Introduction

In a world of industrialization and modernization, air pollution appears to be on the rise. A recent report IQAir Visual 2018 concluded that Delhi is a highly polluted capital across the world and ranked 11th. India rules the list with 22 of the worst 30 cities globally [1]. Air pollution is an insidious destroyer of the country’s health and wealth. With every inhale, microscopic particles get deep into the lungs. Air pollution is the major cause of cancers, strokes, and heart disease [2]. In recent studies, it is found that pollution is stunting children’s growth and development [3, 4]. The total number of cases is likely to rise in the coming years if pollution is not regulated. Moreover, additional deaths as well as disabilities will be introduced [5]. Air pollution has reduced the average life expectancy to 1.8 years. $225 billion is the estimated financial cost in lost labor and a thousand times more cost in healthcare. Rising pollution and its severity have forced researchers and scientists to carry out various studies involving its health effects and future trends. Air pollution dynamics is a complex process due to randomness involved in air pollutants behavior. Adaptive Neuro-Fuzzy Inference System (ANFIS) is effective in dealing with non-linear real-time problems. ANFIS has been widely used in disaster management, rock engineering [6, 7] healthcare sector, finance, and many other real-time fields [8-10]. ANFIS deals with both regression and classification problems [11].

The large error existing in the air quality modeling is discussed and optimized in this study. The major drawback observed in the past studies is the dependency of the model on various parameters. The complexity involved due to the various parameters leads to a large error. The present study aims to develop a less resource-intensive and more effective model in predicting air quality. ANFIS is used as a regression tool. The learning algorithm of classic ANFIS is based on the gradient descent method which is replaced with the evolutionary algorithms in the present study. The efficiency of the proposed model is validated by applying it to the data monitored at the Shadipur area of India. Shadipur is an industrial cum residential area covering all the precursors for polluting air. Hence, appropriate study area to validate the performance of model. The prominent pollutants - nitrogen oxide, particulate matter of size equal to or less than 2.5microns (PM2.5), and sulphur dioxide. The proposed hybrid model using wavelet transform, particle swarm optimization, and adaptive neuro-fuzzy inference system accurately predicts the Air Quality Index and can be used in the public interest to take necessary precautions beforehand.
to June 2019. The model developed is only dependent on the past values of the respective air pollutant. The study is independent of other variables. The authors have not encountered such work to the best of their knowledge. The article is organized into four sections explaining the problem undertaken and the data collected, methodology, results, and the last section concludes the work carried out.

2. Methodology

2.1. Study Area and Dataset
Central Pollution Control Board (CPCB) is India’s apex body that monitors the air quality. The various stations of the organization serve almost all cities. It monitors contaminants and the parameters of the atmosphere. Delhi is one of the most polluted cities. Hence, a residential, industrial, and commercial region of Delhi covering all the human activities resulting in the formation of pollutants is considered for the study. Moreover, the subtropical climate of Delhi makes the behavior of air pollutants chaotic and contains extreme values covering the broader aspect of the applicability of the proposed models. The daily concentration (24-h) of fine particulate matter (PM$_{2.5}$), oxides of nitrogen, and sulphur dioxide from January 2010–June 2019 was collected from CPCB for the current analysis. Fig. 1 depicts the areas monitored by CPCB in Delhi, India, and the study area-Shadipur.

![Fig. 1. Study Area- Industrial cum residential area, Shadipur, Delhi, INDIA (permission from the International Journal of Environmental Science and Technology Springer Nature).](image)

2.2. Wavelet Transform
In real-time data problems, pre-processing of data plays a significant role in any kind of analysis. The presence of extreme outliers, difficulty in feature extraction, and various fluctuations in data lead to errors in modeling real-time problems. The Pre-processing of data is to solve these issues. Pollutant sequence preliminary analysis indicates sudden shifts and spontaneous variations. Wavelet transform is employed to extract all the characteristics of the underline sequence. For further insights refer to Rashmi and Dimple [12]. In this study Daubechies (Db5) wavelet is considered.

2.3. Adaptive-Neuro Fuzzy Inference System
ANFIS is a fusion of neuronal networks and fuzzy systems. It is a widely used method for regression problems. It is based on on-premise and consequent parameters. In the classic ANFIS model, the parameters are tuned using a gradient descent algorithm as used by various researchers in the past [12-14]. ANFIS is based on targeted data and input. The learning algorithm builds the relationship between input and output in the Fuzzy if-then rules structure. The performance of the model is based on the learning algorithm. The parameters are tuned in the present work using various techniques of optimization-swarm intelligence, genetic algorithm, ant colony, and differential evolution, as discussed below.

2.3.1. Particle swarm optimization
In 1995, Eberhart and Kennedy developed the PSO technique [15]. Compared to other evolutionary algorithms the algorithm requires fewer parameters. The output vector $y$ in R$^d$ in PSO is based on position vector $q$ and velocity vector $w$ of the particle. For each iteration, velocities of all variables are modulated based on inertia weight ($W$), cognitive ($d_1$), and social ($d_2$) acceleration. For $(n+1)^{th}$ epoch, velocity and position are updated as:

$$
\begin{align*}
    w^n_{i,j} &= W w^n_{i,j} + d_1 s^n_{i,j} (q^n_{i,j} - y^n_{i,j}) + d_2 s^n_{i,j} (q^n_{p,best,j} - y^n_{i,j}) \\
    y^n_{i,j} &= y^n_{i,j} + w^n_{i,j}
\end{align*}
$$

where, $s_i, s_j \in U(0,1)$ and $d_1 + d_2 \leq 4$. Consider the particle's best position as pbest. Compare the current position of the particle with that of pbest. If the current position is better than pbest, then pbest is the current position otherwise pbest is the best global position (gbest). The optimal values for tuning ANFIS parameters in this study are taken as $W = 1$, $d_1 = 1$, and $d_2 = 2$ using the trial and error process. The maximum iteration for all the algorithms is taken as 1000.

2.3.2. Genetic algorithm
John Holland developed GA in the 1970s [16], and applied henceforth to solve different types of problems. Using evolutionary biology the genetic algorithm is concerted. The algorithm is built based on descent, mutation, and crossover ideas. The important aspect of GA over the other evolutionary strategies is that it successfully works in the presence of more varied. This is a nature-influenced optimization algorithm. The algorithm operates on a chromosome population. The emphasis is not on one search space single point or one chromosome. Properties displayed by the population at all stages are dependent on the characteristics of the preceding stage. The cost function determines chromosomal output by the added fitness function. De et al. [17] will provide further insights into the genetic algorithm. Mutation percentage and mutation rate are taken as 0.4, 0.7, and 0.15, respectively, using the trial and error process crossover ratio.

2.3.3. Ant colony optimization
ACO is influenced by ants. ACO was developed by Dorigo and colleagues in the early 1990s [17]. From then on, the algorithm is incorporated for various optimization problems. The principle of ACO is based on the mechanism of how ant optimizes their path for food. Ants reduce their journey towards food by leaving a (chemical) pheromone trail as they walk. The chemical helps other ants for food search. The shorter path is indicated by a strong
trail of pheromone. The stronger pheromone draws the attention of other ants. The ants mostly choose shorter path until all find the shortest path. The ACO used the mechanism of ants and develop artificial ants starting from the initial node. The artificial ants move to feasible neighbor nodes. Each ant builds a path using the state transition rule;

\[ P(s,v) = \begin{cases} \text{argmax}_{x} \{p(s,v)^{\rho} \}, & \text{if } r \leq r_0 \\ 0, & \text{otherwise} \end{cases} \quad (2) \]

\( \eta(s,v) \) represents pheromone (desirability of \((s,v)\)) on edge \((s,v)\). \( r \) is the parameter that governs the relative value of desired, \( r_0 \) is initialized with \( 0 \leq r_0 \leq 1 \), \( r \) belongs to \( \text{rand}([0,1]) \). \( r_0 = 0.5 \) for carrying out the present study. \( J(r) \) is the set of edges available at point \( r \) of decision. \( S \) is a random variable chosen according to the probability distribution given below:

\[ P(s,r) = \begin{cases} \frac{\{p(s,v)^{\rho} \}}{\sum_{(s,v) \in J(s)} [p(s,v)^{\rho}]}, & \text{if } r \in J(s) \\ 0, & \text{otherwise} \end{cases} \quad (3) \]

Amount of pheromone modifies as

\[ \eta(s,v) \leftarrow (1-\rho)\eta(s,v) + \rho \eta_0, \quad 0 < \rho < 1 \]

where \( \rho \) is the pheromone evaporation coefficient representing. The global modified amount when all ants arrived at the destination is given by

\[ \eta(s,r) \leftarrow (1-\delta)\eta(s,r) + \Delta \eta(s,r) \quad (4) \]

where,

\[ \Delta \eta(s,r) = \begin{cases} \frac{1}{L}, & \text{if } (s,r) \in \text{global best tour} \\ 0, & \text{otherwise} \end{cases} \]

\( L \) denotes the global best tour length from the initial stage; the \( \delta \) is the global evaporation coefficient parameter, and the last term is the increase in desirability [18].

2.3.4. Firefly algorithm

This algorithm is designed on fireflies flashing characteristics [19]. FFA structure is based on (i) Firefly (assumed to be unisex) ability to captivate other; (ii) the intensity of the luminous; and (iii) amount of light emitted by the firefly.

\[ J = J e^{-\alpha d^2} \quad (5) \]

\[ v(d) = v_0 e^{-\beta d} \quad (6) \]

where \( J \) is the light intensity and \( v(d) \) attractiveness at distance \( d \) to the firefly. At \( d = 0 \), \( J = J_0 \) and \( v(0) = V_0 \), and \( \alpha \) is the light absorption coefficient. \( d \) is defined as [20, 21]:

\[ d_j = \sqrt{\sum_{i = 1}^{K} (y_i - y_i) \cdot (y_j - y_j)} \quad (7) \]

where \( y_i \) and \( y_k \) are fireflies positions \( j \) and \( k \). Firefly is captivated by yet brightness-based firefly. The movement of the firefly is given by

\[ y_j = y_j + k \cdot \eta e^{-\alpha d^2} (y_j - y_j) + \eta \eta \quad (8) \]

\( \eta \) is the random movement in case of absence of brightness, \( \eta \) varies between 0 and 1 and the second term in the equation is the captivation factor with coefficient as \( \beta \). FFA is used to tune the premise parameter of ANFIS with optimal values of coefficients as \( \alpha = 1 \) and \( \beta = 2 \), using trial and error process.

2.3.5. Differential evolution

The well-known evolutionary algorithm is based on mutation, crossover, and selection. DE is widely used for optimization problems in various fields [22]. It has been widely applied as opposed to other algorithms due to its simple structure and faster convergence rate. The algorithm is focused on biological processes that involve survival if environmental and genetic features are to be complied with. DE begins its operation by creating a random population, where each person in the population represents a problem-solving solution. The search parameter is initialized, as are the total number of generations. For the description of the algorithm refer to Wei et al. [23]. The optimal crossover probability is taken as 0.2 and scaling factor lower and upper bounds as 0.2 and 0.8.

2.3.6. Proposed algorithm

(1) \( x(t) \) at time \( t \) is considered where \( x(t) \) represents daily (24 h) concentration of air pollutants on day \( t \).

(2) \( x(t) \) is decomposed using the wavelet transform as mentioned in section 2.3.1 using high and low filters. \( x(t) \) is the sum of \( a_n \) (approximation at level \( n \)) and \( d_1, d_2, \ldots, d_n \) (details at level \( n \)). Daubechies wavelet(db5) is considered in the present study due to its property to extract fluctuations nicely.

(3) Instead of \( x(t) \), the smoothed decomposed series are used for further analysis. Let \( y(t) \) denote the approximation at level 5 i.e. \( y(t) = a_5(t) \).

(4) The autocorrelation function (ACF) for \( y(t) \) is computed as described by Mohammad et al. [24] to find out the dependence of \( y(t) \) on lag values. Let the optimized lag is \( \tau \). The input set is past \( \tau \) values of the series and the output value is \( y(\tau + 1) \).

(5) The series is divided into training (70%) and testing (30%) datasets.

(6) The parameters as described in section 2.3 are obtained. Using the algorithm described in sections 2.3.1–5, the premise and consequent parameters are trained.

(7) The parameters are trained to obtain the optimized error. The simulations are carried using Matlab R2019a software.

(8) Steps (4)-(7) are carried out for \( d_1, d_2, \ldots, d_n \). The final output is obtained adding the trained values of approximation and details. The test dataset is simulated using the model. Further effectiveness of the method is verified using trained and tested datasets.

The parameters are trained using different algorithms and the same procedure is followed. The models based on different algorithms are denoted as are compared to obtain the best model.
2.4. Evaluation Criterion
The efficacy of the model is measured against the existing models. The observed-predicted pairs \((y_O(t) \text{ and } y_p(t))\) w.r.t time \(t\) for training and testing datasets are considered. The determinism coefficient \((R^2)\) defines the relation between the aforementioned pairs. The parameters for determining errors are based on various values, and they primarily combine absolute and relative errors. Here are some of the error measurements:

\[
R^2 = 1 - \frac{\sum_{t=1}^{m} (y_O(t) - y_p(t))^2}{\sum_{t=1}^{m} (y_O(t) - \bar{y})^2}
\]

\[
\text{Mean Absolute Error, MAC} = \frac{\sum_{t=1}^{m} |y_O(t) - y_p(t)|}{m}
\]

\[
\text{Mean Absolute Percentage Error, MAPE} = \frac{100 \times \sum_{t=1}^{m} |y_O(t) - y_p(t)|}{\sum_{t=1}^{m} y_p(t)}
\]

\[
\text{Root Mean Square Error, RMSE} = \sqrt{\frac{\sum_{t=1}^{m} (y_O(t) - y_p(t))^2}{m}}
\]

To forecast the air quality, the analysis is carried out on the air pollutants. The Air Quality Index (AQI) is the metric for transmitting air quality to the public whether it is dangerous or safe to go out or take precautions. AQI standard for India was introduced in the year 2014. The best model is evaluated based on equation 9. As per CPCB for India, AQI is evaluated using formulas mentioned in equations 10-12. Using the best model the pollutant is predicted and correspondingly air quality sub-index is evaluated. The prominent sub-index gives the AQI. AQI in the present study is used to validate the proposed model. AQI value is calculated and divided into categories-Good (0-50), Satisfactory (51-100), Moderately polluted (101-200), Poor (201-300), Very poor (301-400), and Severe (401-500) as per CPCB norms.

\[
\text{NOx Air Quality Subindex} = \begin{cases} 
  \frac{X \times 50}{40} & 0 < X \leq 40 \\
  50 + \frac{(X - 40) \times 50}{40} & 40 < X \leq 80 \\
  100 + \frac{(X - 80) \times 100}{100} & 80 < X \leq 180 \\
  200 + \frac{(X - 180) \times 100}{100} & 180 < X \leq 280 \\
  300 + \frac{(X - 280) \times 100}{120} & 280 < X \leq 400 \\
  400 + \frac{(X - 400) \times 100}{120} & X > 400 
\end{cases}
\]

where, \(X\) represents NO concentration in ppb(parts per billion).

3. Results and Discussions
The random behavior of air pollutants makes the task of predicting air quality complex. The study aims to build a model from which high prediction accuracy with minimal input parameters can be obtained. The complexity comes into the frame when the series is non-stationary. The extreme outliers make the problem more challenging. The studies related to model air quality has been carried out for various environmental pollutants like NO2, PM2.5, and SO2 which play a significant role in determining the air quality. Some of the important work is shown in table 1. The problem was previously dealt with using ANFIS modeling but the large error between the predicted and observed value was observed (refer to Table1). The drawback found in classic ANFIS modeling was due to the gradient descent (GD) method. It is observed that in classic ANFIS modeling, the GD algorithm gets trapped in local minima. The tuning of premise parameters was inadequate leading to inadequate predictions. Moreover, too many parameters are also one of the reasons for erroneous prediction. The present work is an attempt to minimize the parameters at the same time improving the air quality prediction. In the current scenario, oxides of nitrogen, particulate matter, and sulphur dioxide are prominent pollutants. Though the air quality index is dependent on NOX, PM2.5, and SO2, NO and NO2 are also predicted in the present study as both have a significant correlation with NOX. The air quality index is dependent on the prominent air quality sub-index corresponding to the pollutant. The proposed hybrid models were validated for the Shadipur area of India. The hybrid model combines the decomposition filter, fuzzy inference system, algorithm for optimizing parameter, and neuronal networks.

The correlation analysis of the pollutants and meteorological parameters was studied to optimize the choice of parameters. It was found that no significant correlation existed between the weather series available (relative humidity, wind speed, and temperature) and the pollutant considered for the available data. The other attempt to find the dependent parameters of the series was to carry out a partial autocorrelation function (PACF). It is observed that in comparison to the past concentration of the pollutant no significant correlation was observed with meteorological and other pollutants. Based on the partial correlation function of pollutant, the optimized lag considered is six. Hence, the concentration of past six days is considered for predicting concentration of the seventh day. For decomposition analysis Daubechies wavelet (Db5) is used. The
The performance analysis of various models in Tables 2-6 depicts the role of wavelet decomposition in analyzing the non-stationary series. Without extracting the features i.e. without using low and high filters, a large difference is observed between predicted and observed data. ANFIS and WAFIS models are also compared for each air pollutant to signify the importance of wavelet transform.
Table 2. Performance Analysis of the Models for Predicting Nitrogen Dioxide (NO)

| MODEL        | WANFIS-PSO | ANFIS-PSO | WANFIS-GA | ANFIS-GA | WANFIS-FFA | ANFIS-FFA | RMSE   | MAE    | R2     |
|--------------|------------|-----------|-----------|----------|------------|-----------|--------|--------|--------|
| Training     | 0.9827     | 0.6161    | 0.9802    | 0.5899   | 0.9491     | 0.5775    | 0.6631 | 0.5484 | 0.0101 |
| Testing      | 0.9713     | 0.5075    | 0.978     | 0.5233   | 0.9508     | 0.5775    | 0.6316 | 0.5362 | 0.0096 |

Table 3. Performance Analysis of the Models for Predicting Particulate Matter PM2.5

| MODEL        | WANFIS-PSO | ANFIS-PSO | WANFIS-GA | ANFIS-GA | WANFIS-FFA | ANFIS-FFA | RMSE   | MAE    | R2     |
|--------------|------------|-----------|-----------|----------|------------|-----------|--------|--------|--------|
| Training     | 0.9877     | 0.7149    | 0.9872    | 0.7042   | 0.9183     | 0.7165    | 0.7812 | 0.6893 | 0.0955 |
| Testing      | 0.9906     | 0.757     | 0.9902    | 0.7533   | 0.928      | 0.6905    | 0.7092 | 0.6984 |urc |

Table 4. Performance Analysis of the Models for Predicting Nitrogen Oxides (NOx)

| MODEL        | WANFIS-PSO | ANFIS-PSO | WANFIS-GA | ANFIS-GA | WANFIS-FFA | ANFIS-FFA | RMSE   | MAE    | R2     |
|--------------|------------|-----------|-----------|----------|------------|-----------|--------|--------|--------|
| Training     | 0.9867     | 0.661     | 0.985     | 0.6393   | 0.9291     | 0.6282    | 0.6648 | 0.6308 | 0.054  |
| Testing      | 0.9595     | 0.5444    | 0.9688    | 0.5572   | 0.856      | 0.5311    | 0.632  | 0.5558 | 0.0187 |

Table 5. Performance Analysis of the Models for Predicting Particulate Matter PM10

| MODEL        | WANFIS-PSO | ANFIS-PSO | WANFIS-GA | ANFIS-GA | WANFIS-FFA | ANFIS-FFA | RMSE   | MAE    | R2     |
|--------------|------------|-----------|-----------|----------|------------|-----------|--------|--------|--------|
| Training     | 0.9877     | 0.6912    | 0.9853    | 0.6759   | 0.9514     | 0.6914    | 0.7355 | 0.6543 | 0.0132 |
| Testing      | 0.9835     | 0.6218    | 0.9854    | 0.6416   | 0.9043     | 0.6091    | 0.7889 | 0.6402 | 0.976E-05 |

Table 6. Performance Analysis of the Models for Predicting Particulate Matter PM2.5

| MODEL        | WANFIS-PSO | ANFIS-PSO | WANFIS-GA | ANFIS-GA | WANFIS-FFA | ANFIS-FFA | RMSE   | MAE    | R2     |
|--------------|------------|-----------|-----------|----------|------------|-----------|--------|--------|--------|
| Training     | 0.9877     | 0.6912    | 0.9853    | 0.6759   | 0.9514     | 0.6914    | 0.7355 | 0.6543 | 0.0132 |
| Testing      | 0.9835     | 0.6218    | 0.9854    | 0.6416   | 0.9043     | 0.6091    | 0.7889 | 0.6402 | 0.976E-05 |

Dimple Pruthi and Rashmi Bhardwaj
Further, the evolutionary algorithm implemented for tuning premise parameters has reduced the error between predicted and observed air pollutant concentrations to a larger extent. The bold glyphs in the tables signify the errors corresponding to the best model. It is observed that for NO, NO₂, NOₓ, and PM₂.₅ proposed Wavelet Transform-Adaptive Neuro-Fuzzy Inference System-Particle Swarm Optimization (WANFIS-PSO) gives better results compared to other models. The genetic algorithm is very much close to PSO. The next step for the best model was to check the computation time.

While tuning the parameters it is observed that differential evolution converges at a very fast rate. Particle Swarm Optimization and Genetic Algorithm take approximately 80 seconds to converge while FFA has a slow convergence rate compared to others. FFA

### Table 6. Performance Analysis of the Models for Predicting Sulphur Dioxide SO₂

| MODEL | WANFIS-PSO | ANFIS-PSO | WANFIS-GA | ANFIS-FFA | WANFIS-DE | ANFIS-DE | WANFIS-FFA | ANFIS-ACO | ANFIS-ACO |
|-------|------------|-----------|-----------|-----------|-----------|-----------|------------|-----------|-----------|
| Training | 0.9815 | 0.6169 | 0.9823 | 0.6031 | 0.9534 | 0.4992 | 0.5435 | 0.5688 | 4.51E-04 | 0.5531 |
| Testing  | 0.9792 | 0.5076 | 0.9817 | 0.5074 | 0.7894 | 0.4383 | 0.4606 | 0.4999 | 9.41E-05 | 0.4967 |
| RMSE Training | 1.3731 | 6.2461 | 1.3419 | 6.3677 | 2.2258 | 8.8814 | 8.601 | 6.6253 | 65.4892 | 6.7444 |
| Testing  | 0.9536 | 4.6432 | 0.8909 | 4.6311 | 3.2581 | 7.3658 | 6.1823 | 4.6738 | 56.0997 | 4.6791 |
| MAE Training | 0.7505 | 3.4773 | 0.7672 | 3.516 | 1.4525 | 7.2184 | 5.5394 | 3.6001 | 45.1907 | 3.6418 |
| Testing  | 0.6069 | 2.9789 | 0.583 | 2.9574 | 1.3042 | 6.2474 | 4.5481 | 2.9866 | 38.7643 | 2.98 |
| MARE Training | 11.4649 | 51.7641 | 11.6339 | 51.372 | 22.2842 | 140.8222 | 82.0887 | 53.3888 | 666.2866 | 55.1915 |
| Testing  | 6.3206 | 33.1995 | 6.0066 | 32.8453 | 14.7799 | 89.5434 | 49.3311 | 33.1355 | 411.338 | 34.0932 |

Fig. 2. Predicted and Observed concentration for the tested dataset of pollutant (a) NO, (b) NO₂, (c) NOₓ, (d) PM₂.₅, (e) SO₂ and (f) AQI values.
gives better results than ACOR and DE. ACOR is most often used for continuous problems. It is observed that ACOR possesses random behavior in case of outliers hence very large relative error. The comparison of algorithms in the study covers almost all aspects in the present scenario to develop a good model for AQI prediction. The daily (24-h) AQI value is evaluated based on the prominent pollutant. The observed and predicted concentrations of pollutants and AQI values corresponding to the best model are very close for the tested data as depicted in Fig. 2(a)-(e) and Fig. 2(f), respectively. AQI prediction met 88.63% accuracy and remaining AQI values lie very close to the breakpoints of the categories as defined in section 2.4. The accuracy level in predicting the air quality index validates the study to be used for making policies and beforehand precautions to be marked safe from air pollution.

4. Conclusion

The atmospheric dynamics is complicated due to the extreme outliers and data unavailability making the study highly expensive. The air quality index model developed in the present study is effective and less complex in forecasting the one-step-ahead quality in case of outliers hence very large relative error. The error. Thus, an effective model for predicting air quality index is obtained which can be adopted by air pollution agencies and the Government for policymaking. The model developed in the present work can be efficiently used and applied to any real-time series to observe future behavior.

Acknowledgment

The authors are thankful to Guru Gobind Singh Indraprastha University, Delhi, India for providing research facilities and financial support.

Author Contributions

D.P. (Ph.D. student) conducted all the experiments and wrote the manuscript. R.B. (Professor) revised the manuscript.

References

1. IQAir. 2018 World Air Quality Report. IQAir AirVisual 2018: c2019 [cited 05 March 2019]. Available from: https://www.iqair.com.
2. Jingchun F, Shulan L, Chunling F, Zhenggang B, Ke-Hu Y. The impact of PM2.5 on asthma emergency department visits: a systematic review and meta-analysis. *Environ. Sci Pollut. Res.* 2016;23:843-850.
3. Cole B, Jeffrey RS, Andrew B, Patrick R. Pediatric Psychiatric Emergency Department Utilization and Fine Particulate Matter: A Case-Crossover Study. *Environ. Health Persp.* 2019;127(9): 097006-1-097006-7.
4. Fu P, Guo X, Cheung FMH, Yung KKL. The association between PM2.5 exposure and neurological disorders: A systematic review and meta-analysis. *Sci. Total. Environ.* 2019;655:1240-1248.
5. Vera P, Kayo U, Shunji K, et al. Acute effects of Ambient PM2.5 on all-cause and cause-specific emergency ambulance dispatches in Japan. *Int. J. Environ. Res. Public Health* 2018;15(2):307.
6. Jalalifar H, Mojedifar S, Sahebi AA, Nezamabadi-pour H. Application of the adaptive neuro-fuzzy inference system for prediction of a rock engineering classification system. *Comput. Geotech.* 2011;38: 783-790.
7. Rajesh S, Ashutosh K, TN S. Estimation of elastic constant of rocks using an ANFIS approach. *Appl. Soft Comput.* 2012;12: 40-45.
8. Rai AA, Pai PS, Rao BRS. Prediction models for performance and emissions of a dual fuel CI engine using ANFIS. *Sadhana* 2015;40:515-535.
9. S Sankar G, Arulmozhiwarman P, Rao T. Prediction of PM 2.5 using an ensemble of artificial neural networks and regression models. *J. Ambient Int. Humanized Comput.* 2018;1-11.
10. Xianghong W, Baozhen W. Research on prediction of environmental aerosol and PM2.5 based on artificial neural network. *Neural. Comput. Appl.* 2019;31:8217-8227.
11. Jyh-Shing RJ. ANFIS: adaptive-network-based fuzzy inference system. *IEEE Trans. Sys. Man Cybern.* 1993;23:665-685.
12. Rashmi B, Dimple P. Development of model for sustainable nitrogen dioxide prediction using neuronal network. *Int. J. Environ. Sci. Technol.* 2020;17:2783-2792.
13. Marija PES, Ivan NM, Živan Ž. An ANFIS based air quality model for prediction of SO2 concentration in urban area. *Serbian J. Manag.* 2013;8(1):25-38.
14. Yegang C. Prediction algorithm of PM2.5 mass concentration based on adaptive BP neural network. *Computing* 2018;100(2): 825-838.
15. David EG, John HH. Genetic algorithms and machine learning. *Mach. Learn.* 1988;3:95-99.
16. Zhe Y, Lin-Na W, Xu J. Prediction of concrete compressive strength: Research on hybrid models genetic based algorithms and ANFIS. *Adv. Eng. Softw.* 2014;67:156-163.
17. De HZ, Yuan L, Lian BJ, Li L, Gang X. A New Approach to Cutting Temperature Prediction Using Support Vector Regression and Ant Colony Optimization. *Adv. Eng. Forum* 2012;4:145-152.
18. Hadi F, Hosine N, Abdulllah M. Hybrid ANFIS with ant colony optimization algorithm for prediction of shear wave velocity from a carbonate reservoir in Iran. *Int. J. Min. Geo-Eng.* 2016;50(2):231-238.
19. Zehir MY, Isa E, Hossein B, et al. Novel approach for streamflow forecasting using a hybrid ANFIS-FFA Model. *J. Hydrol.* 2017;554:263-276.
20. Xin-Shuo Y. Firefly algorithm, stochastic test functions and design optimization. *Int. J Bio-inspired Comput.* 2010;2(2):78-84.
21. Saeed K, MHH Y, Amin P, Mehrdad D. Application of firefly algorithm and ANFIS for optimisation of functionally graded beams. *J. Exp. Theor. Artif. Intell.* 2014;26(2):197-209.

22. Rainer S, Kenneth P. Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *J. Glob. Optim.* 1997;11:341-359.

23. Wei C, Mahdi P, Hamid RP. Performance evaluation of GIS-based new ensemble data mining techniques of adaptive neuro-fuzzy inference system (ANFIS) with genetic algorithm (GA), differential evolution (DE), and particle swarm optimization (PSO) for landslide spatial modelling. *Catena* 2017;157: 310-324.

24. Mohammad V, ME B, Seyyed MRB. Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. *J. Hydrol.* 2013;476:433-441.

25. Fabio B, Marcello B, Marco V et al. Recursive neural network model for analysis and forecast of PM10 and PM2.5. *Atmos. Pollut. Res.* 2017;8(4):652-659.

26. Brunellia U, Piazzaa V, Pignatosa L, Sorbellob F, Vitabilec S. Two-days ahead prediction of daily maximum concentrations of SO2, O3, PM10, NO2, CO in the urban area of Palermo, Italy. *Atmos. Environ.* 2007;41:2967-2975.

27. Ian GM. Evaluation of Artificial Neural Networks for Fine Particulate Pollution (PM10 and PM2.5) Forecasting. *J. Air Waste Manage. Assoc.* 2002;52(9):1096-1101.

28. Polaiah B, Deepika N, Harika C. Prediction Of Nitrogen Dioxide & Ozone Concentrations In The Ambient Air Using Artificial Neural Networks For Visakhapatnam Model. *Int. J. Pure Appl. Mathematics.* 2017;117(19):83-88.

29. Joaquin OM, Eliseo V, Salvador FCR, RE S. Neural network prediction model for fine particulate matter (PM2.5) on the US-Mexico border in El Paso (Texas) and Ciudad Juarez (Chihuahua). *Environ. Modell. Softw.* 2005;20:547-559.

30. Sait CS, Aysun S, Savas B, Gokmen T. Forecasting Ambient Air SO2 Concentrations Using Artificial Neural Networks. *Energy Sources Part B.* 2006;1(2):127-136.

31. Wei S, Jingyi S. Daily PM2.5 concentration prediction based on principal component analysis and LSSVM optimized by cuckoo search algorithm. *J. Environ. Manage.* 2016;188:144-152.

32. Ignacio JT, Francisco JG, Ma LM, Pedro LG. Prediction models of CO, SPM and SO2 concentrations in the Campo de Gibraltar Region, Spain: a multiple comparison strategy. *Environ. Monit. Assess.* 2008;143:131-146.