Journal of Soft Computing Paradigm (ISSN: 2582-2640)
www.irojournals.com/jscp/

Analysis of Statistical Trends of Future Air Pollutants for Accurate Prediction

S. Kavitha[1], J. Manikandan[2]

[1]Professor, Department of ECE, Hindusthan Institute of Technology, Coimbatore, India
[2]Professor, Department of Mechanical Engineering, Hindusthan College of Engineering and Technology, Coimbatore, India

E-mail: [1]kavithamani2003@gmail.com, [2]manikandan.hcet@gmail.com

Abstract

The climate change may be mitigated, and intra air quality assessment and local human well-being can benefit from a decrease in emission of pollutant content in the air. Monitoring the quality of the air around us is one way to do this. However, a location with various emission sources and short-term fluctuations in emissions in both time and space, and changes in winds, temperature, and precipitation creates a complex and variable pollution concentration field in the atmosphere. Therefore, based on the time and location where the sample is obtained, the measurement conducted are reflected in the monitoring results. This study aims to investigate one of India's most polluted cities' air quality measurements by greenhouse gas emissions. Using the Mann-Kendall and Sen's slope estimators, the research piece gives a statistical trend analysis of several air contaminants based on previous pollution data from Mumbai, India's air quality index station. In addition, future levels of air pollution may be correctly forecasted using an autoregressive integrated moving average model. This is followed by comparing different air quality standards and forecasts for future air pollution levels.

Keywords: Air pollutants, statistical approach, air quality index, greenhouse gas, slope estimators

1. Introduction

The Kyoto Protocol was the first treaty to acknowledge human-caused global warming. Climate phenomena such as heatwaves, cyclones, and heavy rain may cause catastrophes and damage because of their unpredictability. The target greenhouse gases include CO_2, CH_4, N_2O, and fluorinated gases, according to the protocol's definition (F-gases), with the unit as Parts Per Million (PPM) respectively. There is a huge amount of carbon dioxide (CO_2) equivalents

DOI: https://doi.org/10.36548/jscp.2021.4.002
Received: 13.11.2021, received in revised form: 12.12.2021, accepted: 20.12.2021, published: 30.12.2021

produced by human activity each year, most of which comes from fossil fuels used in the energy sector [1-5]. For example, cutting gas emissions from various sectors such as energy-related projects and travel-based sectors and industrial smoke, changes the environment while improving air quality and public health.

Regulated air pollution data are commonly used in worldwide population-based epidemiological investigations. Because of their limited geographic coverage, regulatory monitors tend to focus on locations with large population. Efforts to evaluate air quality have been stepped up, including the use of data sources like dispersion models and satellite imaging, in addition to monitoring measures [6-9]. Health effect and health impact evaluations increasingly rely on these data sets, frequently with better spatial-temporal coverage and resolution. Table 1 shows some standard indication of the presence of gases.

Table 1. Standard indication of presence of gases

Scale	**CO (ppm) (Realtime)**	**SO_2 (ppm) (Realtime)**	**NO_2 (ppm) (Realtime)**	**Standard Indication**
Good	0 – 4.4	0 - 35	0 – 50	
Moderate	4.5 – 9.4	36 – 75	51 – 100	
Unhealthy for sensitive people	9.5 – 12.4	76 – 186	101 – 350	
Unhealthy	12.5 – 15.4	187 – 302	351 – 630	
Very Unhealthy	15.5 – 30.4	303 – 605	631 – 1250	
Hazardous	30.5 – 50	606 - 1000	1251 - 2000	

Globally, urban air pollution concentrations have risen during the last year. According to the World Health Organization (WHO), more than 80% of people who live in metropolitan areas where air pollution is measured, are exposed to levels that exceed WHO standards. From 2012 to 2020, the WHO estimates this rise to reach 8%. In major cities, air pollution is a major issue, increasing the risk of heart attacks, strokes, lung cancer, and other acute and chronic respiratory illnesses, such as asthma. Building materials and cultural artefacts may also be damaged [2]. There has been extensive research on the negative effects of air pollution and its sources, and the decline in urban quality is primarily due to an increase in traffic emissions, which account for most of the air pollution.

In environmental research, a variety of multivariate approaches are used since they give information on association, interpretation, and modelling of big ecological datasets. In order to determine the link between pollutants or other variables that impact air quality, a useful statistical technique is correlation analysis, and it is good to know or seek out the most important factors or sources of chemical components [10-14].

This research article contains several sections to explain the entire flow of work; section 2 provides past research work of trend analysis of air pollutants. Section 3 discusses the proposed research idea with statistical analysis. Section 4 delivers obtained results from the proposed work. Finally, section 5 concludes this research work with forthcoming enhancements.

2. Related Works

There are many studies on air pollutants through ozone level concentric regions worldwide. Liu et al. deliver the various regression models used to find the change of ozone and temperature in Taiwan. Besides, tree planting has been shown and tested to increase the air quality measurement [15]. Wang et al. found that air quality improvement zones in Taiwan may reduce global warming by capturing CO_2 and storing it. There is some disagreement on the long-term effects of Electric Vehicle (EV) penetration on Taiwan's air quality, although Li et al. found that traditional fuel-based cars are responsible for a significant portion of the emissions of air pollutants. A correlation between dry eye illness and air pollutants (such as CO and NO_2), as well as temperatures, was also discovered [16,17].

The ARIMA modelling was used to anticipate the pollutants' values with high accuracy after the M-K test and Sen's slope estimator tests were performed to determine the presence of a trend in the pollutants' time-series data. Using data from an Air Quality Index (AQI) station in Varanasi, India, the research investigated the trends and anticipated pollutant concentrations [18].

Using AQI pollution data, many statistical modelling approaches are available for trend analysis and forecasting purposes. The proposed statistical modelling techniques of Zuma Netshiukhwi et al. do not rely on traditional environmental predictions or pollutants formulation through chemical with absolute values to use old data of air pollutants data, for estimating the air quality in the near future [19].

For the suggested research, nonparametric tests are used for statistical analysis. According to Watthancheewakul et al.'s research work, nonparametric tests don't have to adhere to the standard distribution of independent data as parametric tests do [20].

3. Proposed Statistical Trend Analyses

3.1 Collection of Air Quality data

The pollutant data from Mumbai, India, is used in this research. Located in the western Maharashtra area, Mumbai is a significant commercial centre with a big port. Mumbai's air quality was found to be the worst in India in a recent assessment, and according to the municipal pollution control board's 2017 statistics, there was not a single day of "excellent" air quality [21 – 25]. Figure 1 shows the overall proposed architecture for future prediction air pollutants.

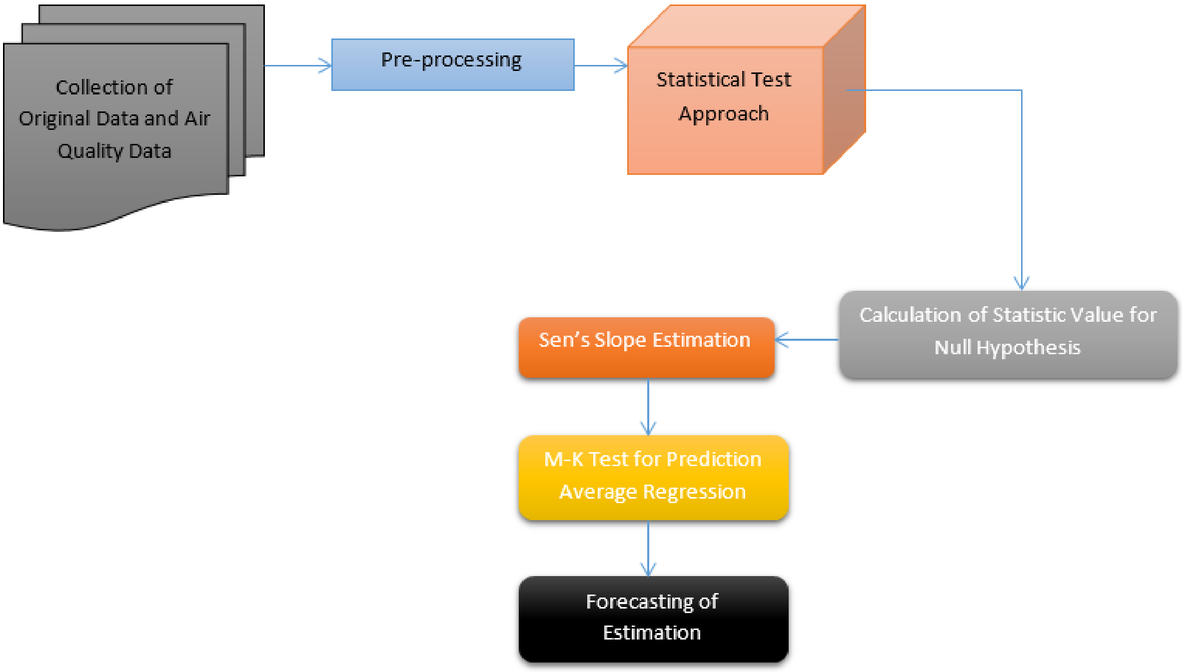

Figure 1. Simplified proposed architecture for future prediction air pollutants

3.2 Statistical Test Approach

When it comes to testing, this study uses a method known as the Mann–Kendall (M–K) test, which was first developed in 1945 by Mann and then revised by Kendell (1975). To reject the null hypothesis, no monotonic trend is assumed in the data. The alternative hypothesis indicates that a monotonic trend exists, either positively or negatively. It is commonly agreed that the M-K test is a proper statistical technique for predicting and analysing various contaminants' statistical patterns over time [26-29]. For the M-K test, the P statistics value is derived in the following manner:

3.3 Calculation of AQI in Mumbai

Currently, the AQI in Mumbai's unhealthy air pollution level is shown in table 2 & 3.

Step 1

$$Statistic\ value = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left|x_j - x_i\right|$$

Pollutant data points with "j" being more significant than "i" is found in time-series, where N is the number of data points in the time-series. According to null hypothesis H, all individual pollutant levels for each day of the year across the year are not in a trend. Alternate hypothesis H claims that pollutant data points show a monotonic trend.

Step 2

$$\left|x_j - x_i\right| = \begin{cases} 1 & if\ x_j - x_i > Zero \\ 0 & if\ x_j - x_i = Zero \\ 1 & if\ x_j - x_i < Zero \end{cases}$$

Step 3

For the powerful Sen's slope estimation with trend estimation,

$$P_i = \frac{x_j - x_i}{b - a}\ ; for\ i = 1,2,3 \ldots .N$$

Step 4

The accurate prediction is defined as,

$$P_{avg} = \begin{cases} p_{\frac{mean+1}{2}} \\ \dfrac{p_{\frac{mean/2}{2}} + p_{\frac{mean+2}{2}}}{2} \end{cases}$$

The M-K test uses a significant level prediction average to anticipate the trend of contaminants, although there is the potential for alternative significant levels. Sen's slope

ISSN: 2582-2640

estimator is used to determine the rate of change for contaminants that do not move in the M-K Test.

3.4 Forecasting of estimation by Autoregressive mode

This Auto Regressive Integrated Moving Average (ARIMA) was developed by Box and Jenkins, initially used to forecast and estimate future values in univariate time-series data (1976). Besides, this technique uses various time series methods to represent better and analyze time-series data. With the ARIMA model, p is the order of the autoregression, d is for differencing order integration, and q is for moving average order. The first stage in the modeling process is to determine whether the time series data is stationary.

4. Results and Discussion

In this portion of the research, the Mann-Kendall test, Sen's slope estimator test, and ARIMA modeling of time-series pollutants from the Maharashtra state [30 -32] AQI sampling station data are used. From the AQI station, the results of the Mann-Kendall test for various pollutants are computed and shown in Table 1-3 for the period from 2017 to 2021.

Table 2. Computed gas values by statistical trend analysis

Year	SO_2 (ppm) (Realtime)	CO (ppm) (Realtime)	NO_2 (ppm) (Realtime)
2018-19	3.12	0.32	13.62
2019-20	2.38	0.31	11.57
2020-21	2.81	0.32	13.20
2021-22	3.41	0.41	14.31

Several factors may impact the air quality in Mumbai, including the city's fast growth based on pollution by smoke creation factories, fire-based power plants in various weather situations etc. In addition, the massive number of automobiles and pollution creating factories multiply the pollution while receiving the dust from desert storms.

Table 3. Year-wise prediction of the status of Mumbai's AQI

Year	Annual Average		Moderate	Unhealthy for sensitive people	Unhealthy
	PM10	NO_2			
2009 -10	178	99	20	123	222
2017-18	121	60	85	154	126
2018-19	129	75	121	181	63
2019-20	95	47	180	145	40
2020-21	130	71	105	143	116
2021-22	142	79	140	115	110

According to an exploratory investigation, many contaminants seem to be uncorrelated after accounting for greenhouse gas emissions in the mean structure. Consequently, the multivariate spatial model is fitted to groups of contaminants that are linked to one another. As a result, it is difficult to establish which contaminants are highly connected.

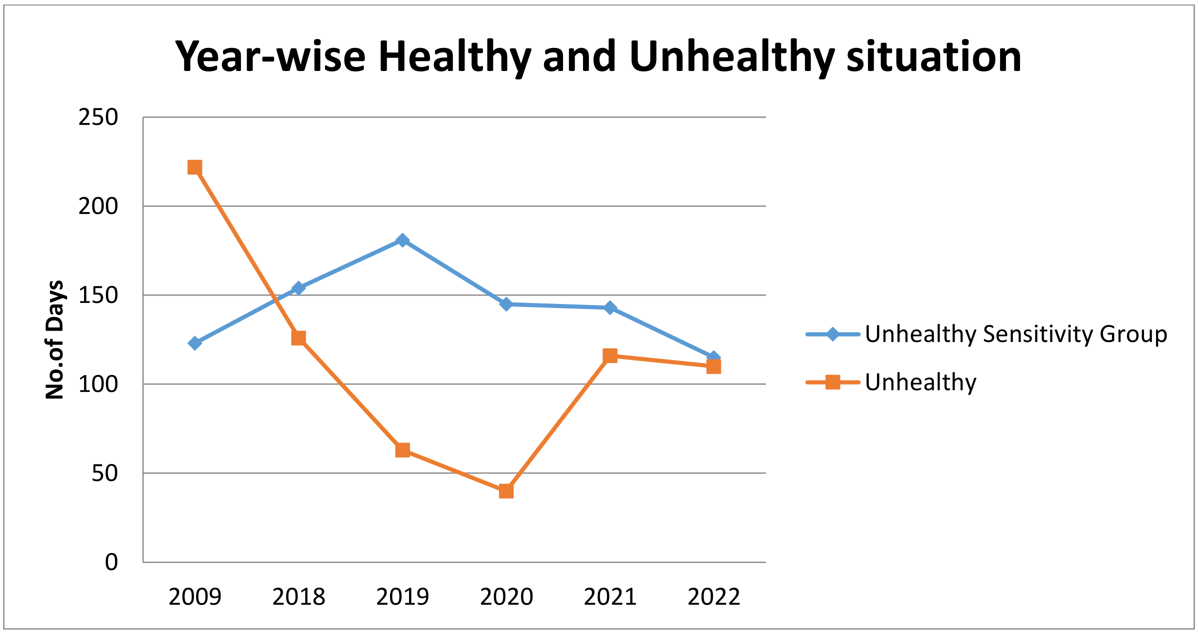

Figure 2. Year-wise Healthy and Unhealthy situation

Consequently, Indian pollution control boards have implemented several legislative and administrative actions in the recent decades. In Table 3, the yearly mean of variation in the air quality of Particulate Matter (PM10) and nitrogen dioxide (NO_2) from 2017 to 2020 is shown. Specifically, during COVID–19 pandemic, a declining trend was found for all air contaminants related to Mumbai city's AQI.

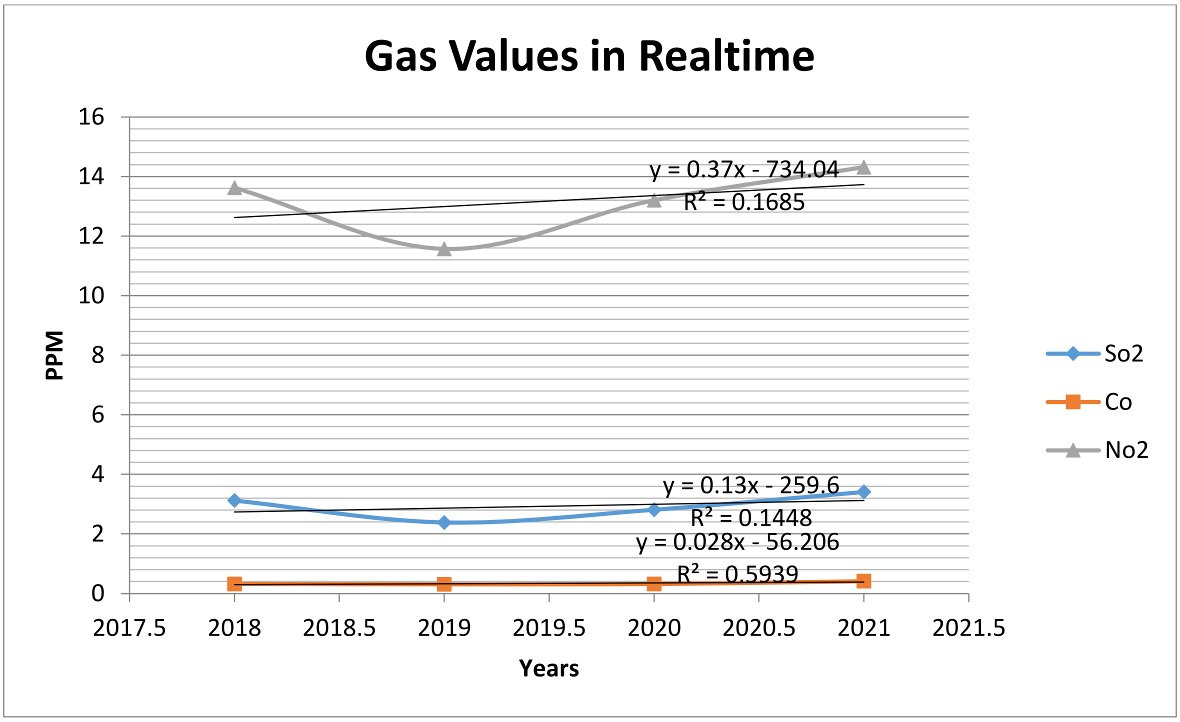

Figure 3. Various gas values in real time

Figure 2-3 depicts the proper prediction parameters for the statistical trend analysis that has been proposed here. The different harmful gas readings rise after the COVID-19 lockdown period, owing to the large number of cars on roads, which caused the values to rise as seen in the graph. For sensitive and ordinary groups of individuals, the chart depicts the number of days during which harmful substances are prevalent in the air as well as for good and unhealthy people.

5. Conclusion

This research article's statistical approach may provide the period's aggressive planning for various equipment and boiler, and air polluted device replacement, such as changing from petrol or diesel engine to natural gas engine. The smoke control equipment is used to increase the air quality in the region with widespread burning boiler devices, and other measures to reduce air pollution must be taken. This research article contains an accurate prediction of air quality in the present and future of Mumbai city, India. The statistical test through the mean and median approach has been conducted for various past resultant data, proving to be accurate.

These steps to improve air quality must be put in place in the future. With the recent establishment of several nations' sustainable development objectives, the competent central agency established the sustainable development targets for air quality and net greenhouse gas emissions shortly. Polluted countries did not see any substantial repercussions from COVID-19, even though it had a worldwide influence on many aspects, including economic activity, dietary habits, air quality, and public health in the year 2020. It is thus possible to achieve sustainable developmental targets within a short period.

References

[1] J. Li and A. D. Heap, “Spatial interpolation methods applied in the environmental sciences: a review,” *Environmental Modelling and Software*, vol. 53, pp. 173–189, 2014.

[2] Anand, C. "Comparison of Stock Price Prediction Models using Pre-trained Neural Networks." Journal of Ubiquitous Computing and Communication Technologies (UCCT) 3, no. 02 (2021): 122-134.

[3] Y. Akita, J. M. Baldasano, R. Beelen et al., “Large scale Air pollution estimation method combining land use regression and chemical transport modeling in a geostatistical framework,” *Environmental Science and Technology*, vol. 48, no. 8,pp. 4452–4459, 2014.

[4] Smys, S., Haoxiang Wang, and Abul Basar. "5G Network Simulation in Smart Cities using Neural Network Algorithm." Journal of Artificial Intelligence 3, no. 01 (2021): 43-52.

[5] Nakajima, T.; Ohara, T.; Masui, T.; Takemura, T.; Yoshimura, K.; Goto, D.; Hanaoka, T.; Itahashi, S.; Kurata, G.; Kurokawa, J.; et al. A development of reduction scenarios of the short-lived climate pollutants (SLCPs) for mitigating global warming and environmental problems. Prog. Earth Planet. Sci. 2020, 7, 33

[6] Karthigaikumar, P. "Industrial Quality Prediction System through Data Mining Algorithm." Journal of Electronics and Informatics 3, no. 2 (2021): 126-137.

[7] Hanaoka, T.; Masui, T. Exploring effective short-lived climate pollutant mitigation scenarios by considering synergies and trade-offs of combinations of air pollutant measures and low carbon measures towards the level of the 2C target in Asia. Environ. Pollut. 2020, 261, 113650.

[8] Pandian, A. Pasumpon. "Artificial intelligence application in smart warehousing environment for automated logistics." Journal of Artificial Intelligence 1, no. 02 (2019): 63-72.

[9] Zhang, H.; Xie, B.; Wang, Z. Effective Radiative Forcing and Climate Response to Short-Lived Climate Pollutants under Different Scenarios. Earth's Futur. 2018, 6, 857–866.

[10] Kottursamy, Kottilingam. "A review on finding efficient approach to detect customer emotion analysis using deep learning analysis." Journal of Trends in Computer Science and Smart Technology 3, no. 2 (2021): 95-113.

[11] Wang, G.; Shen, F.; Yi, H.; Hubert, P.; Deguine, A.; Petitprez, D.; Maamary, R.; Augustin, P.; Fourmentin, M.; Fertein, E.; et al. Laser absorption spectroscopy applied to monitoring of short-lived climate pollutants (SLCPs). J. Mol. Spectrosc. 2018, 348, 142–151.

[12] Vivekanandam, B. "Design an Adaptive Hybrid Approach for Genetic Algorithm to Detect Effective Malware Detection in Android Division." Journal of Ubiquitous Computing and Communication Technologies 3, no. 2 (2021): 135-149.

[13] Kallbekken, S.; Aakre, S. The Potential for Mitigating Short-lived Climate Pollutants. Rev. Environ. Econ. Policy 2018, 12, 264–283.

[14] Shakya, Subarna, and S. Smys. "Big Data Analytics for Improved Risk Management and Customer Segregation in Banking Applications." Journal of ISMAC 3, no. 03 (2021): 235-249.

[15] Liu, P.W.G.; Tsai, J.H.; Lai, H.C.; Tsai, D.M.; Li, L.W. Establishing multiple regression models for ozone sensitivity analysis to temperature variation in Taiwan. Atmos. Environ. 2013, 79, 225–235.

[16] Wang, Y.C.; Liu,W.Y.; Ko, S.H.; Lin, J.C. Tree species diversity and carbon storage in air quality enhancement zones in Taiwan. Aerosol Air Qual. Res. 2015, 15, 1291–1299.

[17] Li, N.; Chen, J.P.; Tsai, I.C.; He, Q.Y.; Chi, S.Y.; Lin, Y.C.; Fu, T.M. Potential impacts of electric vehicles on air quality in Taiwan. Sci.Total Environ. 2016, 566, 919–928.

[18] Sharma, S.; Zhang, Z.; Anshika; Gao, J.; Zhang, H.; Kota, S.H. Effect of restricted emissions during COVID-19 on air quality in India. Sci. Total Environ. 2020, 728, 138878.

[19] Zuma-Netshiukhwi, G.; Stigter, K.; Walker, S., (2013). Use of traditional weather/climate knowledge by farmers in the Southwestern free state of South Africa: Agrometeorological learning by scientists. Atmos., 4(4): 383-410 (28 pages).

[20] Watthanacheewakul, L., (2011). Comparisons of power of parametric and nonparametric test for testing means of several Weibull populations. In Proceedings of the International Multi-Conference of Engineers and Computer Scientist. Hong Kong, 16-18 March.

[21] Haines, A.; Amann, M.; Borgford-Parnell, N.; Leonard, S.; Kuylenstierna, J.; Shindell, D. Short-lived climate pollutants mitigation and sustainable development goals. Nat. Clim. Chang. 2017, 7, 863–869.

[22] Haoxiang, Wang, and S. Smys. "A Survey on Digital Fraud Risk Control Management by Automatic Case Management System." Journal of Electrical Engineering and Automation 3, no. 1 (2021): 1-14.

[23] Bowerman, N.H.A.; Frame, D.J.; Huntingford, C.; Lowe, J.A.; Smith, S.M.; Allen, M.R. The role of short-lived climate pollutants in meeting temperature goals. Nat. Clim. Chang. 2013, 3, 1021–1024.

[24] Manoharan, Samuel. "An improved safety algorithm for artificial intelligence enabled processors in self driving cars." Journal of Artificial Intelligence 1, no. 02 (2019): 95-104.

[25] Gunawan, H.; Bressers, H.; Mohlakoana, N.; Hoppe, T. Incorporating air quality improvement at a local level into climate policy in the transport sector: A case study in Bandung City, Indonesia. Environments 2017, 4, 45.

[26] Balasubramaniam, Vivekanadam. "Artificial Intelligence Algorithm with SVM Classification using Dermascopic Images for Melanoma Diagnosis." Journal of Artificial Intelligence and Capsule Networks 3, no. 1 (2021): 34-42.

[27] Islam, Md Mazharul, Mohammad Shamsul Arefin, Sumi Khatun, Miftahul Jannat Mokarrama, and Atqiya Munawara Mahi. "Developing an IoT Based Water Pollution Monitoring System." In International Conference on Image Processing and Capsule Networks, pp. 561-573. Springer, Cham, 2020.

[28] Rajput, Manita, Mrudula Geddam, Priyanka Jondhale, and Priyanka Sawant. "Smog Detection and Pollution Alert System Using Wireless Sensor Network." In Inventive Communication and Computational Technologies, pp. 1357-1368. Springer, Singapore, 2020.

[29] Patil, Shweta A., and Pradeep Deshpande. "Monitoring Air Pollutants Using Wireless Sensor Networks." In International Conference on Computer Networks and Inventive Communication Technologies, pp. 1-8. Springer, Cham, 2019.

[30] Khairnar, Vaishali, Likhesh Kolhe, Sudhanshu Bhagat, Ronak Sahu, Ankit Kumar, and Sohail Shaikh. "Industrial Automation of Process for Transformer Monitoring System Using IoT Analytics." In Inventive Communication and Computational Technologies, pp. 1191-1200. Springer, Singapore, 2020.

[31] Shilpa, H. L., P. G. Lavanya, and Suresha Mallappa. "Cluster-Based Prediction of Air Quality Index." In Evolutionary Computing and Mobile Sustainable Networks, pp. 899-915. Springer, Singapore, 2021.

[32] “MPCB Home Page: Maharashtra Pollution Control Board.” MPCB Home Page | Maharashtra Pollution Control Board. http://www.mpcb.gov.in/.

Author's biography

S. Kavitha works as a Professor in the Department of ECE at Hindusthan Institute of Technology, Coimbatore, India. Her research are includes WSN, mobile networks, optimization techniques, automation systems, artificial intelligence, computer vision, network security and cloud computing.

J. Manikandan is currently a Professor in the Department of Mechanical Engineering, Hindusthan College of Engineering and Technology, Coimbatore, Tamil Nadu, India. His research interest includes deep learning algorithms, automation systems, process control, computational fluid dynamics, 3D analysis and material sciences.