# Chapter 3

# Instrumentation and Methodology

## 3.1 Instrument Details

### 3.1.1 Sound level meter

SLM used in this research was Fusion Sound Level Meter of 01dB of ACOEM make. It has a Class 1 microphone in accordance with the International Electrotechnical Commission (IEC) 61672 STANDARD had integrated Global Positioning System whose coordinates determined the precise location of the instrument. It had a dynamic measurement range of 21–139 dB(A). Before each field trip, the SLMs were calibrated using Class 1 acoustic calibrator (Cal31 of 01dB) at a sound level of 93.7 dB with frequency of 1 kHz. for ensuring accuracy of filed noise measurements. Various parameters have been derived from measured sound levels, including $L_{\mathrm{Aeq}}$, $L_{10}$, $L_{90}$. Description of these parameters are given below:

a) Equivalent sound pressure level ($L_{eq}$): The equivalent sound pressure level $L_{eq}$ expresses a fluctuating noise's impact as a constant sound level. This constant level would deliver the identical total acoustic energy as the time-varying noise within a defined time period. . In simpler terms, $L_{eq}$ represents a steady, unchanging sound that holds the same A-weighted acoustic energy as the actual, fluctuating noise measured over the same duration. A weighting is the frequency filter applied to adjust the raw sound pressure level based on human hearing sensitivity. This filter mimics this human hearing response. It reduces the weight of low and high frequencies in the measurement, resulting in a value ($L_{\text{Aeq}}$ ) that better reflects how loud the sound actually seems to us. Mathematical equation for ($L_{\text{Aeq}}$ ) is Equation (1.1)

b) 10th percentile noise level ($L_{10}$) :

The 10th percentile noise level, often denoted as $L_{10}$, is a statistical measure used in noise assessment. It represents the noise level that is exceeded for 10% of the measurement period. In other words, it is the level below which the noise stays for 90% of the time. $L_{10}$ is commonly used to indicate the higher, more intrusive noise levels in an environment, providing insight into the frequency and intensity of louder noise events.

c) 90th percentile noise level ($L_{90}$) :

The 90th percentile noise level, often denoted as $L_{90}$, is a statistical measure used in noise assessment. It represents the noise level that is exceeded for 90% of the measurement period. In other words, it is the level below which the noise stays for only 10% of the time. $L_{90}$ is typically used to indicate the background noise level in an environment, giving an indication of the quieter, more constant noise that is present most of the time.

### 3.1.2 Speed gun

Stalker Lidar XLR speed gun was used which can measure speed up to 481km/h with accuracy of ±1 km/h.

### 3.1.3 Weather station

The WSH-500 portable weather station manufactured by SCHELT Technology was used which is capable of measuring temperature ranging from 30 to 70°C, humidity from 0 to 100% RH, wind speed capability of up to 75 m/s, and pressure ranging between 500 to 1100 hPa.

### 3.1.4 Machine for evaluating roughness utilizing low-cost instrumentation

The roughness of a road's surface is a vital indicator of road condition and a significant factor in determining vehicular emission. A simple roughness measuring device, called MERLIN (Machine for Evaluating Roughness using Low-cost Instrumentation), has been specifically designed for use in developing countries. This device can be used for direct measurement or for calibrating response-type instruments like the vehicle-mounted bump integrator. MERLIN consists of a 1.8-meter-long metal frame with a wheel at the front, a foot at the rear, and a probe positioned midway that rests on the road surface. The probe is connected to a pivoted arm, which is weighted in such a manner that it moves downward either until it contacts the road surface or until it reaches the maximum limit of its travel. The opposite end of the arm carries a pointer that traces its movement over a mounted chart. As the machine is wheeled along the road, the position of the pointer is recorded on the chart at regular intervals, typically after each complete revolution of the wheel. The pictorial view of MERLIN apparatus is shown in Figure 3.1.

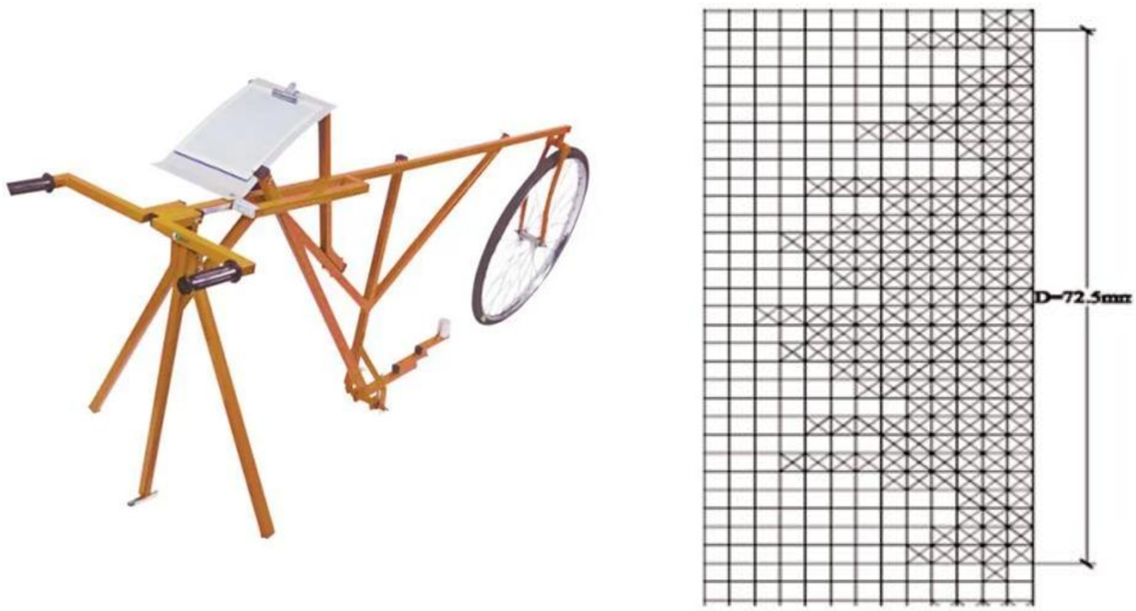



**Figure 3.1.** MERLIN apparatus with tally box

The procedure for measuring road roughness using MERLIN is outlined as follows and followed as per [221] :

1. The MERLIN device is placed at the starting point of the test section, aligned along the direction of travel.
2. The device is steadily wheeled along the road surface. At regular intervals (typically at every full revolution of the wheel), the position of the pointer is automatically recorded on the chart paper, corresponding to the surface profile at that location.
3. A total of 200 readings are collected during the traversal of the road section. A tally box is used to keep track of the number of readings to ensure completeness and accuracy.

4. The MERLIN chart paper is typically divided into grids of 5 mm × 5 mm size, facilitating precise plotting of the displacement values.
5. After all readings are recorded, the top 5% and bottom 5% of the recorded marks (i.e., 10 readings each) are circled off to exclude extreme values that may arise due to anomalies. The remaining 90% of the readings are used to construct a histogram representing the distribution of surface displacements.
6. The width of the histogram denoted as D is measured. This width is indicative of the variability in vertical displacement and thus reflects the roughness of the road surface.
7. The International Roughness Index (IRI) is a standardized parameter used worldwide for quantifying road roughness. The Transport Research Laboratory (TRL) recommended an empirical relationship [221] to compute IRI from the histogram width D for machine-paved roads is shown in Equation (3.1).

$$IRI(m/km) = 0.593 + 0.0471 * D \tag{3.1}$$

### 3.1.5 Camera

Videos of the traffic movement was recorded with the help of EZVIZ HD resolution Wi-Fi camera. Thereafter, bi-directional traffic volume was extracted by manual counting.

## 3.2 Methodology for Heterogeneous Road Traffic Noise Modeling at Mid-block Sections of Mid-sized City in India

### 3.2.1 Geographical location of study area

In India, the Ministry of Education has taken the initiative to address the challenges of Science and Engineering that in future must empower and embolden the nation for selfreliance and inclusive growth. Impacting Research Innovation and Technology (IMPRINT) is the first of its kind with the principal vision to contribute for the domains that are socially relevant. This study is the part of 'IMPRINT' research project whose objective was to study the traffic noise scenario for mid-sized cities that caters to heterogeneous traffic. Gorakhpur mid-sized city (latitude 26°46'N and longitude 83°2'E) in the Indian state of Uttar Pradesh was selected as the study area. Since the past decade, the city has gained remarkable growth in residential, commercial and industrialization sectors with a consequence of increase in traffic noise due to the enhancement of traffic population. In order to collect the traffic noise data, the city landscape was divided into 4 zones i.e. residential, commercial, silence and industrial areas. A total of 23 locations were identified for data collection as shown in Figure 3.2 and detail description is given in Table 3.1.

### 3.2.2 Data collection procedure

Data was collected during November 2018 to March 2022 for an hourly duration to cover the traffic movement in three phases of a day i.e. Morning Peak (MP), intervening Off Peak (OP) and Evening Peak (EP). The respective hourly durations were finalized after numerous trials through local interrogation. Each measuring site may experience peak traffic at differing time spans. For all the data collection sites, the road segment selected was straight, free of speed hump/bump and had a grade of less than 2%. The sites were located in urban environment surrounded by facades of shops and houses/apartments and had bituminous roadway surfacing. The Sound Level Meter (SLM) was installed 1.2m high from the ground, perpendicular to the direction of traffic movement, and at such distance from the centre of the nearby carriageway on the footpath, which ensured safety from traffic and pedestrian movement. This distance, therefore, varied from one study location to the other. Type 1

Fusion SLM of 01dB was used to record the equivalent noise level ($L_{\mathrm{Aeq,1h}}$) and 10-percentile time exceeding level ($L_{10}$). Video recording was done for obtaining the traffic volume, while Stalker lidar speed gun was used for measuring vehicular speeds. Additionally, International Roughness Index, distance of facade from the SLM at same and opposite side of the SLM, average height of building at same and opposite side of SLM. Data collection were done according to the Standard procedure given on IS-10399 1982; IS-3098 1980 [222, 223]. Prior to each field trip, instruments were calibrated and adjusted for accuracy. Cal31 from 01dB is a class 1 acoustic calibrator were used for calibration of SLM at 93.7 dB at 1kHz. Traffic noise dataset of 776 ($L_{\mathrm{Aeq,1h}}$) were collected from the aforementioned 23 locations of Gorakhpur mid-sized city in state Uttar Pradesh in India. After collection of data the extraction of recorded data was done using dB Trait software. dB Trait is a high-performance software program for the post-processing of acoustic, vibration and meteorological data gathered using the following 01dB data collection systems. K-NN algorithm was adopted for traffic noise prediction modeling. Moreover, PCA technique was used for dimensionality reduction and to overcome the problem of multi-collinearity.

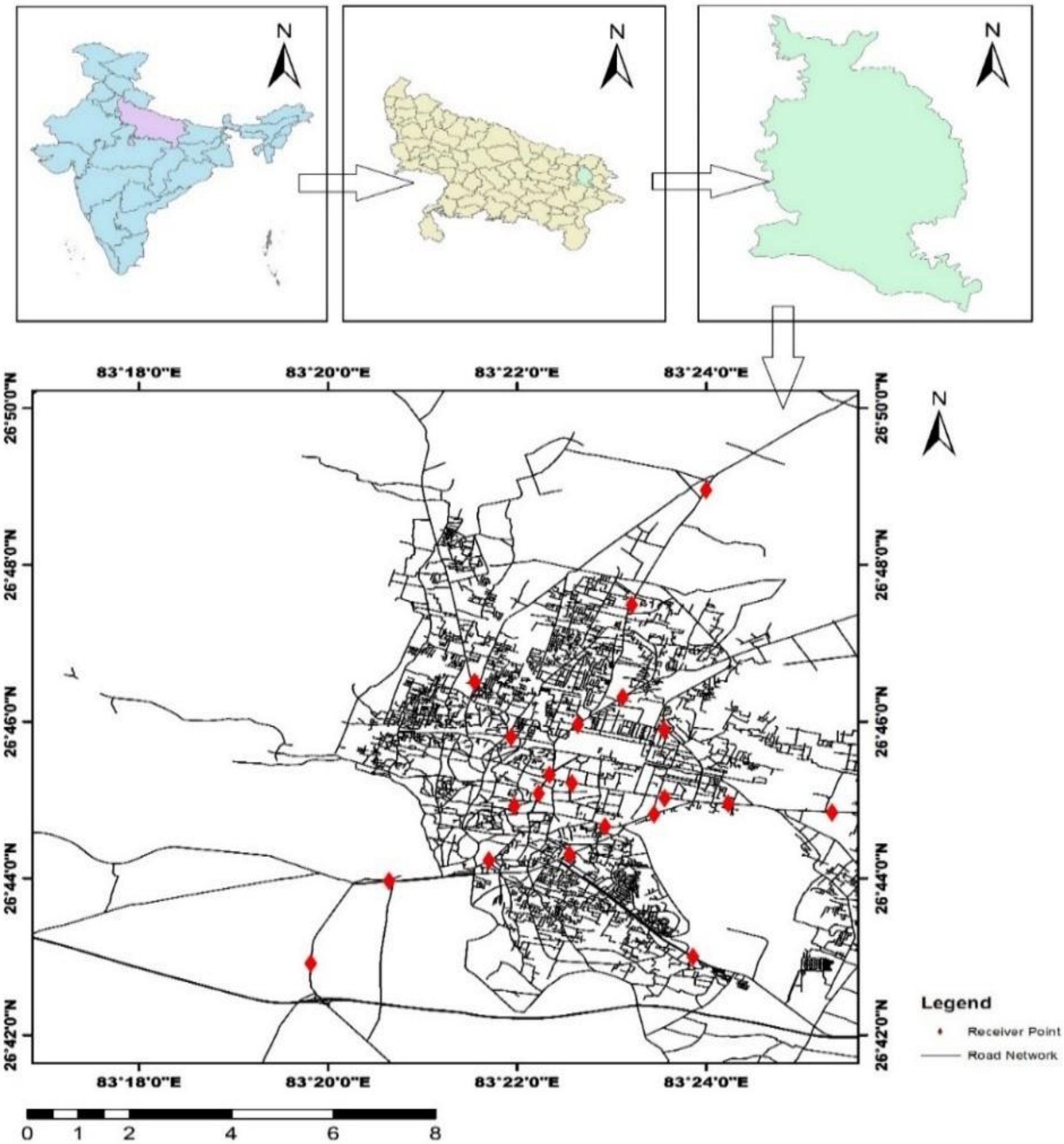



**Figure 3.2.** Geographical location of the study area

**Table 3.1.** Study area

| S.No. | Name of site | Land use | Latitude | Longitude |
|---|---|---|---|---|
| 1 | Arogya Bhawan | Silence | 26°47'29.4"N | 83°23'12.8"E |
| 2 | BRD Medical College | Silence | 26°48'57.2"N | 83°24'00.0"E |
| 3 | AIIMS | Silence | 26°44'50.6"N | 83°25'19.9"E |
| 4 | Mohaddipur-kunraghat | Commercial | 26°44'57.5"N | 83°24'14.0"E |
| 5 | Asuran-Padri | Commercial | 26°46'18.8"N | 83°23'07.1"E |
| 6 | Mohaddipur-Asuran | Residential | 26°45'53.6"N | 83°23'33.4"E |
| 7 | Madan Mohan Malviya University | Silence | 26°43'46.6"N | 83°25'52.7"E |
| 8 | Gorakhpur University | Silence | 26°45'01.5"N | 83°23'33.7"E |
| 9 | Mohaddipur | Commercial | 26°44'49.2"N | 83°23'26.9"E |
| 10 | Paidleyganj | Residential | 26°44'39.7"N | 83°22'55.9"E |
| 11 | Naushad-Rajghat | Commercial | 26°43'58.1"N | 83°20'38.8"E |
| 12 | Naushad-Mau | Commercial | 26°42'55.1"N | 83°19'49.1"E |
| 13 | SBI | Commercial | 26°44'18.3"N | 83°22'33.4"E |
| 14 | Geeda | Industrial | 26°44'36.6"N | 83°16'36.1"E |
| 15 | Taramandal | Residential | 26°43'00.1"N | 83°23'51.7"E |
| 16 | District Hospital | Silence | 26°44'55.3"N | 83°21'58.3"E |
| 17 | Golghar | Commercial | 26°45'05.0"N | 83°22'13.8"E |
| 18 | Kali Mata Mandir Road | Residential | 26°45'48.2"N | 83°21'56.2"E |
| 19 | Assuran | Commercial | 26°45'57.7"N | 83°22'38.6"E |
| 20 | Kalimata intersection | Commercial | 26°45'19.4"N | 83°22'20.6"E |
| 21 | Betiahata | Residential | 26°44'13.9"N | 83°21'42.1"E |
| 22 | Ganesh Chowk | Commercial | 26°45'12.9"N | 83°22'34.9"E |
| 23 | Gorakhnath Mandir | Silence | 26°46'30.4"N | 83°21'33.1"E |

## 3.2.3 Applied machine learning algorithm

### 3.2.3.1 *Principle Component Analysis*

PCA is an unsupervised machine learning technique that aims to minimize the dimensionality of a dataset containing a large number of interrelated variables and retaining maximum amount of variance present in the dataset [224]. PCA is useful in finding the patterns in high-dimensional datasets. It is a projection-based technique which transform the data onto a set of orthogonal axes. Therefore, a new set of variables, known as Principal Components (PCs), with smaller numbers are achieved, which are uncorrelated and are ordered so that first few preserves most of the variation present in the original dataset. To do so, the following steps are generally carried out:

Step 1: Take the whole dataset (without labels) consisting of $d$ dimensions and normalize it.

Step 2: Compute the covariance matrix of the whole dataset.

Step 3: Compute the eigenvectors along with the corresponding eigenvalues.

Step 4: Sort the eigenvectors in the highest to lowest order and select the number of principal components (PCs).

#### 3.2.3.2 *K- Nearest Neighbor (K-NN)*

K-NN is a non-parametric supervised ML algorithm which is based on a fundamental assumption that the predicted value of a particular parameter is more or less similar to its neighborhood value of the past [225]. The algorithm framework consists of three basic elements such as state vector, measuring distance between two state vector and forecasting future state vector through a collection of K-nearest neighbors [226]. This algorithm uses the $k$-number of most similar nearest neighbor from the training dataset [227-230]. The algorithm is also known as a lazy learner algorithm, because, instead of immediate prediction it stores the dataset and categorize it based on their similarities then approaches an action on the dataset. K-NN predicts the new records/data for the regression and classification type problems based on their Euclidean distances, estimated mean, median or model output variable [226, 231-233]. The general formulation of K-NN algorithm has been described in brief in the study of Smith, Williams [234] and Akbari, Overloop [235]. However, the following steps are generally carried out to implement K-NN algorithm:

Step 1: Load the training and testing dataset.

Step 2: Choose the value of K i.e. the nearest data points.

Step 3: Calculate the distance between test data and each row of training data using Euclidean, and Minkowski distance method.

Step 4: Based on the obtained distance value arrange them in ascending order.

Step 5: Select the top K number of rows from the sorted dataset and assign a class to the test point based on most frequent class of these rows.

The K-nearest neighbor (K-NN) algorithm was chosen because of its simplicity, non-parametric nature, and strong ability to model nonlinear relationships without assuming any prior data distribution, which suits the complex and variable nature of traffic noise data. Additionally, K-NN is effective for moderate-sized datasets like ours and offers easy interpretability, which is valuable for developing a practical, real-time noise monitoring system. Preliminary results also indicated that K-NN provided satisfactory prediction accuracy compared to more complex models, justifying its selection. The K-nearest neighbor (K-NN) algorithm selected over traditional regression techniques because it does not assume a linear relationship between the predictors and the target variable, which is important given the complex, nonlinear behaviour of traffic noise data.

### 3.2.4 Method of developing the prediction model

To develop a computer-based software for predicting the traffic noise under heterogeneous traffic condition, the following steps were followed as discussed below and the flowchart for these steps is shown in Figure 3.3.

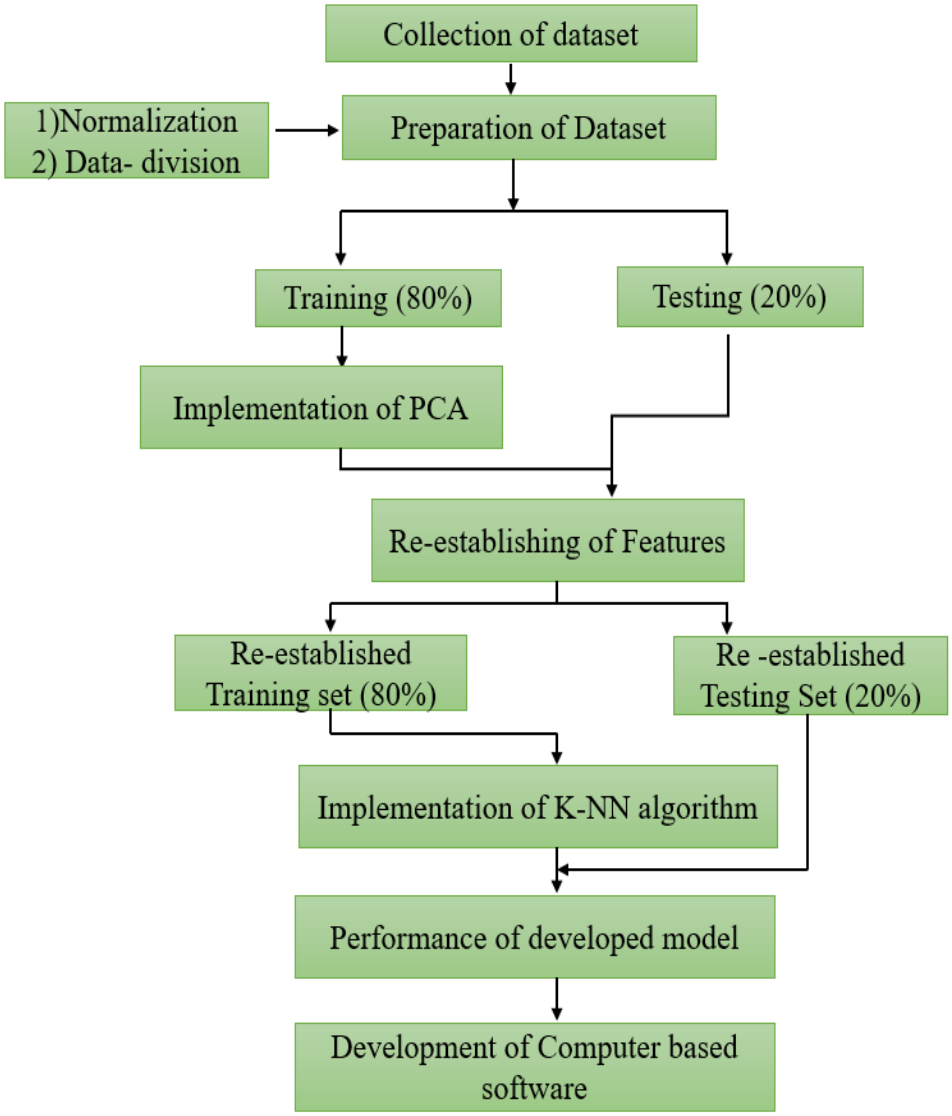



**Figure 3.3.** Flowchart illustrating the complete process of developing the computer-based software for predicting the traffic noise under heterogeneous traffic condition.

1) The collected dataset was analysed and prepared for Modeling purposes. The preparation of dataset was done by normalization. The primary purpose of normalizing the dataset was to improve the computational efficiency, such as decreasing the amount of memory and time required to solve a problem through particular algorithms. Using Equation (3. *2*); all the input and output parameters were normalized between 0 (lower limit) and 1

(upper limit), where $P$ presents the parameters to be normalized; and $P_{max}$ and $P_{min}$ are the maximum and minimum value of the corresponding parameters, respectively.

$$P_{norm} = \frac{(P - P_{min})}{(P_{max} - P_{min})} \tag{3. 2}$$

2) Entire MP, OP and EP dataset was divided into training (TR) and testing (TS) set through K-Fold data division approach. Using five number of folds, 80% of total dataset for each peak was taken for TR purposes and the rest 20% was considered for TS purposes.
3) The PCA technique was implemented for the removal of overfitting and multi-collinearity from the dataset.
4) The K-NN algorithm was employed for developing the traffic noise prediction model. The algorithm hyperparameters tuning were done through grid search technique and the optimized hyperparameters were selected as shown in Table 3.2.
5) The best-fitted model was identified using numerous performance measurement parameters such as coefficient of determination ($R^2$), coefficient of correlation (R), mean absoluter error (MAE), root mean square error (RMSE) which are formulated through Equation (3.3) to (3.6).

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i(a) - y_i(p))^2}{\sum_{i=1}^{N}\left(y_i(a) - \overline{y_i(a)}\right)^2} \tag{3.3}$$

$$R = \frac{\sum_{i=1}^{N}\left(\left(y_i(a) - \overline{y_i(a)}\right)\left(y_i(p) - \overline{y_i(p)}\right)\right)}{\sqrt{\left(y_i(a) - \overline{y_i(a)}\right)^2\left(y_i(p) - \overline{y_i(p)}\right)^2}} \tag{3.4}$$

$$MAE = \left[\frac{1}{N}\sum_{i=1}^{N}|y_i(a) - y_i(p)|\right] \quad (3.5)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i(a) - y_i(p))^2} \quad (3.6)$$

where, $y_i(a)$ = actual value (laboratory obtained value); $y_i(p)$ = predicted value (value obtained through the developed model); $\overline{y_\iota(a)}$= mean of actual value; $\overline{y_\iota(p)}$= mean of predicted value;

6) Finally, based on the results of proposed model, a computer-based software was created through Graphical User Interface (GUI) technique.

**Table 3.2.** Optimial combination of K-NN algorithm hyperparameters for different peak of $L_{eq}$ and $L_{10}$

| **Hyperparameters** | $\boldsymbol{L_{eq}}$ | | | $\boldsymbol{L_{10}}$ | | |
|---|---|---|---|---|---|---|
| | **MP** | **OP** | **EP** | **MP** | **OP** | **EP** |
| K neighbor | 3 | 6 | 3 | 3 | 6 | 3 |
| Leaf size | 30 | | | | | |
| Distance | Minkowski | | | | | |

## 3.3 Methodology for Vehicular Source Emission Model and Modified Highway Noise Model

This study is based on FHWA model which includes development of vehicular source emission model also called as a REMEL model. FHWA has identified REMEL as a function of speed and vehicle category [199]. The REMEL model is also termed as vehicular source emission model. For development of REMEL model the noise emission data of single vehicle is required. The FHWA TNM, like other models in its category, predicts noise levels by adjusting a base sound level, known as the energy mean emission level. This base level represents the average sound energy produced by vehicles. The model then incorporates various factors to refine the noise estimate. These factors include traffic characteristics, such as the type and volume of vehicles, as well as the speed at which they travel. Additionally, adjustments are made for the distance from the roadway to the point where noise is being measured, accounting for the reduction in sound intensity over distance. The model also considers the geometry of the roadway, including the length of the road segment being analyzed, as well as the presence of any barriers or obstructions that could shield or reflect sound. These modifications create a more accurate prediction of the environmental noise impact in specific areas. So, for this research the data collection was done in two part i.e. a) Noise level data collection of single vehicles for developing emission curve for various vehicle type b) Highway traffic noise data was collected from a National Highway (NH) in real running condition. The flow chart of model development is shown in Figure 3.4 and detailed methodology is illustrated in further section.

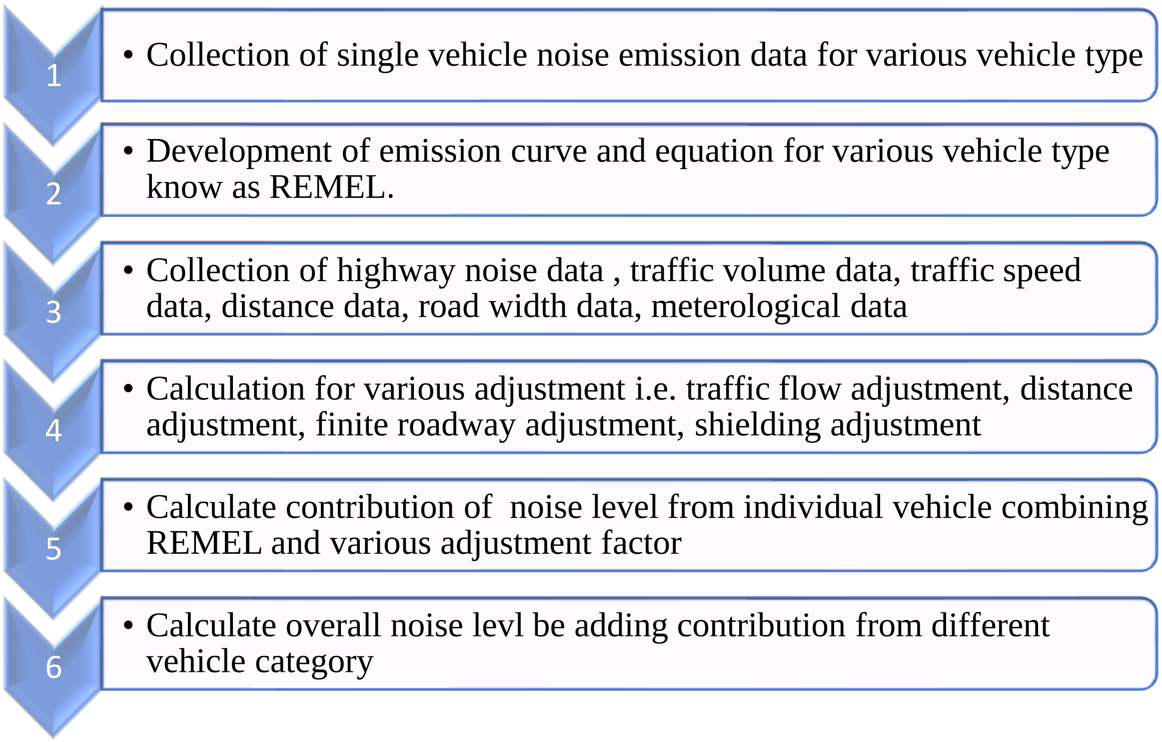



**Figure 3.4.** Flowchart for modified traffic noise model

### 3.3.1 Research site

Noise emission data were collected from two Indian cities—Gorakhpur and Varanasi—at a total of five distinct locations. Each site was carefully selected to minimize the influence of non-traffic noise sources to ensure the accuracy of the measured data. To ensure accurate noise emission measurements for individual vehicles, a site has been carefully selected, distant from prominent sound sources such as busy highways, construction sites, train depots, and airports. The chosen location is strategically positioned away from intersections, lane merges, and other factors that could cause vehicles to accelerate or decelerate. Considering these precautions, the data collection sites were selected as below:

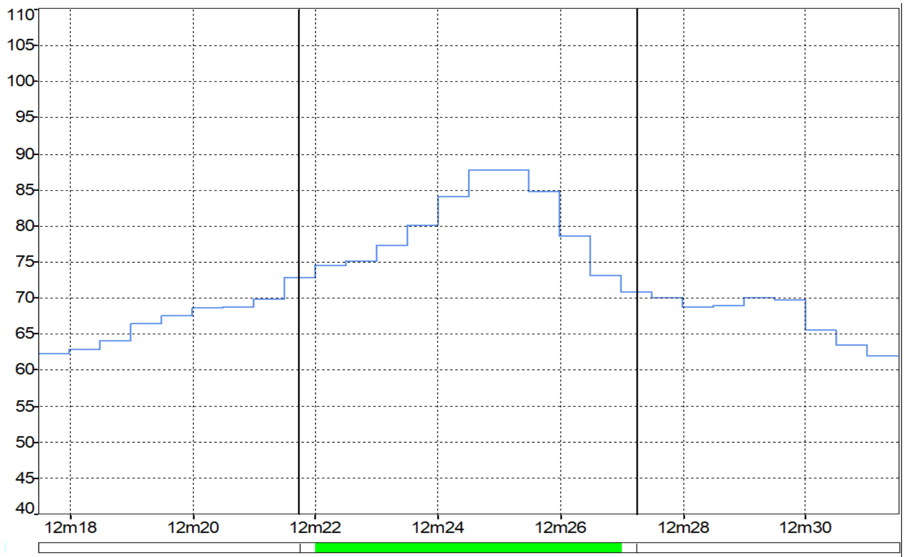
Varanasi: Varanasi - Pryagraj Road - (25°16'49.3"N 82°54'30.4"E); Semi-circle road BHU (25°15'37.1"N 82°59'38.4"E), NH19 (25°14'49.0"N 82°58'47.2"E)

Gorakhpur: NH 727A (26°40'02.0"N 83°32'04.2"E); NH 28 (26°41'40.9"N 83°28'08.0" E).

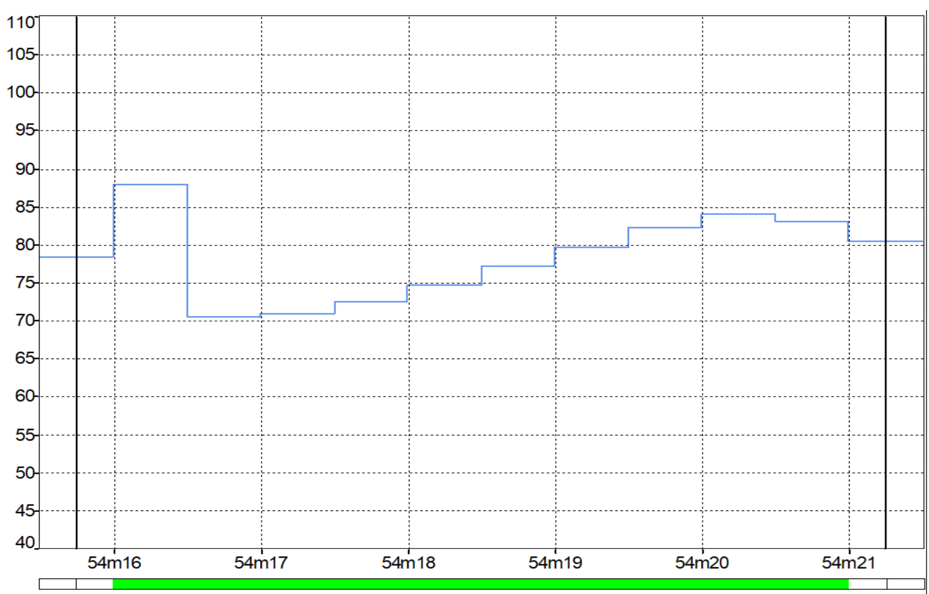
### 3.3.2 Methodology for data collection

The investigation was conducted in accordance with the procedures outlined by the FHWA [39]. At the research site, the road segment ran in a straight line. The vehicle path, or roadway lane, consisted of a dense bituminous mix without debris like gravel. The road maintained a gradient of less than 2%. The measurement points were unobstructed and lacked substantial reflective surfaces. The ground covering comprised paved shoulders with predominantly short grass positioned away from the highway across all sites. The SLM was positioned 7.5 meters from the midpoint of the nearby carriageway and elevated 1.2 meters above the carriageway surface, ensuring an unobstructed view of the roadway with a clear arc of at least 150 degrees for the microphone. Upon the vehicle's approach, the first person activated the recording button on the SLM. The recording ceased after the vehicle passed the SLM and travelled approximately 150-200 meters beyond it; following this procedure, the data on noise emissions from the vehicle were collected. The maximum A weighted sound pressure level with fast response time weighting characteristics ($L_{AF,\max}$) data were recorded for each of the vehicle. Careful attention was given during the data collection process to ensure that extraneous noise sources did not substantially influence the measurements. Specifically, it was ensured that when a vehicle approached a sound level meter, no other vehicles were within a 200-meter radius to the left or right of the SLM. Only these isolated vehicles were considered for the study on vehicle noise emissions. It was also ensured that background noise levels remained at least 10 dB(A) lower than the highest observed value to limit contamination to less than 0.5 dB(A) [7, 9]. In this study, Roadway vehicles were grouped into eight acoustically significant types, i.e. 2-wheelers (2-w), 3-wheelers (3-w), e - rickshaws, Cars (small car, medium car & sport utility vehicle (SUV)), Light Commercial Vehicles (LCV), Trucks, Buses and Tractors.

Approximately 13,684 datasets on noise emissions from individual vehicle passes have been compiled. Aside from emission data, details such as a vehicle's velocity, classification, license plate number, and the count of axles on trucks were also taken. Noise measurements were conducted under fair weather conditions, ensuring that the wind speed remained below 4m/s. The green signal showing in Figure 3.5 is the recording of vehicle's noise as it passes the SLM. The data quality was checked by judging accurate peak, i.e., the first sound level was increased as the vehicle approached the instrument, maximum at the centre and then declined as it drove away. The noise of a single passing vehicle, captured by the SLM during measurement, may also be audible. The signal with only one peak was acceptable for analysis, while the signal with more than one peak was disregarded due to sound contamination; in this way, data quality was checked.




**a) Accepted single pass by event**




**b) Denied single pass by event**

**Figure 3.5.** Selection of corrected noise signal

### 3.3.3 Methodology for model development

A vehicular source noise emission model for various vehicle categories was developed using regression analysis on a comprehensive dataset of noise levels and spot speed measurements. Noise emission data were collected for different vehicle types, including passenger cars, light commercial vehicles, trucks, and motorcycles, under real-world conditions. To ensure data reliability and accuracy, a thorough quality assessment was conducted. The following procedures were adopted for the modeling process.

a) Data categorization and pre-processing:

1. The collected data was categorized into 5 km/h speed bands for each vehicle category. For example, vehicles traveling at speeds between 20 and 25 km/h were grouped together in a single speed band.
2. This banding approach allows for a more granular analysis of how noise emissions change with speed, capturing subtle variations within the data that might be overlooked if broader categories were used.
3. Within each speed band, two key statistical measures were calculated:
    - 50th Percentile Noise Level (Median Value): This value represents the middle point of the noise emission data within each speed band, providing a robust measure that is less affected by outliers compared to the mean.
    - Mean Vehicle Speed: The average speed of vehicles within each speed band was calculated to represent the central tendency of speed for that group.

b) Regression analysis and model development:

1. The median noise levels (50th percentile values) were plotted against the corresponding mean vehicle speeds for each speed band.
2. For each vehicle category, a regression model was developed to describe the relationship between noise emissions and vehicle speed. This involved fitting a mathematical equation to the data points, typically in the form of a linear or logarithmic regression model, depending on the characteristics of the data.
3. The resulting regression equations represent the noise emission curve for each vehicle category, allowing for predictions of noise levels at different speeds.

c) Model validation and statistical significance:

1. The correlation coefficient for each regression model was calculated to quantify the strength of the relationship between noise emissions and vehicle speed. The Pearson correlation coefficient ($r_{xy}$) for n number of paired data can be found in the equation [236] given below:

$$r_{xy} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum {x_i}^2 - \left(\sum {x_i}^2\right)\sqrt{\sum y^2 - \left(\sum {y_i}^2\right)}}} \tag{3.7}$$

Where n is sample size; $x_i$ is speed and $y_i$ is noise level associated with ith pair

2. To ensure that the models are statistically valid and reliable, the correlation coefficient for each vehicle category must be greater than the critical value at a 95% confidence level ($p < 0.05$). This critical value is derived from statistical tables and depends on the number of data points and degrees of

freedom in the analysis. The critical value of correlation coefficient $r_{min}$ is given by following equation [236] given below:

$$r_{min} = \frac{t_{\alpha,n}}{\sqrt{m - 2 + t^2{}_{\alpha,n}}} \tag{3.8}$$

3. A high correlation coefficient (close to 1) indicates a strong linear relationship between speed and noise emissions, suggesting that the model accurately captures the underlying trend in the data.
4. If the correlation coefficient is lower than the critical value, it implies that the relationship is weak or non-existent, and the model may not be a good fit for the data.

This approach results in a detailed, statistically reliable model that illustrates how noise emissions change as vehicle speed varies for each category of vehicles. The noise emission curve is an essential tool for predicting traffic noise levels under different traffic conditions, making it highly useful for noise impact assessments and urban planning.

### 3.3.4 Methodology for collection of highway traffic noise data

The National Highway were selected for highway traffic noise data collection during the period January 2022 to March 2022 at three sites in Varanasi: NH19 (25°14'57.4"N 82°56'41.9"E), NH19 (25°15'07.6"N 83°04'15.5"E), and NH24 (25°15'13.9"N 83°21'37.6"E). These sites were selected for model development.  for model development. Additionally, data from a fourth site in Gorakhpur, NH28 (26°41'49.7"N 83°27'00.8"E), were used for model validation. A total of 106 hours of data were collected for model development, while 36 hours were gathered for model validation. For all the data collection sites, the road segment was straight, free of speed hump/bump and had a grade of less than 2%. The SLM was installed 1.2m high from the ground, perpendicular to the direction of traffic movement, and at such distance from the centre of the nearby carriageway on the footpath, which ensured safety from traffic and pedestrian movement. The SLM were placed on both sides of the road. Traffic noise data ($L_{\mathrm{Aeq,1h}}$) were recorded for 1 hour.

### 3.3.5 Methodology for modified highway traffic noise model development

A modified FHWA traffic noise prediction model has been developed by replacing FHWA REMEL curve with emission curve for different vehicle category developed in present research based on Indian noise emission regulation standards, current vehicle technology, Indian pavement type. After that various adjustment factor added to REMEL model provided by FHWA i.e. traffic flow adjustment, distance adjustment etc. The resulting equation for traffic noise prediction is given below:

$$L_{eq}(h_i) = \overline{(L_0)}_{E_i} + traffic\ flow\ adjustment + Distance\ adjustment + finite\ roadway\ adjustment + shielding\ adjustment$$

$$L_{eq}(h_i) = \overline{(L_0)}_{E_i} + 10\log\left(\frac{N\pi D_0}{ST}\right) - 25 + 10\log\left(\frac{D_0}{D}\right)^{1+\alpha} + 10\ \log\left(\frac{\varphi_a\ (\emptyset_1, \emptyset_2)}{\pi}\right) + \Delta_s \tag{3.9}$$

$L_{eq}(h_i)$ stands for the equivalent sound level if *i*th class of vehicle, $\overline{(L_0)}_{E_i}$ stands for the reference energy mean emission level of *i*th class of vehicle, N represents the number of vehicle passing a specific point within a 1 hour period, D represents the perpendicular distance, in meters, from the centre line of the traffic lane to the receiver, $D_0$ is the reference distance, $\alpha$ represents the site parameter depends upon site conditions, S represents the average speed of vehicle, T represents the duration, usually 1 hour over which equivalent sound level is computed. $\emptyset_1 and\ \emptyset_2$ are the angles from the perpendicular that defines the limits of the observers view of a section of the roadway, $\Delta_s$ is the excess attenuation due to barriers, buildings, wood.

In equation given above ; $\overline{(L_0)}_{E_i}$ is REMEL for different vehicle which were replaced with present study emission curve. FHWA REMEL ( $\overline{(L_0)}_{E_i}$ ) equation were shown in Table 2.5 at S/N 7. Adjustment was remained unchanged and used same as provided by the FHWA model. In traffic flow adjustment the reference distance $D_0$ were taken 15m from centre of the nearby carriageway as recommended by FHWA. The perpendicular distance ( D ), in meters, from the centre line of the traffic lane to the receiver changed from location to location to ensured safety of the instrument from traffic and other disturbances. In the research $D$ were varied from 5.5m to 10.2 m. $\alpha$ represents factor for ground surface (ground between the roadway and observer), which is 0.5 for soft and 0 for hard ( site is reflective). In this research the surface between roadway and observer were acoustically hard. Finite roadway adjustment and shielding adjustment were taken zero because road section was straight and there was no barrier between source and receiver. By addition of all vehicle's classes, noise level of each side (near and far ) can be calculated by Equation (3.10) and (3.11) given below:

$$L_{eq}(NE) = 10 \log \left( \sum_{all\ i} 10^{\frac{L_{eq}\ (i)(NE)}{10}} \right) \tag{3.10}$$

$$L_{eq}(FE) = 10 \log \left( \sum_{all\ i} 10^{\frac{L_{eq}\ (i)(FE)}{10}} \right) \tag{3.11}$$

In the equations provided, $L_{eq}(NE)$ and $L_{eq}(FE)$ were equivalent noise level for near end side and far end side of the road. $L_{eq}\ (i)(NE)$ and $L_{eq}\ (i)(FE)$ were equivalent noise level from ith vehicle calculated on inner and far end of the road . The overall noise level was calculated using the logarithmic sum of the noise levels from both the near and far end , as represented by the Equation (3.12).

$$L_{eq}(Total) = 10 \log\left( 10^{\frac{L_{eq}\,(NE)}{10}} + 10^{\frac{L_{eq}\,(FE)}{10}} \right) \tag{3.12}$$

## 3.4 Methodology for Measurement and Prediction of Road Traffic Noise at Different Floor Levels of Buildings in a Mid-size Indian City

### 3.4.1 Geographical location of study area

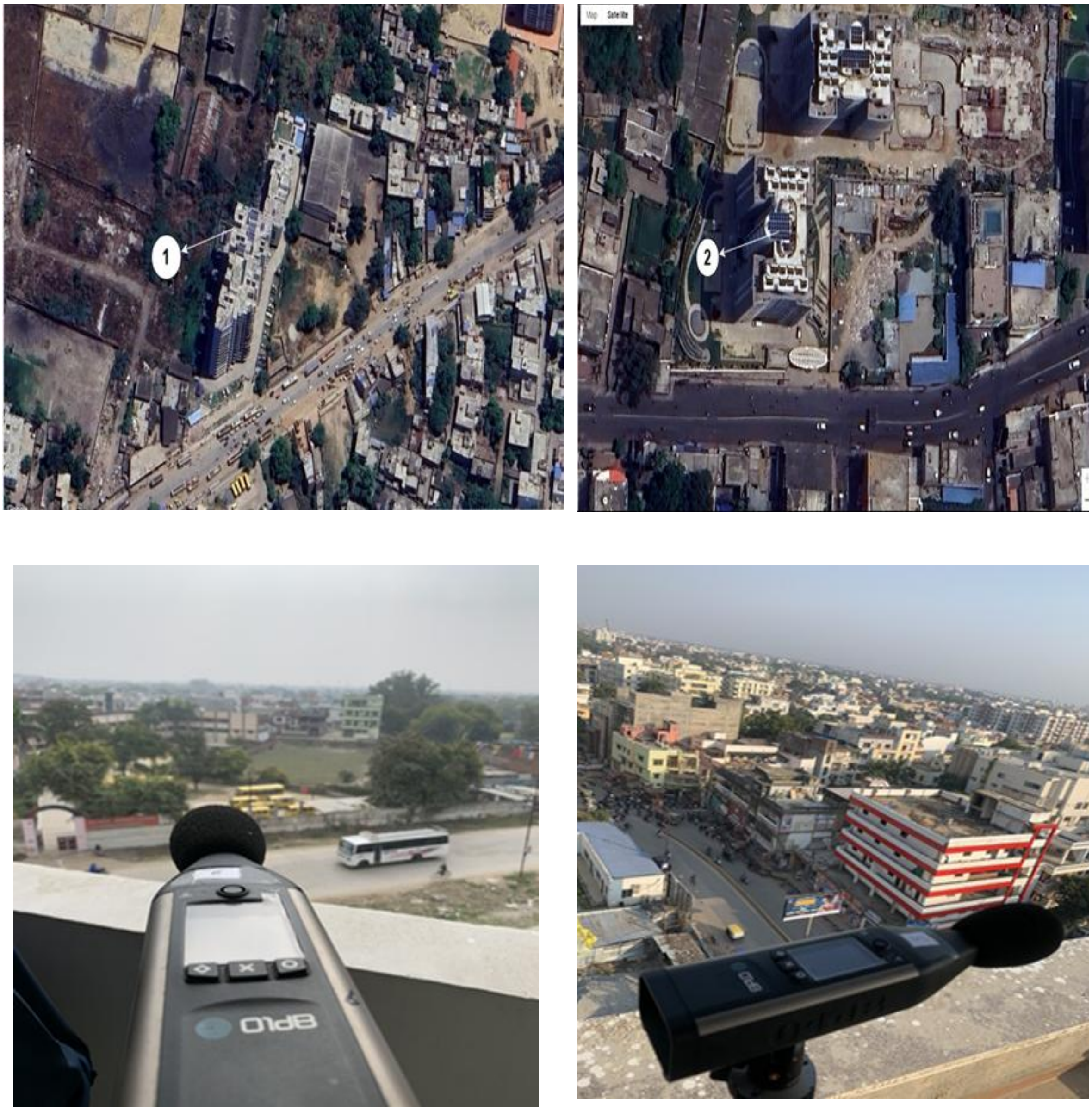
Varanasi beholds the reputation of being the holy city of northern India for various religions of the sub-continent. It is also a major destination for education and health care. Therefore, it has good presence of tourists apart from the locals that makes up most of the city's vehicular traffic. Being a historical city, which has experienced the evolution of human civilization through ages, it has narrow carriageways with the least possibility for widening. Excessive vehicular traffic on these carriageways leads to slower movement of traffic and frequent traffic jams during peak hours. Most commuters were observed to be using personal vehicles instead of public ones. The public transport is sizeably on 3-wheelers which has limited passenger capacity as compared to buses. Buses operate only on few arterial routes. The IC engine-operated 3-wheelers (known to generate high noise levels) are rapidly getting replaced with the electric motor-operated ones. Continuous rows of houses/apartments and shops are built parallel and close to the roads. Due to a lack of convenient parking, drivers often park on the carriage-way, restricting the space for traffic movement. Such a complex and restrictive situation alongside the roadways exposes apartment dwellers and business establishments to the constant hum of traffic. Much of the populace lives in structures that are significantly shorter than 20 metres in height. In this city, buildings taller than 40 metres are uncommon.

### 3.4.2 Case studies

From the perspective of traffic noise pollution, the receiver's distance from the source is a significant factor, as it makes a difference in the receiver's health. Concerns are being raised persistently about the impact of traffic noise on the health of those who reside in high-rise buildings next to roads. In light of this, the noise propagation in high-rise buildings in a medium-sized city road network like Varanasi was investigated. Two case studies were taken up namely, building 1 and building 2 whose details are shown in Figure 3.6 and Table 3.3. Both the buildings were used for residential dwellings.

### 3.4.3 Data collection procedure

Data was collected on regular business days to ensure noise data capturing predominantly due to road traffic operation, which was not possible on weekends and holidays due to community activities in the building that may contribute significantly to the overall noise. The protocol of data collection has been taken from the Indian standard IS 3028:1998, International Organization for Standardization (ISO) 1996-2 and ISO 1996-2016. Equivalent noise levels for the hourly duration ($L_{\mathrm{Aeq,1h}}$) on each floor of both the buildings were measured for 2-cycles of data between 10:00AM and 6:00PM. The sound level metre (SLM) was placed on the balcony on each floor of the building on the façade line, each at 1.2m above the floor [237-240]. This arrangement ensured consistency and comparability of noise measurements. The roadside noise level was measured at a distance of 7.5 m from the centre of the nearby carriageway (adjacent to the instrument) with the SLM being positioned at a height of 1.2 m above the roadway surface. Noise measurements were done during fair weather after ensuring that the wind speed was below 4 m/sec [160] . Time stamp of instances when horns were blown in the traffic were recorded. This was confirmed with the noise signal during data extraction in dB Trait software. The pictorial representation of the data collection site is shown in Figure 3.7.




a) Building 1. Besides Ashapur road

b) Building 2. Besides Mahmoorganj- Sigra road

**Figure 3.6.** Data collection site of traffic noise study for high-rise residential building

**Table 3.3.** Building details

| S. No. | Height (m) of each floor | No. of floors | Building distance (m) from edge of the nearby carriage-way |
|---|---|---|---|
| Building – 1 (25°21'50.2"N 83°01'58.4"E) | 2.8 | 15 | 33.3 |
| Building – 2 (25°18'22.5"N 82°59'10.6"E) | 3.0 | 15 | 18.4 |

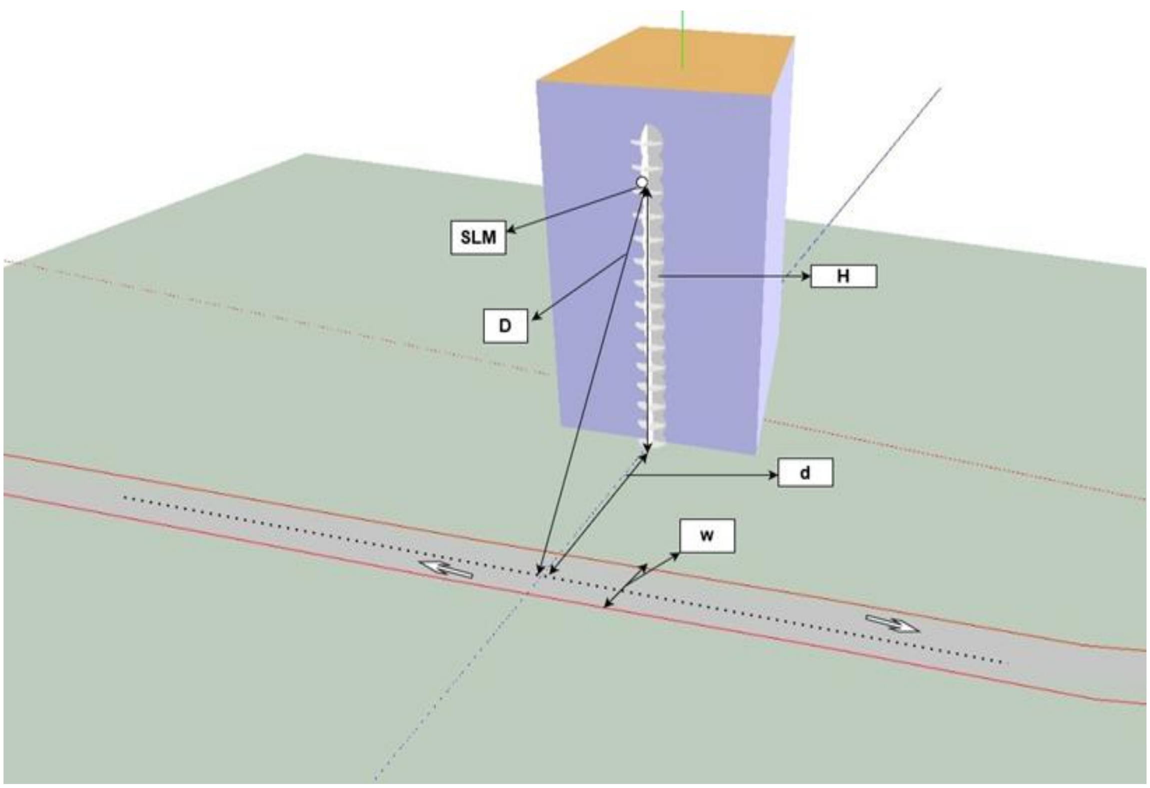



**Figure 3.7**. Pictorial representation of data collection site

The noise source was assumed to be lumped at the centre of the carriageway, which propagates to various floors of the building in which the receivers are present. The length of the propagation line shall be minimum at the ground floor façade (parallel to the ground) and shall increase with floor height. Various parameters have been derived from measured sound levels, including $L_{\mathrm{Aeq}}$, $L_{10}$, $L_{90}$, Traffic Noise Index (TNI), and Noise Climate (NC). The $L_{\mathrm{Aeq}}$ was chosen because many guidelines use this as the parameter for describing noise in legislations i.e. WHO [18] , Indian Central Pollution Control Board (CPCB) [241] etc. To assess acoustic comfort [242] and fluctuations in noise levels [243], TNI and NC were calculated using Equation (3.13) and (3.14) respectively [244].

$$\mathrm{TNI} = [\,4\,(L_{10} - L_{90}) + L_{90} - 30\,]\,, \mathrm{dB(A)} \tag{3.13}$$

$$\mathrm{NC} = (L_{10} - L_{90})\,, \mathrm{dB(A)} \tag{3.14}$$

### 3.4.4 Artificial neural network (ANN)

In artificial intelligence, ANN denotes a subfield influenced by neuroscience and conceived as a human brain simulation [245]. As a result of researchers' efforts to duplicate the human brain, a new discipline known as Deep Learning emerged. The human brain is an intricate network of billions of neurons. Just as the neurons in a human brain are connected in different layers, so are the neurons in ANNs. Nodes refer to these neurons. The three main components of any ANN are the input, hidden, and output layers. Artificial neurons (nodes) are the building blocks of a neural network. All the neurons are interconnected so signals can be sent from the input layer to the output layer via the hidden layer. In an Artificial Neural Network (ANN), the strength of the connections between neurons is represented by synaptic weights, which store the network's knowledge. Information is encoded in these weights and moves from the input layer, through the hidden layer, and finally to the output layer [244]. The ANN model operates based on three fundamental principles: multiplication, summation, and activation. At the ANN's input layer, inputs are multiplied by their assigned weights. These weighted inputs, along with biases, are then summed in the hidden layers. The result of this summation is processed by an activation (transfer) function to produce the output of the ANN. The ANN consists of several key components, including biases, weights, transfer functions, and the training process.

a) Weights

   Each input is multiplied with some weight. Weights are the numbers assigned to each feature or input and represent how crucial that feature is in determining the final prediction. Weights, denoted by $W_i$, represents the information the neural network uses to solve problems. The net inputs are the sum of the product of the weights and input signals.

b) Bias

   A bias unit is an extra neuron added to each pre-output layer, set to hold a constant value of 1. The summation function computes the weighted sum of the inputs; if the weighted total is zero, bias is applied to make the output non-zero, or any other value, to increase the system's reaction.

c) Activation or transfer function

   A transfer function acts as a rule for converting input signals into output signals. These rules can be either linear or non-linear and fall into four main categories: unit step (threshold), sigmoid, piecewise linear, and Gaussian functions the unit step function computes output from net input for single-layer networks. An S-shaped curve represents the Sigmoid function, which are logistic and tangential functions. These logistic functions have ranges from 0 to 1, while the tangential function have ranges from -1 to +1. A piecewise linear function is a type of mathematical function defined by multiple linear segments, each applicable over specific intervals of the input domain. These functions can be either continuous, where the segments connect smoothly without gaps, or discontinuous. Within each segment, the output is directly proportional to the input, with the slope of the segment determining the proportionality constant. Gaussian functions in ANNs are primarily used for their smooth, localized response properties, which are beneficial for various tasks such as pattern recognition, function approximation, and regularization. The symmetry of Gaussian functions is demonstrated by the property $f(-x)=f(x)$.

d) Training and Learning

To achieve a desired output, networks undergo adjustments by fine-tuning the connections (weights) between layers. This process of refinement continues as long as mistakes (errors) decrease. However, training stops when errors begin to rise again, preventing the network from memorizing inaccurate patterns. This internal process of network improvement is referred to as learning.

This study utilized artificial neural networks to develop scientific equations for predicting vertical traffic noise. The model, which included one input layer, one hidden layer, and one output layer, demonstrated success. The number of neurons in the hidden layer was carefully selected; too few neurons might not adequately capture the input signals, while too many could result in overfitting. Since there is no universal algorithm for determining the optimal number of neurons, the study relied on experimental results to identify the ideal number. The framework of the ANN model is an essential factor in the success of a produced model; hence, careful consideration must be given [71]. Many authors [71, 75, 79, 106, 119, 246] have successfully implemented the ANN approach for Modeling traffic noise the standard ANN design (multilayer feed-forward with back-propagation) was used in the present study. The most frequent type of neural network is the Feed-Forward Neural Network  trained using the back-propagation (BP) method because of its ease of use [247, 248]. Any model's ability to predict traffic noise largely depends on the selection of input parameters [53, 133].