Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(5), 718-725 | ISSN:2251-6727

Development of ANN model for Prediction of AQI at Sanathnagar, Hyderabad, India.

[1]Bhavana. Hemavani, [2]G. V. R. Srinivasa Rao

[1]Corresponding Author: Research Scholar, Department of Civil Engineering, University College of Engineering (A), Andhra University, Visakhapatnam, India.

[2]Professor, Department of Civil Engineering, Andhra University College of Engineering (A), Andhra University, Visakhapatnam, India.

ORCiD ID: [a]Bhavana Hemavani, https://orcid.org/0009-0009-4244-2070.

[b]G.V.R.Srinivasa Rao, https://orcid.org/0000-0001-7558-4440

(Received: 02 September 2023 Revised: 14 October Accepted: 07 November)

KEYWORDS

Artificial Neural Networking, Air Quality Index, Prediction, MATLAB.

ABSTRACT:

Air pollution is not only environmental problem, but also a major health issues in many countries. The effective methods in controlling the air pollution is by a proper air quality management. One of effective methods is by predicting the air quality index. One such method is based on prediction of data is by using Artificial neural networking (ANN). In this study 21 models were developed for prediction of pollutants. The required data for development of ANN models were collected from the TSPCB (Telangana State Pollution control board), Hyderabad, India. The data comprises 2007-2017 years of hourly data. Best suited models were selected on the basis of statistical tools such as R and MSE. This study was conducted seasonally-Summer, Post-monsoon, Winter. The predicted data is for 2017 – 2023. The AQI (predicted) is Computed by using IND-AQI CPCB method and the quality rating is giving by color coding. After the prediction a comparison was drawn between AQI (observed) and AQI (predicted), then the conclusions were drawn from this table.

1. Introduction

Air quality management is a crucial process that involves monitoring air quality, assessing the impact of human activities, taking measures to improve the situation, and ensuring that these measures are effective. This system is designed to reduce the emission of pollutants and other harmful substances in the atmosphere and to sustain ambient air quality. In recent years, effective methods have been developed globally to understand and summaries the importance of good air quality. One of the best ways to maintain air quality is by computing the air quality index, which helps to categorize the meteorological conditions relevant to the Air Quality of that country or city.

In recent times many methods have been developed for prediction of AQI. Recent Studies have proving that ANN method is the most interesting approach to predict AQI.

ANNs are used in complex situations, especially in predicting air pollutants, to provide better results. ANN models can predict air quality over various time periods, making them ideal for predicting regional or global air quality during specific periods. ANNs differ from traditional methods in terms of mapping variables, prediction accuracy, and handling missing values. ANNs can be categorized into two types: Feed-forward networks and recurrent networks. Feed-

Journal of Chemical Health Risks
www.jchr.org
JCHR (2023) 13(5), 718-725 | ISSN:2251-6727

forward networks have one layer of hidden neurons, while recurrent networks have at least two layers of hidden neurons.

When describing air quality, it is important to report the concentration of all pollutants in that location. Air quality parameters include gases, volatile organic compounds, and meteorological parameters. In India, rapid urban development has led to severe crises in air quality, affecting biology, physics, and economic system. Based on the above passages, developing an approach to interpret and predict future air concentrations in urban environment is urgently required. One such work is aimed at in this study.

2. Objective of the Study

The primary objective of this study is to predict the ambient air quality of the study area by developing an artificial neural network model using real time data and finding the significant meteorological factors affecting the air pollutants concentration in a given period.

3. Study area

The study area Sanathnagar is both industrial cum residential where the industrial area is home to a variety of small and medium scale chemical, pharmaceutical, electrical and mechanical industries. The connectivity of study area to neighboring places is via road, which implies that most of the pollution concentrations are related to vehicular emissions. These conditions make it an ideal study area for understanding and predicting ambient air quality. The Telangana state Pollution Control Board (TSPCB) provides the hourly atmospheric concentration of all the pollutants that fall under the Central Pollution Control Board (CPCB).

4. Data collection and aggregation

The data relating to the sampling location Sanathnagar is collected from TSPCB, Hyderabad, Telangana, India for the years 2007-2017. The following Table 1 describes the types of parameters collected from TSPCB for this study.

Table 1: Type of parameters collected

S.No	Parameters	Type of parameters	Unit
1	Carbon Monoxide CO	Ambient air pollutant	$\frac{\mu g}{m^3}$
2	Sulphur dioxide SO_2	Ambient air pollutant	$\frac{\mu g}{m^3}$
3	Particulate Matter PM	Ambient air pollutant	$\frac{\mu g}{m^3}$
4	Oxides of Nitrogen NO_x	Ambient air pollutant	$\frac{\mu g}{m^3}$
5	Benzene C_6H_6	Ambient air pollutant	$\frac{\mu g}{m^3}$
6	Toluene $C_6H_5CH_3$	Ambient air pollutant	$\frac{\mu g}{m^3}$
7	Xylene $(CH_3)_2C_6H_4$	Ambient air pollutant	$\frac{\mu g}{m^3}$
8	Atmospheric Temperature AT	Meteorological Parameter	°C
9	Relative Humidity RH	Meteorological Parameter	%
10	Wind Speed WS	Meteorological Parameter	$\frac{m}{s}$

Journal of Chemical Health Risks
www.jchr.org
JCHR (2023) 13(5), 718-725 | ISSN:2251-6727

11	Sun radiation SR	Meteorological Parameter	$\frac{w}{m^2}$
12	Wind Direction WD	Meteorological Parameter	deg

5. Artificial Neural network (ANN) modelling

ANN modelling has several advantages over statistical methods with an accurate prediction. The objective of this study is to implement ANN modelling for predicting air quality based on the previously available data. Therefore, in this modelling the prediction is carried out for the air pollutants such as CO, SO_2, NO_x, PM, Benzene, Toluene, and Xylene. The prediction is carried out by using meteorological parameters as set of inputs. here the input data is obtained from Principal component analysis and Cluster analysis outputs.

The air quality data obtained from TSPCB is divided into two parts: Seasonal data from 2007 to 2016, used for model development and training of the network, whereas 2017 data, used for additional testing of the network. Therefore, the model prediction is given for 2017 to 2026, and the predicted data was validated by comparing with 2017 (observed) data.

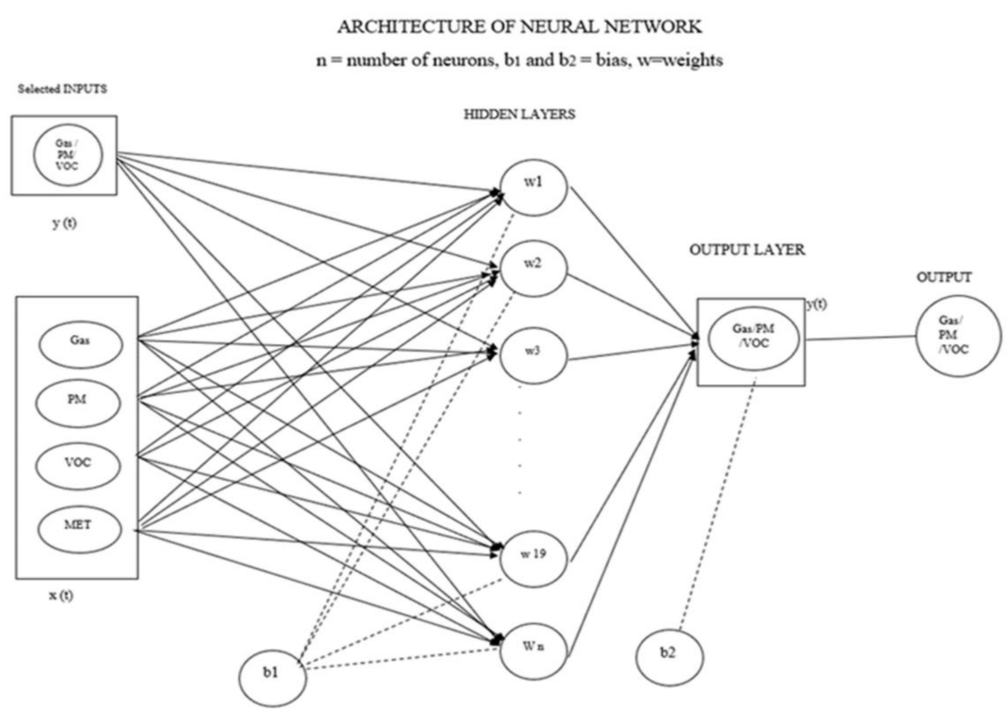

Figure 1: Arcitecture of Nueral Network

Journal of Chemical Health Risks
www.jchr.org
JCHR (2023) 13(5), 718-725 | ISSN:2251-6727

6. Results and Discussion

6.1 Architecture of the neural network

This network model utilizes non-linear time series prediction and modelling tool (ntstool) with two layered architecture. The network is Non- Linear auto regression exogenous model (NARX) network, this network predicts the series y(t) based on the past values of y(t) and another series x(t). This is feed-forward network.

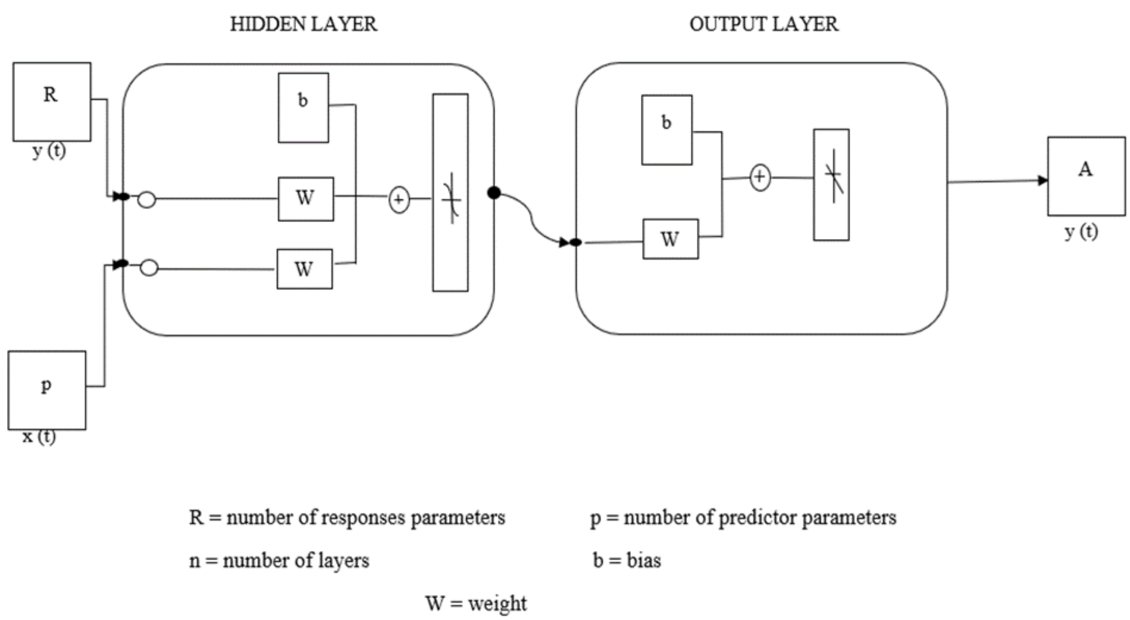

Figure 2: NARX Network Model

The figure 2 displayed above represents the model structure developed in this research shows a two layered simple feed forward network. It comprises of four sections. The first section includes y(t) and x(t). The second section consists of a hidden layer that comprises of weights, bias neurons, and input variables. This is where the network training takes place. The third section includes output variables that are influenced by the weights and bias. This is where the network validation and testing are conducted. The final stage involves obtaining predicted variables through prediction.

The input data for modelling varies based on the season and model. The number of neurons and hidden layers in the network architecture effect the output, therefore different sets of neurons (n) are used, specifically 10,15 and 20. Here the tool has segregated the input data for each model is segregated into three categories training data (70%), validation data (15%), and test data (15%), with a time delay of 2. The training algorithm: Levenberg- Marquardt is applied for all the models.

6.2 Selected ANN models for Data Prediction

A total of 180 ANN models were created for all the three seasons to predict and develop AQI in the study area. Out of 180 models, 21 were selected because they had the best suited values of R and MSE. The selection was based on accuracy, efficiency, and the ability to handle large data sets, and by taking into consideration the statistical values of both MSE, R and testing of the network. The Table 2 consist of the structural architecture of the different parameters as follows

Journal of Chemical Health Risks
www.jchr.org
JCHR (2023) 13(5), 718-725 | ISSN:2251-6727

"IN-n-OP". where IN = number of input variables, n= number of layers, OP= number of output variables. The time mentioned in the table head indicates the time required by the model to produce results, which is in minutes.

Table 2: ANN models for Data Prediction

S.NO	SEASON	PARAMETER	NETWORK ARCHITECTURE (IN-n-OP)	TIME (MIN)	TRAINING RESULTS	TESTING RESULTS
1	SUMMER	CO	4-20-1	0	MSE = 0.034 R=0.749	MSE =0.034 R=0.673
		SO_2	4-15-1	0	MSE =34.323 R=0.7161	MSE = 4.616 R=0.3302
		NOx	6-10-1	0	MSE = 455.33 R=0.7852	MSE =260.76 R=0.7376
		PM	6-20-1	0	MSE =842.9 R=0.854	MSE =355.79 R=0.741
		BENZENE	4-15-1	0	MSE =1.152 R= 0.858	MSE = 0.235 R= 0.230
		TOLUENE	4-20-1	0	MSE =130.33 R= 0.818	MSE = 55.186 R= 0.3011
		XYELENE	4-20-1	0	MSE = 3.028 R= 0.881	MSE = 0.9768 R= 0.290
2	POST-MONSOON	CO	5-15-1	0	MSE = 0.027 R=0.698	MSE = 0.029 R=0.029
		SO2	5-15-1	0	MSE =61.68 R=0.859	MSE = 15.327 R=0.0095
		NOx	7-15-1	0	MSE = 516.37 R=0.775	MSE = 266.90 R=0.270
		PM	4-15-1	0	MSE = 861.35 R=0.900	MSE =303.83 R=0.743
		BENZENE	6-15-1	0	MSE =18.722 R=0.957	MSE = 2.089 R=0.150
		TOLUENE	6-15-1	0	MSE =306.69 R=0.983	MSE = 183.50 R=0.457
3	WINTER	CO	4-15-1	0	MSE =0.0269 R=0.814	MSE =0.042 R=0.620
		SO2	4-15-1	0	MSE = 87.27 R=0.795	MSE = 80.780 R=0.780
		NOx	6-20-1	0	MSE =674.53 R=0.846	MSE =602.13 R=0.611
		PM	4-10-1	0	MSE	MSE = 509.94

Journal of Chemical Health Risks
www.jchr.org
JCHR (2023) 13(5), 718-725 | ISSN:2251-6727

					=1.0170e+03 R=0.885	R=0.555
BENZENE	6-20-1	0	MSE = 13.929 R=0.854	MSE =2.331 R=0.7910		
TOLUENE	4-20-1	0	MSE = 389.54 R=0.793	MSE = 190.618 R=0.795		
XYELENE	6-20-1	0	MSE =7.436 R=0.916	MSE =2.60 R=0.628		

The three plots used for understanding the accuracy of the model, are the performance plot, error histogram plot and the response plot.

6.3 Comparison study on AQI Observed and Predicted

The comparison is observed in the tabular form to facilitate seasonal study. Therefore, two set of data sets are used for the same purpose where the predicted data is observed from 2024 – 2026 and for the verification of model is done with the data 2017-2023. For the verification purpose a set of AQI data (observed) was downloaded from the CPCB website.

The comparison of AQI is shown in the below mentioned table 3, this table consist of observed and predicted values of AQI for all the three seasons.

Table 3: AQI: Predicted and Observed Value

Season	Year	Month	AQI Predicted		AQI CPCB	
			Value	Quality status	Value	Quality status
SUMMER	2017	MARCH	128	MODERATE	90	SATISFACTORY
		APRIL	91	SATISFACTORY	107	MODERATE
		MAY	79	SATISFACTORY	73	SATISFACTORY
		JUNE	62	SATISFACTORY	38	GOOD
	2018	MARCH	67	SATISFACTORY	122	MODERATE
		APRIL	99	SATISFACTORY	83	SATISFACTORY
		MAY	102	MODERATE	69	SATISFACTORY
		JUNE	99	SATISFACTORY	45	GOOD
	2019	MARCH	117	MODERATE	89	SATISFACTORY
		APRIL	110	MODERATE	69	SATISFACTORY
		MAY	107	MODERATE	77	SATISFACTORY
		JUNE	57	SATISFACTORY	52	SATISFACTORY
	2020	MARCH	111	MODERATE	66	SATISFACTORY
		APRIL	92	SATISFACTORY	56	SATISFACTORY
		MAY	75	SATISFACTORY	63	SATISFACTORY
		JUNE	57	SATISFACTORY	34	GOOD
	2021	MARCH	143	MODERATE	128	MODERATE
		APRIL	117	MODERATE	90	MODERATE
		MAY	140	MODERATE	48	GOOD

Journal of Chemical Health Risks
www.jchr.org
JCHR (2023) 13(5), 718-725 | ISSN:2251-6727

Season	Year	Month	AQI Predicted		AQI CPCB	
			Value	Quality status	Value	Quality status
		JUNE	88	SATISFACTORY	36	GOOD
	2022	MARCH	62	SATISFACTORY	123	MODERATE
		APRIL	116	MODERATE	81	SATISFACTORY
		MAY	120	MODERATE	74	SATISFACTORY
		JUNE	63	SATISFACTORY	47	GOOD
	2023	MARCH	143	MODERATE	74	SATISFACTORY
		APRIL	117	MODERATE	82	SATISFACTORY
		MAY	140	MODERATE	67	SATISFACTORY
		JUNE	88	SATISFACTORY	46	GOOD
POST-MONSOON	2017	OCTOBER	114	MODERATE	113	MODERATE
		NOVEMBER	135	MODERATE	107	MODERATE
	2018	OCTOBER	123	MODERATE	86	SATISFACTORY
		NOVEMBER	145	MODERATE	163	MODERATE
	2019	OCTOBER	116	MODERATE	93	SATISFACTORY
		NOVEMBER	108	MODERATE	188	MODERATE
	2020	OCTOBER	79	SATISFACTORY	103	MODERATE
		NOVEMBER	107	MODERATE	126	MODERATE
	2021	OCTOBER	116	MODERATE	94	SATISFACTORY
		NOVEMBER	134	MODERATE	106	MODERATE
	2022	OCTOBER	102	MODERATE	97	SATISFACTORY
		NOVEMBER	121	MODERATE	178	MODERATE
WINTER	2017	JANUARY	125	MODERATE	173	MODERATE
		FEBRUARY	114	MODERATE	203	POOR
		DECEMBER	129	MODERATE	243	POOR
	2018	JANUARY	125	MODERATE	229	POOR
		FEBRUARY	110	MODERATE	125	MODERATE
		DECEMBER	192	MODERATE	202	POOR
	2020	JANUARY	146	MODERATE	211	POOR
		FEBRUARY	140	MODERATE	100	MODERATE
		DECEMBER	122	MODERATE	176	MODERATE
	2021	JANUARY	107	MODERATE	104	MODERATE
		FEBRUARY	103	MODERATE	95	SATISFACTORY
		DECEMBER	123	MODERATE	192	MODERATE
	2022	JANUARY	122	MODERATE	159	MODERATE
		FEBRUARY	136	MODERATE	158	MODERATE
		DECEMBER	136	MODERATE	185	MODERATE
	2023	JANUARY	138	MODERATE	141	MODERATE
		FEBRUARY	147	MODERATE	107	MODERATE

Season	Year	Month	AQI Predicted		AQI CPCB	
			Value	Quality status	Value	Quality status
		DECEMBER	110	MODERATE	155	MODERATE

The Central Pollution Control Board (CPCB) has defined an AQI of 50 as fairly clean atmospheric conditions. However, it has been observed that the predicted AQI for all three seasons has been consistently above this value throughout the study period. This indicates that the air quality in the region needs improvement.

7. Conclusion

Out of the three seasons, winter can be named as the worst season or highly polluted season with the worst air quality rating. After the Table 3 observation and the quality rating table it can be concluding that the model is “correct” and giving satisfactory results as both the predicted and observed status are similar.

8. Acknowledgement

We sincerely thank Department of Civil Engineering, University College of Engineering (A), Andhra University, Visakhapatnam, India for constant encouragement. We extended our gratitude towards Telangana state pollution control board (TSPCB), Hyderabad, India for providing the required data.

References

[1] Abdul-Wahab SA, Al-Alawi SM. Assessment and prediction of tropospheric ozone concentration levels using artificial neural networks. Environmental Modelling & Software. 2002 Jan 1;17(3):219-28.

[2] Anggraini, Asfilia Nova, et al. "Air Quality Forecasting in DKI Jakarta Using Artificial Neural Network." *MATICS: Jurnal Ilmu Komputer dan Teknologi Informasi (Journal of Computer Science and Information Technology)* 14.1 (2022): 1-5

[3] Baawain, Mahad S., and Aisha S. Al-Serihi. "Systematic approach for the prediction of ground-level air pollution (around an industrial port) using an artificial neural network." *Aerosol and air quality research* 14.1 (2014): 124-134.

[4] Díaz-Robles, Luis A., et al. "A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: The case of Temuco, Chile." *Atmospheric Environment* 42.35 (2008): 8331-8340.

[5] Hamdan, Mohammad A., Mohammad F. Bani Ata, and Ahmad H. Sakhrieh. "Air Quality Assessment and Forecasting Using Neural Network Model." *Journal of Ecological Engineering* 22.6 (2021): 1-11.

[6] Kumar, Navneet, Anirban Middey, and Padma S. Rao. "Prediction and examination of seasonal variation of ozone with meteorological parameter through artificial neural network at NEERI, Nagpur, India." *Urban Climate* 20 (2017): 148-167.

[7] Liu, Jiaqi. "Method of outdoor PM2. 5 concentration prediction based on MATLAB." IOP Conference Series: Earth and Environmental Science. Vol. 332. No. 2. IOP Publishing, 2019.