and predicted $L_{eq.}$

Model	Equation	R^2
	y = 4.05 + 0.94 x	0.93

5.5.2 Relative Importance of Predictors

This section provides a thorough sensitivity analysis for predictive variables in an ANN model, demonstrating the effect of various input attributes on the model's predictions. This assessment is critical for determining which variables have the most significant impact on the model's output, which aids in feature selection, model interpretation, and possible modifications. Table 10 provides insightful data about the sensitivity of various predictive variables in ANN models and their contribution to forecasting equivalent noise levels. Each row in the table represents a different independent variable, including several vehicle types and average speed.

To determine the normalized importance of these neurons, we calculated their relative importance by dividing each neuron's significance by the maximum importance observed. The outcome is denoted as a percentage, and analysis revealed that the 2-wheelers emerge as the most crucial predictor, with a normalized importance of 100%. Subsequently, the tractors and trailers factor holds a significant normalized importance of 52.6%. In descending order, the 3-wheelers follow with a 46.3% normalized importance, trailed by the buses at 41.2%. Furthermore, cars, trucks, and LCVs carry substantial significance, with normalized values of 33.6%, 23.5%, and 22.6% respectively. Lastly, the Average Speed holds a significant normalized importance of 18.4%. Besides, according to the normalized importance of all the predictor variables, ranking is assigned to each, as shown in Figure 13.

Table 10. Sensitivity analysis results.

Network (N)	N (1)	N (2)	N (3)	N (4)	N (5)	N (6)	N (7)	N (8)	N (9)	N (10)	Importance (Average)	Normalized Importance in %
2-W	1	1	1	1	1	1	1	1	1	1	1	100
Car	0.39	0.21	0.48	0.33	0.13	0.19	0.23	0.51	0.31	0.62	0.34	33.6
3-W	0.6	0.44	0.49	0.44	0.28	0.43	0.4	0.52	0.49	0.67	0.48	46.3
Bus	0.42	0.25	0.54	0.47	0.27	0.4	0.35	0.54	0.32	0.58	0.41	41.2
LCV	0.26	0.26	0.22	0.14	0.14	0.36	0.22	0.27	0.19	0.24	0.23	22.6
Truck	0.21	0.18	0.34	0.22	0.12	0.21	0.21	0.29	0.3	0.24	0.23	23.5
Tractor/ Trailer	0.43	0.53	0.6	0.44	0.41	0.44	0.48	0.54	0.59	0.7	0.52	52.6
Average Speed	0.27	0.09	0.2	0.19	0.14	0.13	0.17	0.21	0.18	0.33	0.19	18.4

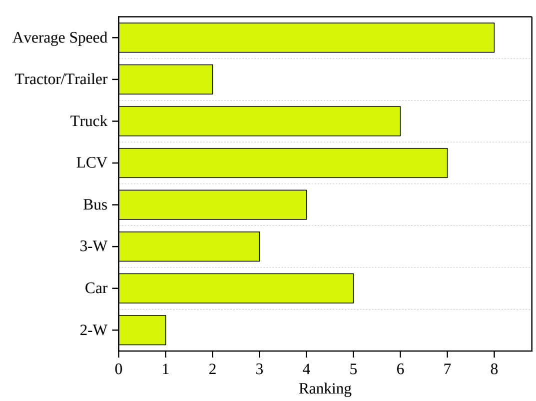

Figure 13. Ranking of the predictor variables.

6. Comparative Analysis Based on Goodness of Fit

The results of the present research were validated through comparison with existing studies in the academic literature. Table 11 provides a comparison of the results obtained in the present study with those from previous research in the field of traffic noise prediction modeling. It evaluates the effectiveness of these models in predicting traffic noise in different cities, both in India and abroad, with a focus on the use of MLR and ANN modeling. Research conducted at various study locations using the regression approach yielded R^2 values ranging from 0.54 to 0.85 in previous literature (Khajehvand et al., 2021; Motylewicz & Gardziejczyk, 2020; Singh et al., 2013; Thakre et al., 2020). R^2 values in the range of 0.81 to 0.85 were attained by studies in Indian cities like Delhi, Patiala (Garg et al., 2015; Kumar et al., 2014; Singh et al., 2016), Chongqing city, China (Chen et al., 2020), and New Klang Valley, Malaysia (Ahmed et al., 2021). However, studies in Surat City, India, and Granada, Spain, showed a slightly lower R^2 ranging from 0.75-0.76 (Genaro et al., 2010; Ranpise et al., 2021). A novel approach employed an Emotional Neural Network in Nicosia, North Cyprus, resulting in an R^2 of 0.81 (Nourani et al., 2020).

All the studies discussed in this context exhibit relatively modest outcomes, with R^2 values below 0.87, except for research conducted in Serbia by Nedic et al. (2014), which achieved a notably higher R^2 value of 0.93, consistent with the current study's findings. This higher value of R^2 for ANN approach in the present study demonstrates a high level of predictive capability and the ability to explain a significant portion of the variance in traffic noise data. The primary reason behind these results can be attributed to the appropriateness of the neural network's design, specifically its feed-forward architecture introducing eight input parameters and a hidden layer that serves as the activation function for weighted input layers. Furthermore, it is notable that the sum of squares error values during the network's training and testing phases remained remarkably minimal.

Table 11. Prior research conducted on linear regression and ANN modeling.

References	Employed Approach	Study area	R^2 of MLR/ANN
Genaro et al. (2010)	ANN	Granada, Spain	0.76
Singh et al. (2013)	MLR	Delhi, India	0.70
Kumar et al. (2014)	ANN	Patiala, India	0.85
Nedic et al. (2014)	ANN	Serbia	0.93
Garg et al. (2015)	ANN, MLR	Delhi, India	0.81 (ANN), 0.54 (MLR)
Singh et al. (2016)	ANN, RF, DT, and Linear regression	Patiala, India	0.83
Thakre et al. (2020)	MLR	Nagpur, India	0.65
Chen et al. (2020)	ANN	Chongqing City, China	0.82
Motylewicz & Gardziejczyk (2020)	MLR	Bialystok, Poland	0.85
Nourani et al. (2020)	Emotional neural network	Nicosia, North Cyprus	0.81
Ranpise et al. (2021)	ANN	Surat, India	0.75
Khajehvand et al. (2021)	MLR	Karaj, Iran	0.82
Kumar (2021)	ANN, RSM	Delhi, India	0.86
Ahmed et al. (2021)	ANN, CFS-ANN, Ensemble RF-ANN	New Klang Valley, Malaysia	0.83
Present study	ANN	Dhanbad, India	0.84 (MLR), 0.93 (ANN)

7. Conclusions

The findings from this study offer a positive outlook for the development of an additional tool to assess and predict urban environmental noise, aiming to mitigate acoustic pollution. In the present study, various traffic parameters that impact the generation of traffic noise were explicitly chosen. These parameters function as the independent variables in the modeling procedure. Two widely used soft computing approaches, namely

MLR and ANN, are used to model traffic noise. The following are the significant findings of the present study:

- The developed MLR model demonstrated its ability to predict noise levels with RMSE of 1.55 and 1.69 in training and testing, respectively, with R^2 value of 0.84. This R^2 value indicates that the MLR model explains 84% of variations in noise levels.
- The application of ANN for forecasting traffic noise yields favorable outcomes, with an RMSE of 0.195 and 0.204 during the training and testing phases, respectively.
- The R^2 value for ANN model is 0.93, highlighting the model's substantial accuracy and efficiency. Furthermore, sensitivity analysis demonstrates 2-wheelers as the major contributing factors for predicting traffic noise in the case of the ANN approach.
- The ANN model outperforms the MLR model, exhibiting outstanding accuracy and showcasing its capacity to manage complex datasets effectively.
- The proposed model in the present study can be implemented in the other mid-sized cities of India with heterogeneous traffic mix and similar kind of traffic flow pattern.

The accuracy and efficiency of the model suggest its potential usefulness in various practical applications. By helping to identify noise hotspots, these models enable urban planners to focus interventions where traffic noise is most detrimental. Urban planners can strategically distribute spaces and create zones that minimize noise impact on sensitive regions such as residential neighborhoods and educational institutions by including noise considerations in land use planning. Based on noise forecasts, transportation infrastructure can be optimized by modifying the layout of roads and highways, adding noise barriers, and using pavement materials that reduce noise. Furthermore, traffic control methods, such as optimized traffic signals and speed restrictions, can effectively reduce noise levels. The models are also used to guide the layout of urban green spaces and the construction of noise barriers. Consequently, this research contributes valuable insights to the field of noise prediction and its applications in addressing environmental and urban challenges.

7.1 Limitations and Scope for Future Work

The developed model is based on the supervised learning technique and data dependent which is calibrated and validated in Indian traffic conditions. The model is site-dependent, with training primarily based on the unique traffic noise situations of a mid-sized Indian city. Local traffic flow patterns, road geometry, driving conditions, and environmental conditions are few of the input parameters that contribute to this site dependability. When the model is applied to sites that have extremely distinct contexts, these input parameters may not be properly captured. Therefore, a foreseen challenge is the potential degree of failure when testing the model in a country different from the one used for calibration. To confront this challenge adeptly, it is essential to integrate a set of recommendations into the course of future research initiatives. Firstly, considering localized calibration based on local data to account for region-specific traffic noise characteristics. Secondly, exploring data augmentation techniques to simulate variations in traffic noise conditions, especially in areas with limited available data. Lastly, implementing cross-validation frameworks that involve training the model on data from one country and validating it on data from another country to assess its transferability.

The proposed model produced satisfactory outcomes, although there is potential for improvement. Our models did not account for noise-related factors such as pavement type, gradient, acceleration, deceleration, honking, and distance from the edge of the. As a result, future studies should explore including some of these factors. Additionally, more advanced soft computing approaches such as Decision Tree, Extra Tree, Random Forest, Gradient Boosting, and KNN can be applied to predict traffic noise. Moreover, in future research, expanding the number of sampling locations can be pursued, leveraging substantial datasets. Subsequent exploration may involve conducting a comprehensive analysis of the correlation existing between different vehicle categories and their associated noise levels.

Conflict of Interest

The authors declare that there is no conflict of interest for this publication.

Acknowledgments

The authors would like to express sincere thanks to the Department of Environmental Science and Engineering, IIT (ISM) Dhanbad, for their outstanding assistance and for providing the resources that enabled us to conduct this research.

References

Ahac, S., Ahac, M., Domitrović, J., & Dragčević, V. (2021). Modeling the influence of roundabout deflection on its efficiency as a noise abatement measure. *Sustainability, 13*(10), 5407. https://doi.org/10.3390/su13105407.

Ahac, S., Ahac, M., Majstorovi, I., & Basic, S. (2023). Speed reduction capabilities of two-geometry roundabouts. *Applied Sciences, 13*(21), 11816. https://doi.org/10.3390/app132111816.

Ahmed, A.A., Pradhan, B., Chakraborty, S., & Alamri, A. (2021). Developing vehicular traffic noise prediction model through ensemble machine learning algorithms with GIS. *Arabian Journal of Geosciences, 14*, 1564. https://doi.org/10.1007/s12517-021-08114-y.

Babisch, W., & Van Kamp, I. (2009). Exposure-response relationship of the association between aircraft noise and the risk of hypertension. *Noise and Health*, *11*(44), 161-168. https://doi.org/10.4103/1463-1741.53363.

Calixto, A., Diniz, F.B., & Zannin, P.H.T. (2003). The statistical modeling of road traffic noise in an urban setting. *Cities*, *20*(1), 23-29. https://doi.org/10.1016/s0264-2751(02)00093-8.

Chang, T.Y., Liang, C.H., Wu, C.F., & Chang, L.T. (2019). Application of land-use regression models to estimate sound pressure levels and frequency components of road traffic noise in Taichung, Taiwan. *Environment International, 131*, 104959. https://doi.org/10.1016/j.envint.2019.104959.

Chen, L., Tang, B., Liu, T., Xiang, H., Sheng, Q., & Gong, H. (2020). Modeling traffic noise in a mountainous city using artificial neural networks and gradient correction. *Transportation Research Part D: Transport and Environment, 78*, 102196. https://doi.org/10.1016/j.trd.2019.11.025.

Chen, X., Chen, H., Yang, Y., Wu, H., Zhang, W., Zhao, J., & Xiong, Y. (2021). Traffic flow prediction by an ensemble framework with data denoising and deep learning model. *Physica A: Statistical Mechanics and Its Applications, 565*, 125574. https://doi.org/10.1016/j.physa.2020.125574.

Chevallier, E., Leclercq, L., Lelong, J., & Chatagnon, R. (2009). Dynamic noise modeling at roundabouts. *Applied Acoustics, 70*(5), 761-770. https://doi.org/10.1016/j.apacoust.2008.09.009.

Debnath, A., & Singh, P.K. (2018). Environmental traffic noise modelling of Dhanbad township area - A mathematical based approach. *Applied Acoustics, 129,* 161-172. https://doi.org/10.1016/j.apacoust.2017.07.023.

Debnath, A., Singh, P.K., & Banerjee, S. (2022). Vehicular traffic noise modelling of urban area-a contouring and artificial neural network based approach. *Environmental Science and Pollution Research, 29*(26)*,* 39948-39972.

Garg, N., Mangal, S.K., Saini, P.K., Dhiman, P., & Maji, S. (2015). Comparison of ANN and analytical models in traffic noise modeling and predictions. *Acoustics Australia, 43*(2),179-189. https://doi.org/10.1007/s40857-015-0018-3.

Garg, N., Sinha, A.K., Dahiya, M., Gandhi, V., Bhardwaj, R.M., & Akolkar, A.B. (2017). Evaluation and analysis of environmental noise pollution in seven major cities of India. *Archives of Acoustics*, *42*(2), 175-188. https://doi.org/10.1515/aoa-2017-0020.

Genaro, N., Torija, A., Ramos-Ridao, A., Requena, I., Ruiz, D.P., & Zamorano, M. (2010). A neural network based model for urban noise prediction. *The Journal of the Acoustical Society of America, 128*(4), 1738-1746. https://doi.org/10.1121/1.3473692.

Gilani, T.A., & Mir, M.S. (2021). Modelling road traffic noise under heterogeneous traffic conditions using the graph-theoretic approach. *Environmental Science and Pollution Research, 28*(27), 36651-36668. https://doi.org/10.1007/s11356-021-13328-4.

Givargis, S., & Karimi, H. (2010). A basic neural traffic noise prediction model for Tehran's roads. *Journal of Environmental Managemen*t*, 91*(12), 2529-2534. https://doi.org/10.1016/j.jenvman.2010.07.011.

Golmohammadi, R., Abbaspour, M., Nassiri, P., & Mahjub, H. (2009). A compact model for predicting road traffic noise. *Iranian Journal of Environmental Health Science and Engineering, 6*(3), 181-186.

Gupta, A., Gupta, A., Jain, K., & Gupta, S. (2018). Noise pollution and impact on children health. The *Indian Journal of Pediatrics, 85*(4), 300-306. https://doi.org/10.1007/s12098-017-2579-7.

Hamad, K., Ali Khalil, M., & Shanableh, A. (2017). Modeling roadway traffic noise in a hot climate using artificial neural networks. *Transportation Research Part D: Transport and Environment, 53*, 161-177. https://doi.org/10.1016/j.trd.2017.04.014.

Khajehvand, M., Rassafi, A.A., & Mirbaha, B. (2021). Modeling traffic noise level near at-grade junctions: Roundabouts, T and cross intersections. *Transportation Research Part D: Transport and Environment, 93*, 102752. https://doi.org/10.1016/j.trd.2021.102752.

Konbattulwar, V., Velaga, N.R., Jain, S., & Sharmila, R.B. (2016). Development of in-vehicle noise prediction models for Mumbai Metropolitan Region, India. *Journal of Traffic and Transportation Engineering*, *3*(4), 380-387. https://doi.org/10.1016/j.jtte.2016.04.002.

Kumar, K., Parida, M., & Katiyar, V.K. (2011). Road traffic noise prediction with neural networks - A review. *An International Journal of Optimization and Control: Theories & Applications*, *2*(1), 29-37. https://doi.org/10.11121/ijocta.01.2012.0059.

Kumar, P. (2021). Traffic noise prediction and optimization using response surface method (RSM). *Arabian Journal of Geosciences, 14*(21), 2181. https://doi.org/10.1007/s12517-021-08547-5.

Kumar, P., Nigam, S.P., & Kumar, N. (2014). Vehicular traffic noise modeling using artificial neural network approach. *Transportation Research Part C: Emerging Technologies*, *40*, 111-122. https://doi.org/10.1016/j.trc.2014.01.006.

Lipowicz, A., & Lopuszanska, M. (2005). Marital differences in blood pressure and the risk of hypertension among polish men. *European Journal of Epidemiology, 20*(5), 421-427. https://doi.org/10.1007/s10654-005-1752-x.

Ma, T., Antoniou, C., & Toledo, T. (2020). Hybrid machine learning algorithm and statistical time series model for network-wide traffic forecast. *Transportation Research Part C: Emerging Technologies, 111*, 352-372. https://doi.org/10.1016/j.trc.2019.12.022.

Manojkumar, N., Basha, K., & Srimuruganandam, B. (2019). Assessment, prediction and mapping of noise levels in Vellore city, India. *Noise Mapping, 6*(1), 38-51. https://doi.org/10.1515/noise-2019-0004.

Mishra, R.K., Nair, K., Kumar, K., & Shukla, A. (2021). Dynamic noise mapping of road traffic in an urban city. *Arabian Journal of Geosciences*, *14*(2), 122. https://doi.org/10.1007/s12517-020-06373-9.

Motylewicz, M., & Gardziejczyk, W. (2020). Statistical model for traffic noise prediction in signalised roundabouts. *Bulletin of the Polish Academy of Sciences: Technical Sciences*, *68*(4), 937-948.

Münzel, T., & Daiber, A. (2018). Environmental stressors and their impact on health and disease with focus on oxidative stress. *Antioxidants and Redox Signaling*, *28*(9), 735-740. https://doi.org/10.1089/ars.2017.7488.

Nedic, V., Despotovic, D., Cvetanovic, S., Despotovic, M., & Babic, S. (2014). Comparison of classical statistical methods and artificial neural network in traffic noise prediction. *Environmental Impact Assessment Review, 49,* 24-30. https://doi.org/10.1016/j.eiar.2014.06.004.

Nourani, V., Gökçekuş, H., Umar, I.K., & Najafi, H. (2020). An emotional artificial neural network for prediction of vehicular traffic noise. *Science of the Total Environment, 707*, 136134. https://doi.org/10.1016/j.scitotenv.2019.136134.

Osborne, M.T., Radfar, A., Hassan, M.Z.O., Abohashem, S., Oberfeld, B., Patrich, T., Tung, B., Wang, Y., Ishai, A., Scott, J.A., Shin, L.M., Fayad, Z.A., Koenen, K.C., Rajagopalan, S., Pitman, R.K., & Tawakol, A. (2020). A neurobiological mechanismlinking transportation noise to cardiovascular disease in humans. *European Heart Journal, 41*(6), 772-782. https://doi.org/10.1093/eurheartj/ehz820.

Pascale, A., Fernandes, P., Guarnaccia, C., & Coelho, M.C. (2021). A study on vehicle noise emission modelling: correlation with air pollutant emissions, impact of kinematic variables and critical hotspots. *Science of the Total Environment*, *787*, 147647. https://doi.org/10.1016/j.scitotenv.2021.147647.

Patel, R., Kumar Singh, P., & Saw, S. (2022). Recent advancements in the challenges and strategies of globally used traffic noise prediction models. *Environmental Science and Pollution Research, 29*(32), 48168-48184. https://doi.org/10.1007/s11356-022-20693-1.

Peng, J., Parnell, J., & Kessissoglou, N. (2019). A six-category heavy vehicle noise emission model in free-flowing condition. *Applied Acoustics, 143,* 211-221. https://doi.org/10.1016/j.apacoust.2018.08.030.

Ranpise, R.B., Tandel, B.N., & Singh, V.A. (2021). Development of traffic noise prediction model for major arterial roads of tier-II city of India (Surat) using artificial neural network. *Noise Mapping, 8*(1), 172-184. https://doi.org/10.1515/noise-2021-0013.

Road Transport Year Book, R.T. (2019). *Road Transport Year Book*. https://morth.nic.in/sites/default/files/RTYB_Publication_2019_20 (1).pdf

Singh, D., Nigam, S.P., Agrawal, V.P., & Kumar, M. (2016). Vehicular traffic noise prediction using soft computing approach. *Journal of Environmental Management, 183*(1), 59-66. https://doi.org/10.1016/j.jenvman.2016.08.053.

Singh, D., Prakash, A., Srivastava, A.K., Kumar, K., & Jain, V.K. (2013). The effects of meteorological parameters in ambient noise modelling studies in Delhi. *Environmental Monitoring and Assessment, 185*(2), 1873-1882. https://doi.org/10.1007/s10661-012-2674-1.

Steinbach, L., & Altinsoy, M.E. (2019). Prediction of annoyance evaluations of electric vehicle noise by using artificial neural networks. *Applied Acoustics, 145*, 149-158. https://doi.org/10.1016/j.apacoust.2018.09.024.

Thakre, C., Laxmi, V., Vijay, R., Killedar, D.J., & Kumar, R. (2020). Traffic noise prediction model of an Indian road: an increased scenario of vehicles and honking. *Environmental Science and Pollution Research*, *27*(30), 38311-38320. https://doi.org/10.1007/s11356-020-09923-6.

Tiwari, S.K., Kumaraswamidhas, L.A., Gautam, C., & Garg, N. (2022). An auto-encoder based LSTM model for prediction of ambient noise levels. *Applied Acoustics*, *195*, 108849. https://doi.org/10.1016/j.apacoust.2022.108849.

Tiwari, S.K., Kumaraswamidhas, L.A., Kamal, P.M., & Rehman, M.U. (2023). A hybrid deep leaning model for prediction and parametric sensitivity analysis of noise annoyance. *Environmental Science and Pollution Research, 30*(17), 49666-49684. https://doi.org/10.1007/s11356-023-25509-4.

To, W.M., Ip, R.C.W., Lam, G.C.K., & Yau, C.T.H. (2002). A multiple regression model for urban traffic noise in Hong Kong. *The Journal of the Acoustical Society of America, 112*(2), 551-556. https://doi.org/10.1121/1.1494803.

Ul-Saufie, A.Z., Yahaya, A.S., Ramli, N.A., Rosaida, N., & Hamid, H.A. (2013). Future daily PM10 concentrations prediction by combining regression models and feedforward backpropagation models with principle component analysis (PCA). *Atmospheric Environment, 77,* 621-630. https://doi.org/10.1016/j.atmosenv.2013.05.017.

Umar, I.K., Gökçekuş, H., & Nourani, V. (2022). An intelligent soft computing technique for prediction of vehicular traffic noise. *Arabian Journal of Geosciences, 15*, 1571. https://doi.org/10.1007/s12517-022-10858-0.

Umar, I.K., Nourani, V., Gökçekuş, H., & Abba, S.I. (2023). An intelligent hybridized computing technique for the prediction of roadway traffic noise in urban environment. *Soft Computing, 27*(15)*,* 10807-10825. https://doi.org/10.1007/s00500-023-07826-5.

Uyanık, G.K., & Güler, N. (2013). A study on multiple linear regression analysis. *Procedia - Social and Behavioral Sciences, 106,* 234-240. https://doi.org/10.1016/j.sbspro.2013.12.027.

Vijay, R., Chakrabarti, T.A.P.A.N., & Gupta, R. (2018). Characterization of traffic noise and honking assessment of an Indian urban road. *Fluctuation and Noise Letters*, *17*(04), 1850031. https://doi.org/10.1142/S0219477518500311

Yadav, A., Parida, M., & Kumar, B. (2023). Statistical modeling of traffic noise at intersections in a mid-sized city, India. *Noise Mapping*, *10*(1), 20220164. https://doi.org/10.1515/noise-2022-0164.

Yankoty, L.I., Gamache, P., Plante, C., Goudreau, S., Blais, C., Perron, S., Fournier, M., Ragettli, M.S., Fallah-Shorshani, M., Hatzopoulou, M., Liu, Y., & Smargiassi, A. (2021). Manuscript title: Long-term residential exposure to environmental/transportation noise and the incidence of myocardial infarction. *International Journal of Hygiene and Environmental Health, 232*, 113666. https://doi.org/10.1016/j.ijheh.2020.113666.

Zhang, X., Kuehnelt, H., & De Roeck, W. (2021). Traffic noise prediction applying multivariate bi-directional recurrent neural network. *Applied Sciences, 11*(6), 2714. https://doi.org/10.3390/app11062714.

Zhu, Z., Zhu, S., Zheng, Z., & Yang, H. (2019). A generalized Bayesian traffic model. *Transportation Research Part C: Emerging Technologies, 108*, 182-206. https://doi.org/10.1016/j.trc.2019.09.011**.**

Publisher's Note- Ram Arti Publishers remains neutral regarding jurisdictional claims in published maps and institutional affiliations.