Predicting Vehicular SO2 Emissions using Artificial Neural Networks and Mamdani Fuzzy Logic: A Comparative Analysis

Nixon Deb Pritom
Shahjalal University of Science and Technology
Department of Civil and Environmental Engineering
Sylhet, Bangladesh

Tanmoy Goswami
Shahjalal University of Science and Technology
Department of Civil and Environmental Engineering
Sylhet, Bangladesh

ABSTRACT
Air pollution, particularly SO2 emission from vehicular sources due to rapid urbanization and mass migration towards city, poses significant environmental and health risks. This study aims to develop and evaluate predictive models for vehicular SO2 concentration using Artificial Neural Network (ANN) and Mamdani Fuzzy Logic. The hourly SO2 concentration in the study areas exceeds the 0.075 ppm, 1hr-SO2 standard set by the Environmental Protection Agency.

In this investigation, an artificial neural network (ANN) was applied to forecast SO_2 concentrations based on the types and quantities of vehicles using a Multi-Layer Perceptron (MLP) architecture which was improved by backpropagation. A Mamdani Fuzzy Logic inference system was constructed and simulated using MATLAB. Six different vehicles were used as independent variables in this study to predict SO2 levels using trapezoidal and triangular membership function.

The findings reveal that the Mamdani Fuzzy Logic model and the Artificial Neural Network (ANN) model can both accurately estimate the concentrations of SO_2 in vehicles, however, the ANN model performs better and provides more insightful information than the fuzzy logic model due to its human-like reasoning. The correlation coefficient between predicted and observed SO2 concentration in the neural network model is R=0.9294, with a root mean square error of 0.0267, the ANN model performed better than FLA model (RMSE = 0.04895). In the future, it will be necessary to incorporate meteorological variables to create the optimal model.

Keywords
Air Pollution, Sulfur Dioxide, Artificial Neural Networks, Mamdani Fuzzy Logic, Air Quality Prediction

1. INTRODUCTION
With a significant implication for human well-being, air pollution is a drastic environmental health crisis. The World Health Organization [1] report highlighted that 92% of the global population lives in areas where the air quality guidelines are surpassed by air pollution. SO2, as a common type of pollutant from vehicular emissions, predominantly dominates the health burden. Due to an exponential rise in urbanization and population, rapid development of modern industry and transportation, air pollution has turned into a global issue [2]. Poor air quality not only has an impact on people's lives and jobs but also impedes economic development and causes climate change [3]. Long-term exposure to air pollution can also damage the immune, neurological, reproductive, and respiratory systems in addition to increasing the risk of cancer. Under severe circumstances, it may even be fatal [4, 5]. According to a new World Bank research, air pollution kills 15,000 Bangladeshis per year [6]. In Bangladesh's urban areas, unburned gasoline from two stroke motor vehicles is the main cause of air pollution and one of the biggest contributors [7, 8]. Moreover, diesel vehicles comprise roughly 20-22% of all vehicles, and gasoline vehicles make up approximately 80%. Of these 80% gasoline-powered vehicles, about 40% are motorcycles [5]. Two major causes of concern today are rising pollution levels brought on by fast urbanization and rising emissions from motor vehicles [9]. Acid rain is a widespread issue that has several negative environmental repercussions. It is among the worst environmental issues of the modern era. Two of the most important gases that contribute to the formation of acids in the atmosphere are nitrogen oxides and sulfur dioxide. Acid rain has adverse effect on human health, forest and costal ecosystem [10]. Despite the number of death due to air pollution decreased, a 14% of increase in mortality was experienced in developing countries like India between 1990 and 2017.

Approximately 5% of all SO2 emissions are directly caused by transportation (but this percentage can reach 17% in some nations), with diesel fuel creating more SO2 per liter than gasoline [11]. Global SO2 emissions are estimated to be 294 million tons per year. Anthropogenic sources account for 160 million tons of this total [12].

2. LITERATURE REVIEW
In 2015, SO2 concentrations at the Islamic University of Indonesia's Integrated Campus were measured using the Gauss Dispersion Model and the para-rosanilin method. The model revealed that motorcycles emit more SO2 than other vehicles, though new SO2 sources were introduced during the study. Despite this, automobile exhaust contained less than 1% of the SO2 predicted by the model [13]. Motor vehicle emissions along Ilorin Lagos Road pose health risks, with SO2 levels exceeding NESREA limits. The Altair 5X sensor was effective for testing, while the Gliar 3 excelled in ambient measurements. A filter pack system achieved over 95% efficiency in collecting SO2, HNO3, and NH3, making it ideal for daily use. Passive samplers are recommended for long-term monitoring, and maintaining roadside trees is advised [14]. The SO2 results from the monitor were 6% lower than the two-stage sampler, and passive samplers showed slightly higher results but were generally consistent with other methods. Tetrachloromercurate (II) impregnated Whatman No. 17 filter paper has been shown by [15] to be an excellent method of sampling atmospheric sulfur dioxide (SO2) at very low concentrations, down to 0.05 ppbv. The process included pumping air through the treated filters at predetermined flow rates, then utilizing a modified West-Gaeke technique for spectrophotometric measurement. So, Filters coated with lithium hydroxide or oxalic acid and potassium carbonate can collect SO2 with about 100% efficiency [16]. A spectrophotometric method was developed

to measure trace SO2 levels using a reaction with iodine monochloride. While nanomaterials have improved sensor sensitivity and portability, future research should focus on enhancing response times, stability, and reliability through new fabrication techniques. The study also reviews recent advancements in electrochemical SO2 detection [17, 18].

A study developed a meso-scale emission inventory for Dhaka City, incorporating vehicle travel characteristics and fuel types, using both top-down and bottom-up approaches. The study highlighted that congestion and fuel types are major emission contributors. Another study in Hanoi tested the Operational Street Pollution Model (OSPM) by measuring SO2 concentrations on five streets and used the model to estimate emission factors. Motorcycles, comprising over 92% of vehicles, were identified as the main source of pollution [19, 20].

Now, it is pertinent to explore the comprehensive history and evolution of ANNs [21]. ANNs are ideal for complex data modeling, such as environmental pollution assessment, due to their non-linear and self-learning capabilities. Their ability to model the opaque relationship between SO2 concentration and vehicular emissions makes them effective. This study employs a three-layer neural network with a hidden recurrent layer to predict SO2 concentrations [22]. The neural network model, trained with the Levenberg-Marquardt algorithm, outperformed multivariate regression, yielding better predictions with lower error. It effectively modeled SO2 behavior across multiple stations using data from a single location, though no single statistical method is universally optimal for SO2 prediction [23]. However, artificial neural network prediction models are recommended for implementing an air quality control warning system. Whereas a study used a multi-layer perceptron (MLP) artificial neural network (ANN) to predict daily ground-level SO2 concentrations in Istanbul [24], representing the seasonal changes of SO2 and provided that the ANN model provided better results than convolutional deterministic models.

According to the findings, fuzzy logic system uses linguistic variables and rules similar to human reasoning and prediction [25]. Fuzzy logic, while less accurate than neural networks, excels in handling uncertainty and imprecise data. The FLA-based model calculates air quality using SO2, NO2, CO, O3, and PM10 inputs, creating an Air Quality Index (AQI) with categories from 'Good' to 'Hazardous.' This approach offers adaptable, human-like reasoning, effectively capturing non-linear interactions between contaminants [26]. A study demonstrated that using pollutant concentrations of NO2, SO2, and PM2.5, fuzzy logic algorithms can forecast the air quality index (AQI) for Kampala city with accuracy [27]. Using triangle membership functions and Mamdani's fuzzy inference system, they created a fuzzy logic inference system using MATLAB's Fuzzy Logic Toolbox. According to their research, fuzzy logic models can predict AQI with high accuracy and dependability, making them a useful tool for managing and monitoring air quality in real time.

3. MATERIALS AND METHODS

3.1 Study Area and Sampling technique

Sylhet is a developing city of Bangladesh. The number of vehicles on the road is regularly increasing. This research predominantly focuses on the Sylhet-Sunamgonj highway. Three stations on the Sylhet-Sunamgonj highway were selected for this research as shown in the Figure 1, such as Modina Market (Station 1) (24º54′37.1″N, 91º50′53.9″E), Subidbazar (Station 2) (24º54′26″N, 91º51′39.1″E), and Amberkhana (Station 3) (24º54′18.3″N, 91º52′11.7″E). No significant industrial activity was seen in stations at the time of data collection.

Figure 1: Location of the three stations

Figure 2: SENKO SGT-P portable single-gas detector (SO2)

A portable single gas detector (SGT-P) for detecting toxic gases in the air used in the research as shown in Figure 2. The SGT-P is a compact, lightweight monitor designed for hazardous areas. Within one minute of the vehicle's emission, data on SO2 concentration is displayed on the LED screen. It was measured using the diffusion method. Gas sensors are becoming increasingly popular because of their simplicity of use, portability, speed, and consistency in data collecting; gas sensors are becoming increasingly popular; gas sensors are becoming increasingly popular [18].

As this study assumes no emissions from industries, coal burning, or brick kilns, vehicular emissions were the primary source of SO2 in the stations. The National Ambient Air Quality Standard (NAAQS) value of the 1-hr average SO2 concentration is 75 ppb (0.075 ppm) [28]. The 99th percentile of maximum daily SO2 concentrations is 0.075 ppm, and the 1-hour SO2 NAAQS was established as such. As a result, these standards have a chance of 1.00 - 0.99 = 0.01. That means 1 exceedance every 100 days (1/100 = 0.01). The calculation results in approximately (365*3*0.01 ≈11) 11 exceedances over three years when this ratio is projected to the number of days in that period [29–31]. As a result, four data might exceed 0.075 ppm in a year. As a preliminary assessment, this study had a sampling period of 8 hours per day, seven days a week. The sampling took place on the outskirts of the stations. For vehicle counting and concentration sampling, multiple people were present simultaneously. Each station had one week for data collection.

3.2 Data Analysis Procedure

3.2.1 Artificial Neural Network Analysis:

Artificial neural networks, like the Multi-Layer Perceptron (MLP), are used to analyze complex relationships between predictor factors and predicted parameters, such as SO2 concentrations. The MLP typically consists of three layers: input, output, and hidden layers, each made up of interconnected neurons or nodes.

Key Performance Indicators (KPI):

i. Correlation coefficient (R): The strength of a link between two variables is measured using correlation coefficients. This metric assesses the strength and direction of a two-variable liner connection.

ii. Coefficient of determination (R^2): The coefficient of determination describes the amount of variability induced by a factor's relationship to another. The R2 coefficient is sometimes known as the "goodness of fit" coefficient.

iii. Mean Absolute Error (MAE): It is a basic arithmetical average of the difference between the predicted value and the observed value [32]. The formula for calculating this error is shown below-

$$MAE = \left(1/_{N}\right) \sum_{i=1}^{n} \left| C_{p_i} - C_{o_i} \right| \ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

Here, C_{pi}= predicted concentration for the i^{th} `observation.

C_{oi}= observed concentration for the i^{th} observation.

N= total number of observations.

iv. Root Mean Square Error (RMSE): The standard deviation of the projected values from the best-fitted line is computed using residuals or anticipated values, and the RMSE indicates their dispersion. It's commonly used for experimental outcomes or machine learning results [32].

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(C_{p_i} - C_{o_i}\right)^2}{N}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

Here, C_{pi}= predicted concentration for the i^{th} observation.

C_{oi}= observed concentration for the i^{th} observation.

N= total number of observations.

3.2.2 Mamdani Fuzzy Inference System:

In this study, a quantitative method for determining the SO2 concentration is used directly from field stations, which use gas detecting sensors. The primary goal is to establish a link between vehicles and SO2 concentrations. The concentration of SO2 in individual vehicles was also measured, with no interference from other vehicles. A SENKO SGT-P single gas detector was used for sampling SO2. The data was analyzed based on the vehicle number at each station. Specifically, each vehicle's contribution to the natural air's concentration of SO2 was evaluated as a percentage of that vehicle's SO2 emissions. A simulation of fuzzy logic approaches was implemented using the software simulation environment of MATLAB version R2024a. The suggested model was simulated using the MATLAB Fuzzy Logic toolbox. Then, using 6 input variables—the consist of three-wheelers (3W), two-wheelers (2W), cars, microbuses, buses, and trucks— its performance behavior was observed by developing fuzzy prediction model. The output parameter was the estimated concentration of SO2 (ppm). Mamdani fuzzy inference technique was implemented to process the fuzzy logic model [33].

Later, the sets that need to be established within these universes are identified in order to carry out the fuzzification process, and the boundary values of the universal sets for the chosen input and output parameters are established. Determining the input variables, fuzzification, creating fuzzy inference rules, defuzzification, and model evaluation are the design steps that are included in the process.

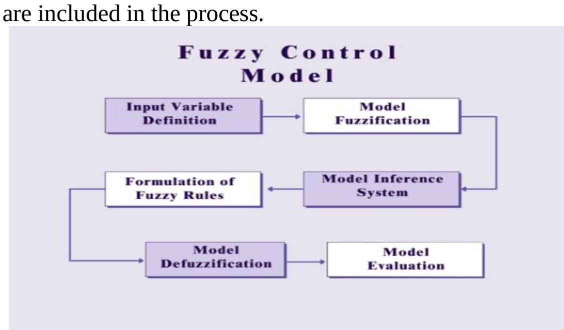

Figure 3: Main components of fuzzy model designed

3.2.3 Membership function selection:

An expression that indicates the amount that an input fits into a set is known as a membership function (MF) [34]. The non-fuzzy input parameters can be converted to fuzzy linguistic terms through membership functions which are used in the fuzzification and defuzzification process. The degree of membership function is limited to 0 and 1. The Mamdani fuzzy toolbox is used to design the fuzzy inference rules [33]. The associated fuzzy membership functions for all inputs are specified as Low, Medium and High, where low and high membership functions are represented as trapezoidal membership function (trapmf) and medium is represented as triangular membership function (trimf) [27].

Table 1: The universe of discourse membership function, domain ranges, and input factor boundary values for crisp and fuzzy sets of data.

Crisp Input Variables	Fuzzy Input membership functions	Boundary Values for Universal sets	Universe of Discourse for MFs
Three-Wheeler (3W)	Low, Medium, High	1400-2400	1400-1891, 1892-2053, 2054-2400
Two-Wheeler (2W)	Low, Medium, High	500-800	500-649. 650-696, 697-800
Car	Low, Medium, High	80-270	80-166, 167-188, 189-270
Microbus	Low, Medium, High	130-180	130-151, 152-158, 159- 173
Bus	Low, Medium, High	0-40	0-13, 14-20, 21-40
Truck	Low, Medium, High	4-60	4-38, 39-45, 46-60

The five fuzzy sets that reflect the chosen output variable—SO2 concentration (ppm)—are Good, Moderate, Unhealthy, Sensitive, Unhealthy, and Dangerous.

Table 2: The universe of discourse membership function, domain ranges, and output factor boundary values for crisp and fuzzy sets of data.

Output Variables	Fuzzy Output membership function	Boundary Values for Universal sets	Universe of Discourse for MFs
SO2 Concentration (ppm)	Good, Moderate, Sensitive, Unhealthy, and Dangerous	0-1600	0-100, 101-500, 501-1100, 1101-1500, 1501-1600

In this step of modelling, the rule base representing the relationship between input variables and output variables formed, called fuzzy inference rule. As there are 6 inputs in the model representing 3 fuzzy membership functions for every input, a total of 729 (36) rules are used in the model. A cross-section from the rules range from scenarios where all vehicle types emit low levels of pollutants, resulting in "Good" air quality with low SO_2 concentrations, to cases where all vehicles emit high levels of pollutants, leading to "Dangerous" levels of

SO_2 concentration. Intermediate rules illustrate the gradual degradation of air quality as emissions from specific vehicle types, particularly buses and trucks, increase. For example, Rule 5 indicates that moderate emissions from buses and medium emissions from trucks result in a "Moderate" SO_2 concentration.

3.2.4 Design of the Input/Output Fuzzy Membership Functions

Figure 4 illustrates an example input/output design of the fuzzy inference system parameters and corresponding membership inputs. In Figure 4, plots for Three-wheeler, Two-wheeler, Car, Microbus, Bus, Truck created using two trapezoidal type membership function as low and high and a triangular type of membership function as medium are displayed. Here, Figure 4(a-f) also demonstrates membership function of the input variables such as Low, Medium, and High, where Figure 4(g) represents membership function plot of an output variable as Good, Moderate, Sensitive, Unhealthy, Dangerous.

3.2.5 Proposed Fuzzy Logic Model for Predicting SO2 Concentration Assessment

Figure 5 illustrates the fuzzy inference rules that were suggested in Table 3 to evaluate the output of the suggested fuzzy-based SO2 concentration prediction model. The fuzzy inference rules were created using a set of fuzzy input membership functions and utilizing the "AND" connection.

Every fuzzy inference rule is regarded to possess a weighted priority function that is equivalent (W = I) in the model evaluation. This indicates that the order is irrelevant because every rule has the same priority while evaluating the model.

4. RESULT AND DISCUSSION

4.1 Rule and Surface Viewer:

The fuzzy inference illustration makes use of the rule viewer. It determines the effect of rules and how membership function affects. In the Mamdani Fuzzy Logic Algorithm based vehicular SO2 concentration presented in the previous chapter, the SO2 concentration in air is actively calculated using the maximum values of input variables. Figure 6, the dual combination of vehicle types used as input variables demonstrated in the surface graphic which represents the changes in SO2 concentration (Three-wheeler, Two-wheeler, Car, Microbus, Bus, and Truck). It is clear and concise that the dual combinations of vehicle types predominantly affect the SO2 concentration. The deterioration of air quality through the increase in SO2 concentration is experienced when there is a increase in the boundary values of the input variables, and good air quality is ensured when there is a decrease in SO2 concentration.

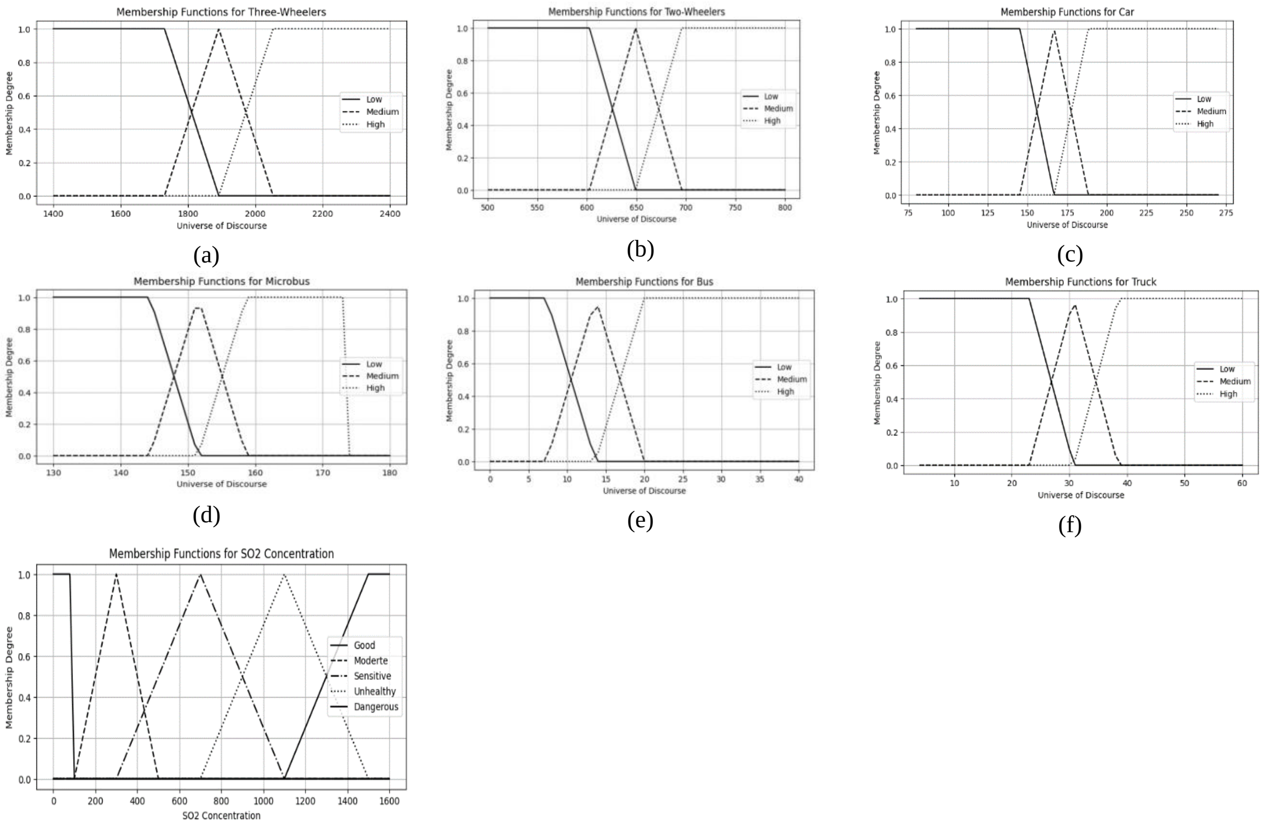

Figure 5: Triangular type membership functions defined for six different vehicle types: (a) for Three-wheeler, (b) for Two-wheeler, (c) for Car, (d) for Microbus, (e) for Bus, (f) for Truck, and (g) for SO2 Concentration.

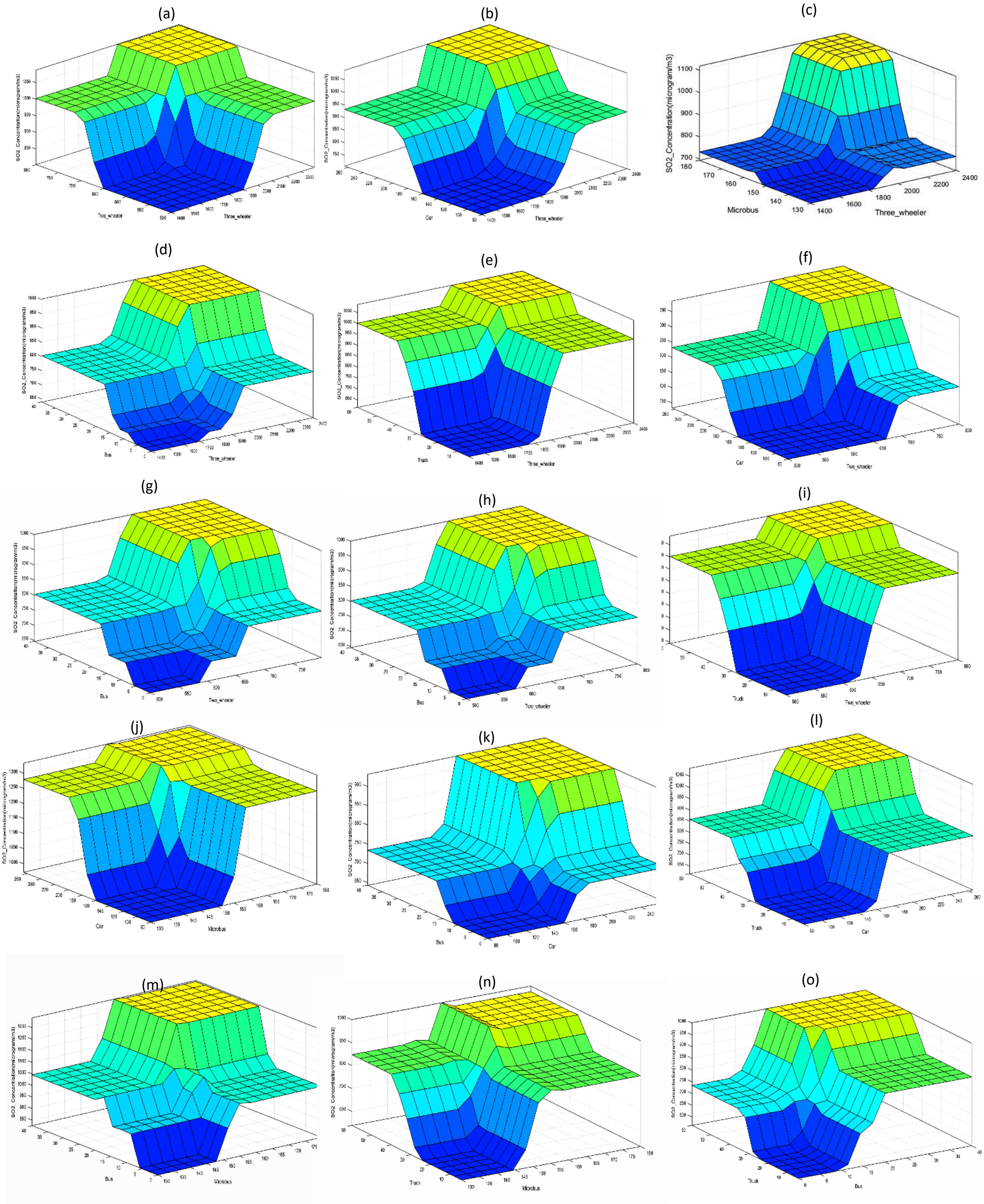

Figure 6: Surface graphics representing the modifications to SO2 concentration (µg/m3) regarding to different vehicle types used as inputs: (a) modifications to SO2 concentration regarding three-wheeler and two-wheeler, (b) modifications to SO2 concentration regarding three-wheeler and car, (c) modifications to SO2 concentration regarding three-wheeler and microbus, (d) modifications to SO2 concentration regarding three-wheeler and bus, (e) changes in SO2 concentration regarding three-wheeler and truck, (f) modifications to SO2 concentration regarding two-wheeler and car, (g) modifications to SO2 concentration regarding two-wheeler and microbus, (h) modifications to SO2 concentration regarding two-wheeler and bus, (i) modifications to SO2 concentration regarding two-wheeler and truck, (j) modifications to SO2 concentration regarding car and microbus, (k) modifications to SO2 concentration regarding car and bus, (l) modifications to SO2 concentration regarding car and truck, (m) modifications to SO2 concentration

regarding microbus and bus, (n) modifications to SO2 concentration regarding microbus and truck, and (o) modifications to SO2 concentration regarding bus and truck.

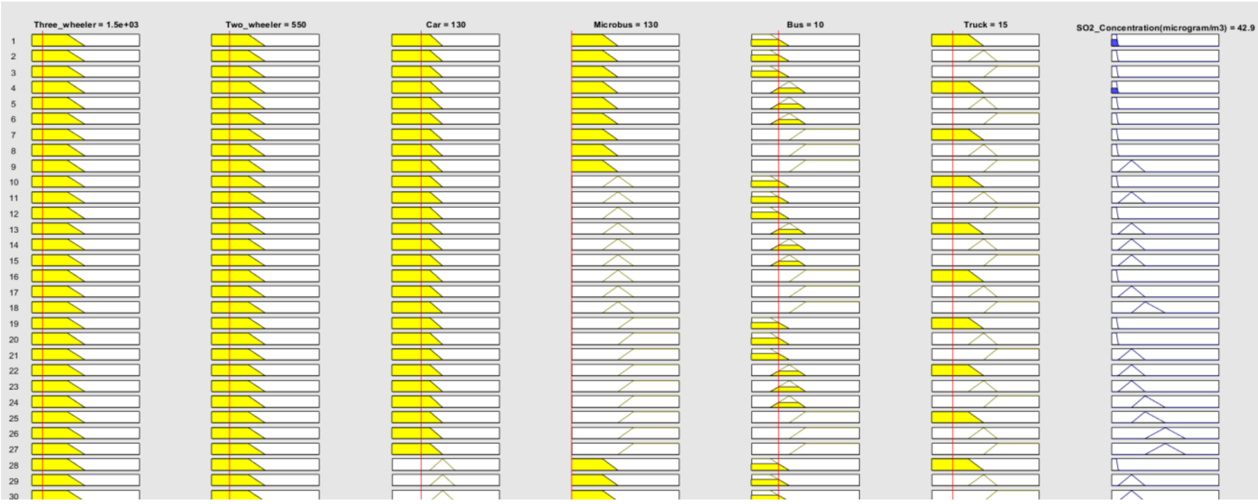

Figure 7: Active rules and Membership functions of input three-wheeler, two-wheeler, car, microbus, bus, and truck vs. output of SO2 concentration (ppm).

Figure 7 illustrates how, in the rule view, varying vehicle quantities have a significant impact on the air's SO2 concentration level. The resulting SO2 concentration, for instance, is shown in the figure given the numbers of the three-wheeler, two-wheeler, automobile, microbus, bus, and truck. For example, if the numbers of the three-wheeler, two-wheeler, car, and truck are 1500, 550, 130, 130, 10, and 15, respectively, the expected SO2 concentration is 42.9 µgm/m3.

Pairwise Relationships Among Inputs and SO2 Concentration:

The pairwise associations between several vehicle types (3W, 2W, Car, Micro Bus, Bus, and Truck) and the SO2 concentration in parts per million are displayed in Figure 8. The relationships between the concentration of SO2 emissions and the various vehicle types are shown visually by the scatter plots and histograms.

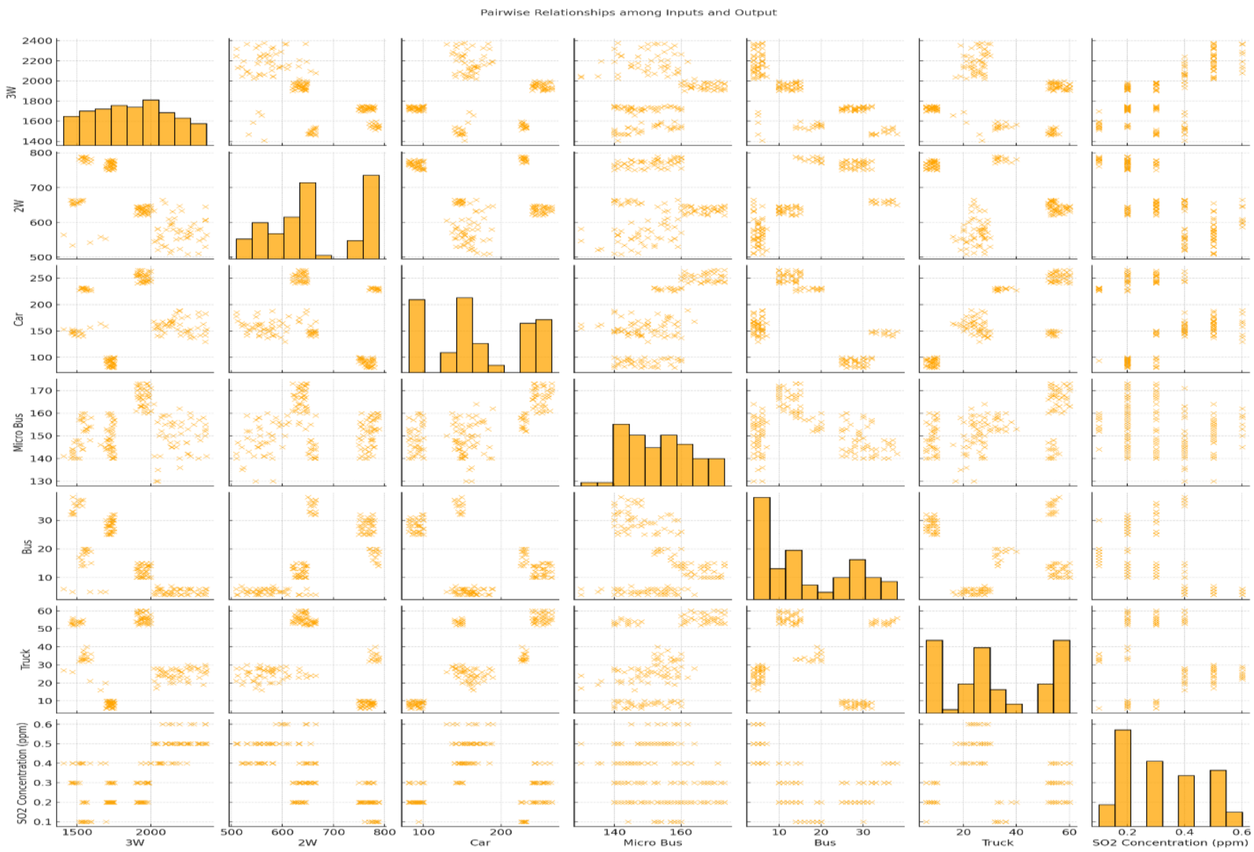

Figure 8: Pairwise relationship among inputs and outputs.

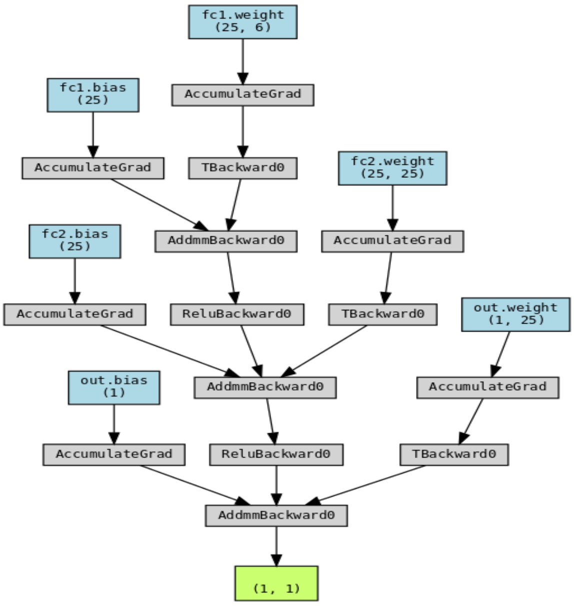

Figure 9: Artificial Neural Network model architecture

We can see from the scatter plots that the distribution pattern of SO2 concentration varies noticeably depending on the kind of vehicle. Some vehicle types, such trucks and 3W, seem to have higher amounts of SO2 emissions than other vehicle types, including cars and 2W. The frequency distribution of each variable, or the number of occurrences within ranges, is displayed by the histograms down the diagonal.

4.2 Defuzzification to Crisp Sets:

To produce a readable result, defuzzifying involves turning a fuzzified output into a single crisp value with respect to the fuzzy set [27]. The process that needs to be controlled is represented by the defuzzied version of the Fuzzy Inference System controller. There are several defuzzification techniques, and they all provide distinct results. The type of problem being solved determines which approach is used.

The concept of the Center of Gravity (COG) technique is the most often used among the other methods, the weighted average method, the mean of maxima approach, and the maximum membership principal approach, which are utilized in this work. In comparison with the other methods, COG produces a more accurate result since it indicates the value that correlates to the center of gravity of the generated curve [26].

To calculate the output crisp value in this work using the COG, assuming “Z” is “C”, then the formula for the expression is formulated as follows, where Z is the final output SO_2 Concentration:

$$Z = \frac{\int \mu_C(Z) Z \, dz}{\int \mu_C(Z) \, dz} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.. (3)$$

Therefore, the defuzzified value of the output for the input values of Three-wheeler = 1400, Two-wheeler = 500, Car = 170, Microbus = 155, Bus = 7 and Truck = 28 is the predicted SO_2 Concentration = 40 µgm/m^3 as shown in Figure 7.

4.3 Artificial Neural Network Analysis:

The backpropagation (BP) training method was utilized in this work because of its simplicity, extensive use, and powerful nonlinear training applications. ‘Pandas’ was used for data handling. The samples were divided into three subgroups using ‘train-test-split’ from ‘scikit-learn’ library based on randomization, with 70% being used for training, 15% for validation, and 15% for testing. To ensure effective training of the model, the input features were normalized using ‘StandardScaler’ from ‘scikit-learn’. In the ANN model the ReLU (Rectified Linear Unit) activation function was used for the hidden layers as it is non-linear, computationally simple and efficient, and it mitigates vanishing gradient problem [35]. The activation function ReLU is defined as f(x)=max(0, x) which means it outputs the inputs directly if it is positive, otherwise it outputs zero.

The neural network was designed using pytorch, characterized by:

- An input layer with 6 neurons representing 6 input features (three-wheeler, two-wheeler, car, micro bus, bus and truck).
- Two hidden layers each with 25 neurons which are activated using ReLU activation function.
- An output layer with a single neuron to predict the S02 concentration.

The Adam optimizer was utilized with learning rate 0.001 to update the model weights and it was chosen for its adaptive learning rate, efficient momentum, handling and effective gradient descent for training neural networks.

As demonstrated in Figure 9 the input data set, consisting of 6 features, was proceeded through the neural network. Initially, the input was weighted in the first layer (with 25 neurons) and transmitted to the second hidden layer (with 25 neurons) and weighted, and finally the output was calculated by conveying through final output layer. Which gave the predicted SO2 concentration.

In the process it performs a forward pass to compute prediction and calculate training loss, then it evaluates validation and test loss to monitor performance and finally backpropagates the loss to adjust model weights. The training, validation and test losses are recorded at each epoch and plotted in Figure 10(a).

The Mean Squared Error (MSE) loss function from ‘torch.nn’ modules was used to measure the model's performance and direct the training procedure. In regression tasks, MSE is frequently used to calculate the average squared difference between the target values and the predicted values.

In this research, an Artificial Neural Network (ANN) was developed to predict vehicular SO2 concentration in the air. The model was implemented using PyTorch on Google Colab.

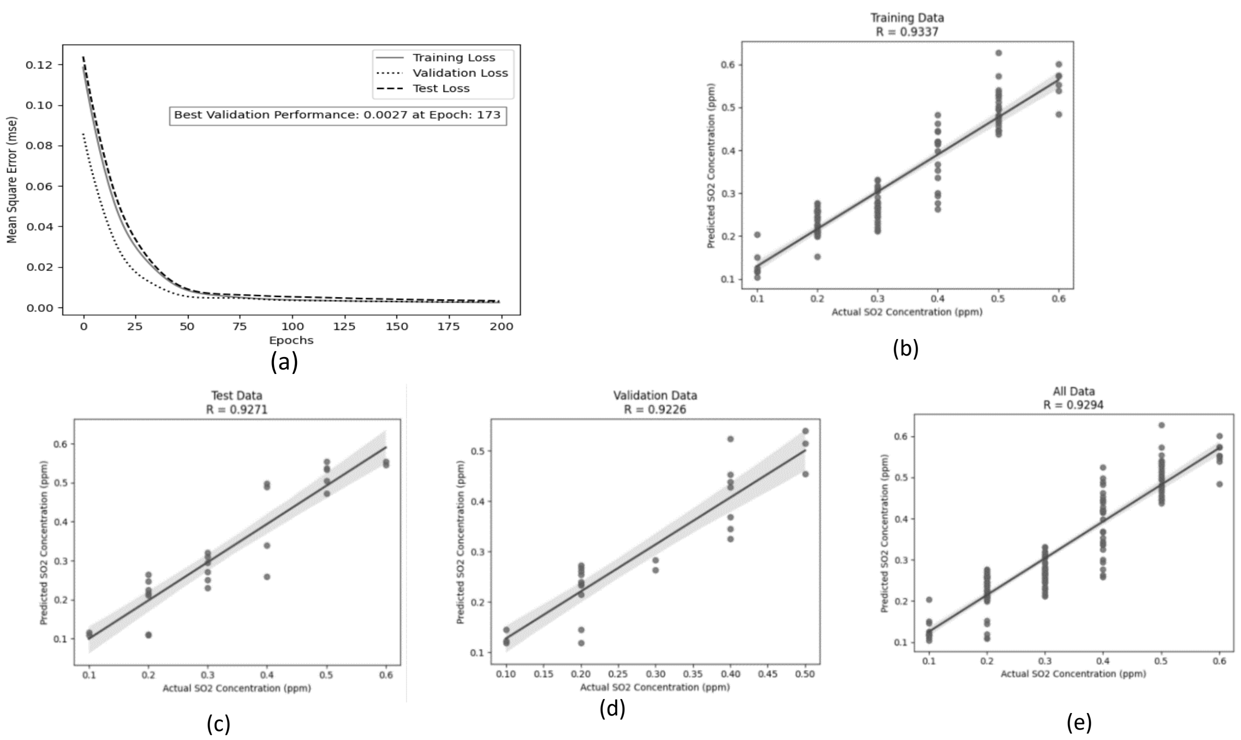

Figure 10: Predicted and observed value of SO2 concentration (ANN) for all stations; (a) error vs. epochs, (b) correlation coefficient of train data, (c) correlation coefficient of test data, (d) correlation coefficient of validation data, and (e) correlation coefficient of all data.

The ANN architecture consists of two hidden layers, each containing ten neurons. This structure was chosen to balance model complexity and computational efficiency [35–37].

$$SO_2(ppm) = \frac{SO_2(\mu g/m^3)\times 24.45}{molecular\ weight\ of\ SO_2 \times 10^3} \ldots\ldots\ldots\ldots\ldots (4)$$

Eqn. 4 is the unit conversion equation which converts SO2 concentration in µg/m3 to ppm, as µg/m3 unit was utilized in FLA membership functions.

It can be seen from Figure 10(a) that the model shows excellent performance with the best validation of .0027 mean square error achieved at 172 epochs. The MSE of both training and validation decreases consistently which represents effective learning and generalization and appropriate hyperparameters (Number of neurons, Layers, Learning rate, optimizer, Batch size, Epochs etc. for the task. Again, the test MSE curve follows a similar pattern which further validates the model's robustness. Moreover, the convergence of training, validation, and test MSE curve implies that neither overfitting nor underfitting is occurring.

The Pearson correlation coefficient (R) is determined to see the accuracy of the ANN model which is presented in Figure 10(b, c, d, e). The correlation between observed and fitted values for the training is 0.9337 which indicated the model learned effectively. Similarly, the higher correlation value 0.9226 of validation data indicates the model has good predictive performance and is capturing a significant portion of the variance in the data. Again, a higher R value of 0.9271 in test data represents model predicted unseen data much more accurately. And finally, all data passed through the model, and it correlates highly between observed and predicted values with R = 0.9294.

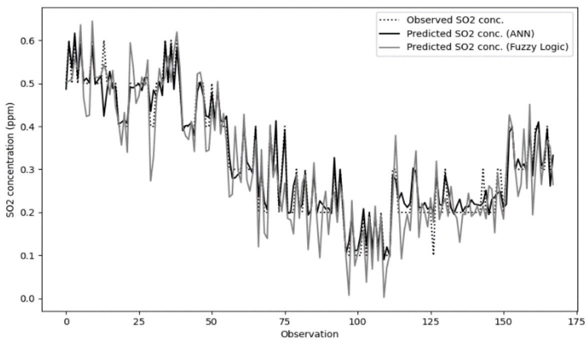

Figure 11: Comparison of Model predicted and observed values of SO2 concentration.

Comparison of ANN and FLA model:

Other parameters such as coefficient of determination (R2), mean absolute error (MAE), root mean square error (RMSE) defined which model fitted the best.

Across three monitoring stations (Madina Market, Subidbazar, and Amberkhana), the performance indicators for forecasting SO2 concentration using Fuzzy logic Algorithm (FLA) and Artificial Neural Network (ANN) performed well from both of its perspectives. The ANN model consistently provided better results for SO2 prediction than FLA based model. The ANN model represents higher correlation coefficient (R) with a value of 0.9294 and coefficient of determination (R2) with

with a value of 0.8637, indicating a stronger fit to the observed data. Moreover, lower value of Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) is achieved by the ANN model with MAE of 0.0175 ppm, RMSE of 0.0267 ppm. These findings highlight the efficiency of advanced machine learning techniques over traditional models in environmental data prediction. Figure 11 shows that the predicted values of SO2 concentration were the most closely related to the observed value. ANN analysis, rather than FLA analysis, revealed that the model performed best in this case.

Table 4: Comparison of KPIs of FLA and ANN

Key Performance Indicator	Madina Market (Station 1)		Subidbazar (Station 2)		Amberkhana (Station 3)		Combined	
	FLA	ANN	FLA	ANN	FLA	ANN	FLA	ANN
R^2	0.6819	0.7816	0.7759	0.9652	0.6475	0.8173	0.7169	0.8637
R	0.8257	0.8841	0.8809	0.9825	0.8047	0.9041	0.8467	0.9294
MAE (ppm)	0.03947	0.01728	0.03943	0.01315	0.03807	0.02361	0.03899	0.0175
RMSE (ppm)	0.04814	0.03035	0.05054	0.01767	0.04813	0.03133	0.04895	0.0267

5. CONCLUSION

This study effectively created and examined predictive models on vehicular SO2 concentration using both Artificial Neural Network (ANN) and Mamdani Fuzzy Logic Algorithm (FLA). This investigation was based on three monitoring stations along the Sylhet-Sunamgonj highway, where the ANN model consistently provided better performance than the FLA based model. However, both models performed accurately on the basis of their algorithm, despite FLA focusing on logical terms. Particularly, all of the key performance indicators such as, coefficient of determination (R2), correlation coefficient (R), Mean Absolute and Root Mean Square errors indicated stronger fit to the observed data for ANN model. Higher correlation coefficient (R) value of 0.9294 and coefficient of determination (R2) value of 0.8637 indicated a stronger where lower Means Absolute error (MAE) value of 0.0175 ppm and Root Mean Square error (RMSE) value of 0.0267 ppm reflected superior accurate.

In comparison with FLA based model for managing and monitoring air quality, the ANN highlights better prediction. The study validates the use of advanced machine learning for future research, such as ANNs in environmental data prediction, offering a promising approach for future research by considering other parameters which influence SO2 and other contamination concentration in air and practical applications in air quality management.

6. REFERENCES

[1] World Health Organization. *WHO Global Air Quality Guidelines. Particulate Matter (PM2.5 and PM10), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide*. (2021).

[2] Zhou, Y., Zhao, X., Lin, K.-P., Wang, C.-H. & Li, L.: A Gaussian process mixture model-based hard-cut iterative learning algorithm for air quality prediction. *Appl. Soft. Comput.* **85**, 105789 (2019).

[3] Kampa, M. & Castanas, E.: Human health effects of air pollution. *Environ. Pollut.* **151**, 362–367 (2008).

[4] Haque, A. H., Huda, N., Zaman Tanu, F., Sultana, N., Shahid Hossain, M. A., & Rahman, M. H.: Ambient air quality scenario in and around Dhaka city of Bangladesh. *Barisal Uni. J. Part* **1**, 203–218 (2017).

[5] Khan, A. S. R.: Diesel vehicles emissions control in Bangladesh. *International Conference on Mechanical Engineering* (Dhaka, 2003).

[6] Ahmed, S. & Mahmood, I.: Air pollution kills 15,000 Bangladeshis each year: The role of public administration and government's integrity. *Afr. J. of Poli. Sci.* **11**, 1–012 (2017).

[7] Alam, G. M. J.: Environmental pollution of Bangladesh - its effect and control. *Pulp and Paper*. 13–17 (2009).

[8] Kojima, M., Bank, W. & Brandon, C.: *Improving Urban Air Quality in South Asia by Reducing Emissions from Two-Stroke Engine Vehicles*. https://www.researchgate.net/publication/242772958 (2014).

[9] Bhanarkar, A., Bhanarkar, A. D., Goyal, S. K., Sivacoumar, R., & Rao, C. V. C.: Assessment of contribution of SO_2 and NO_2 from different sources in Jamshedpur region, India. *Atmos. Environ.* **39**, 7745–7760 (2005).

[10] Ahmed, H. Y.: Management of Acid rain for environmental pollution. (Salauddin University, Erbil, 2021).

[11] Kebin, H. E., Zhang, Q. & Hong, H.: Types and Amounts of Vehicular Emission. *Point Sources of Pollution: Local Effects and its Controls.* **1**, (2009).

[12] Fioletov, V. E., McLinden, C. A., Krotkov, N., Li, C., Joiner, J., Theys, N., Carn, S., and Moran, M. D.: A global catalogue of large SO_2 sources and emissions derived from the Ozone Monitoring Instrument. *Atmos. Chem. Phys.*, 16, 11497–11519, https://doi.org/10.5194/acp-16-11497-2016, 2016.

[13] Supriyanto & Ansharullah, A. W.: Contribution of Sulfur Dioxide Concentration from Vehicle Emissions at Integrated Campus of Islamic University of Indonesia. in *3rd International Conference on Civil, Biological and Environmental Engineering (CBEE-2016) Feb. 4-5, 2016, Bali (Indonesia)* (International Institute of Chemical, Biological & Environmental Engineering, 2016). doi:10.15242/IICBE.C0216028.

[14] Makkonen, U. & Juntto, S.: Field comparison of measurement methods for sulphur dioxide and aerosol sulphate. *Atmos. Environ.* **31**, 983–990 (1997).

[15] Axelrod, H. D. & Hansen, S. G.: Filter sampling method for atmospheric sulfur dioxide at background concentrations. *Anal. Chem.* **47**, 2460–2462 (1975).

[16] Quinn, P. K. & Bates, T. S.: Collection Efficiencies of a Tandem Sampling System for Atmospheric Aerosol Particles and Gaseous Ammonia and Sulfur Dioxide. *Environ. Sci. Technol.* **23**, 736–739 (1989).

[17] Gayathri, N. & Balasubramanian, N.: Spectrophotometric Determination of Sulfur Dioxide in Air, Using Thymol Blue. *J. AOAC Int.* **84**, 1065–1069 (2001).

[18] Khan, M., Rao, M. & Li, Q.: Recent Advances in Electrochemical Sensors for Detecting Toxic Gases: NO_2, SO_2 and H_2S. *Sensors* **19**, 905 (2019).

[19] Iqbal, A., Allan, A. & Zito, R.: Meso-scale on-road vehicle emission inventory approach: a study on Dhaka City of Bangladesh supporting the 'cause-effect' analysis of the transport system. *Environ. Monit. Assess.* **188**, 149 (2016).

[20] Hung, N. T., Ketzel, M., Jensen, S. S. & Oanh, N. T. K.: Air Pollution Modeling at Road Sides Using the Operational Street Pollution Model—A Case Study in Hanoi, Vietnam. *J. Air Waste Manage Assoc.* **60**, 1315–1326 (2010).

[21] Wu, Y. chen & Feng, J. Wen.: Development and Application of Artificial Neural Network. *Wirel. Pers. Commun.* **102**, 1645–1656 (2018).

[22] Chelani, A. B., Rao, C. V. C., Phadke, K. M. & Hasan, M. Z.: Prediction of sulphur dioxide concentration using artificial neural networks. Environ. Model. & Soft. 17, 161–168 (2002).

[23] Nunnari, G., Dorling, S., Schlink, U., Cawley, G., Foxall, R., & Chatterton, T.: Modelling SO2 concentration at a point with statistical approaches. Environ. Model. & Soft. 19, 887–905 (2004).

[24] Saral, A. & Ertürk, F.: Prediction of Ground Level SO2 Concentration using Artificial Neural Networks. Water, Air and Soil Pollut.: Focus 3, 307–316 (2003).

[25] Shah, D., Patel, K. & Shah, M.: Prediction and estimation of solar radiation using artificial neural network (ANN) and fuzzy system: a comprehensive review. Int. J. of Energy and Water Resources 5, 219–233 (2021).

[26] Arici, N. & Atacak, I.: Modelling and Evaluating Air Quality with Fuzzy Logic Algorithm-Ankara-Cebeci Sample. Int. J. of Intel. Sys. and App. in Engg. 4, 263–268 (2017).

[27] Katushabe, C., Kumaran, S. & Masabo, E.: Fuzzy Based Prediction Model for Air Quality Monitoring for Kampala City in East Africa. Appl. Sys. Innov. 4, 44 (2021).

[28] Environmental Protection Agency.: National Ambient Air Quality Standard for Sulfur Dioxide; Final Rule. (2010).

[29] Guerra, S. A., Olsen, S. R. & Anderson, J. J.: Evaluation of the SO 2 and NO X offset ratio method to account for secondary PM 2.5 formation. J. Air Waste Manage Assoc. 64, 265–271 (2014).

[30] Shie, R. H., Yuan, T. H. & Chan, C. C.: Using pollution roses to assess sulfur dioxide impacts in a township downwind of a petrochemical complex. J. Air Waste Manage Assoc. 63, 702–711 (2013).

[31] Wickham, L.: The revised US national standard for sulphur dioxide. J. of Air qual. and clim. change 45, (2011).

[32] Pak, U., Kim, C., Ryu, U., Sok, K. & Pak, S.: A hybrid model based on convolutional neural networks and long short-term memory for ozone concentration prediction. Air Qual. Atmos. Health. 11, 883–895 (2018).

[33] Mamdani, E. H. & Assilian, S.: An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller. Int. J. Man-Machine Studies 7, 1–13 (1975).

[34] Adil, O., Ali, A., Ali, M., Ali, A. Y. & Sumait, B. S.: Comparison between the Effects of Different Types of Membership Functions on Fuzzy Logic Controller Performance. Int. J. of Emerg. Engg. Res. and Tech. 3, 76 (2015).

[35] Safari, A. & Ghavifekr, A. A.: International Stock Index Prediction Using Artificial Neural Network (ANN) and Python Programming. in 2021 7th International Conference on Control, Instrumentation and Automation (ICCIA) 1–7 (IEEE, 2021). doi:10.1109/ICCIA52082.2021.9403580.

[36] Srivastava, S., Sharma, L., Sharma, V., Kumar, A. & Darbari, H.: Prediction of Diabetes Using Artificial Neural Network Approach. in 679–687 (2019). doi:10.1007/978-981-13-1642-5_59.

[37] Gupta, A., Parmar, R., Suri, P. & Kumar, R.: Determining Accuracy Rate of Artificial Intelligence Models using Python and R-Studio. in 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N) 889–894 (IEEE, 2021). doi:10.1109/ICAC3N53548.2021.9725687.