Grenze International Journal of Engineering and Technology, Jan Issue

Analysis of Pollution Level Near Mining Area using Neural Network

Ramesh Bantwal[1] and K Panduranga Vittal[2]
[1-2]KLS Gogte Institute of Technology, Belgaum, Karnataka, Sr., Member, IEEE

***Abstract*— The curve fitting analysis is performed on time-stamped signal values. Then a situational analysis graphical user interface (GUI) is developed to analyze the data which is from a mining area. The resulting waveform plots and the designed GUI are presented. The data is from the mining area namely Singrauli. Five environmental parameters related to pollution, namely CO, SO_2, PM10, PM2.5, and Ozone are considered for the analysis using neural network.**

***Index Terms*— curve fitting, CO, environmental parameters, graphical user interface, mining, neural network, ozone, pollution, SO_2 , situational analysis**

I. Introduction

The environment pollution control is of paramount importance in the CO analysis of atmospheric conditions. The sad demise of several workers in the city of New Delhi due to carbon monoxide CO inhalation during the year 2019, has led to believe the significance of this study. The measurement and analysis of atmospheric conditions of hazardous industrial environments are very important for the safety of personnel. Safety will be compromised if these pollutants are not monitored in and around hazardous industries.

II. Literature Survey

The global dispersals of CO are presented by Barret et.al. [1] and deliberated with the assistance of simulations. Additional air pollutants emanated comprise SO2 and NO2 produced from the furnace and drying procedures. SO2 is produced from the sulfur composites in the combusted fuel. The burning of coal in revolving cement kilns produces NO2 from the nitrogen in the coal and incoming ignition.
Detecting air pollution from space can observe drifts and deviations of atmospheric toxins affecting human health. The detection of air pollution over the cities, with satellites, has the challenge of low CO sensitivity of measurements [2].
There seems to be a risk of exceeding of health-related exposure levels caused by the current ambient and environmental levels of CO in Milan [3].
The possibility of using satellite images to measure contaminants in natural conditions and large urban areas and even countries are explored. The use of this information helps to reduce the costs and labor compared with land-based techniques. The convenience of satellite services in the terrain helps in reducing the cost of air pollution checks using satellite imagery [5]. CO discharges from vehicles, petrol, and diesel generators, show a concentration of carbon monoxide that is highly contaminated. It is good to stay away from these sources. CO will be emitted into the atmosphere when fossil fuels are used to run the vehicles and generators whereby the unsafe condition prevails [6].

Grenze ID: 01.GIJET.10.1.325_1

For comparison, CO concentrations were measured over continental areas in Brazil. Much larger values have been found in moderate 'burning' regions, such as the south of the state of Mato Grosso and the north-western part of the state of Parana, where 200 ppb in the dry season has been observed. Since normally the air masses have traveled for several days over the ocean, the air masses over the site present low chemical activity. Daily variations of CO2 are very small of the order of a few percent relative to the diurnal mean [7].
The environmental impact of drilling equipment is determined in a surface coal mining operation. The approach used in this paper allowed the determination of exhaust and dust emissions, and sound pressure level. Modeling of the environmental impact of the equipment was conducted using the Microsoft Visual Studio.NET software package [8].
Kasibhatla P, and Arellano, 2002 [9] had a study that emissions of CO from fuel use in China, a country that accounts for a significant fraction of total Asian emissions and where there has been a rapid growth in the industrial and transportation sectors are estimated to have increased by only 16 Tg year between 1990 and 1995 due to almost no change in the relatively large biofuel combustion source. Thus, it is also unlikely that a significant portion of the discrepancy between the top-down and bottom-up estimates in Asia can be explained by growth in Asian CO emissions from 1990 to 1994.
Sarwa et al., 2017 showed the lung performance of non-coal personnel was better than the of coal personnel [10]. The authors recommend further study with regards to this work.

III. METHODOLOGY

A. Neural Network Curve Fitting, Time series analysis, and Situation analysis

Curve fitting is the method of postulating the model that offers the finest fit to the particular curves in the data set. Curved correlations between variables are not as easy to fit and understand as a linear relationship.
For linear relationships, when the independent variable is advanced by one unit, the average of the dependent variable changes by a definite amount.
Prediction can be a linear combination of different features as indicated in equation (1)

$$\hat{y} = \omega 1.x1 + ... + \omega d.\ xd + \beta \ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

Considering the input data values x and the target value $\hat{y}$, one can find a weight vector ω and bias term β that nearly associates data values xi with their targets $\hat{y}$. Mathematically, we can denote the predictions y corresponding to a set of data points x using matrix notation as in equation (2). Equation (3) is used to measure the quality of the model.

$$\hat{y} = x.\omega + \beta \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2)$$

$$E\ (y, \hat{y}) = \sum\_i\hat{}n\blacksquare.\ (\hat{y}i - yi)2 \ldots\ldots\ldots\ldots \quad (3)$$

For example, for the value xi = {0.5,0.6, 0.45, 0.9, ...}, equations (4) and (5) could be written.

$$y1 = 0.5\ x1 + \beta \ldots\ldots\ldots\ldots\ldots\ldots \quad (4)$$

$$y2 = 0.6\ x2 + \beta \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (5)$$

Square Loss: For a one-directional vector, one can determine the relationship between the one feature and the target value. It is also possible to determine a linear predictor and its error. The squared loss penalizes outer data points.
Manipulating the model: To reduce the error, one has to alter the model. This is achieved by choosing values of the parameters ω and β. For example, for 0.0504 as the value for ωi, 0.858 for β, 0.77 for target, the y value is obtained as in equation (6)

$$0.90 = 0.0504 * 0.77 + 0.858 \ldots\ldots\ldots\ldots\ldots \quad (6)$$

1) Neural Network

For linear models, the simplest possible useful neural network is to determine the output of the model by multiplying the input with the model's weight ω and the offset β is added.
Training the model makes it better throughout a period of training. To find out if the goal has been reached or not, the squared distance between the prediction and the true value is calculated.

2) Curved relationship

The data may have curved relationships among variables. In a curved relationship, the alteration in the dependent variable associated with a single unit change in the independent variable varies based on the position in the observation space. The curve fitting figure is to obtain a mathematical representation that fits the data. Each row in the input data 'x' comprises sensor data and the rows in the 'y' comprise the target value. For a given point, a linear combination yields a prediction that is close to the target value. The curve fit finds the parameters with which the correlation between the function and the data is as close as possible.
The data set x could be the multi-dimensional input array if more than one feature is chosen, and the y matrix should be of the same length as that of matrix x.
The dataset, which is obtained for the air pollution, the curve fitting structure Figure is developed for 2150 samples of carbon monoxide, 2150 samples of SO2, 2150 samples of PM10, 1008 samples of PM2.5, and 1008 samples of ozone for the year 2019 and 2018

B. Time series analysis and prediction

To forecast upcoming values of a time series p(t) (equation (7)), from previous values of that time series and previous values of another time series q(t) is called non-linear auto-regression with external input.

$$p(t) = f(p(t+1) \ldots p(t-d), q(t-1), q(t-d)) \ldots\ldots\ldots\ldots\ldots\ldots (7)$$

The neural network structure used for the mining environment of Singrauli (which is in Madhya Pradesh, India), for all the five parameters, is shown in Figure. The hidden layer comprises 10 neurons. The step ahead analysis can be performed using this network.

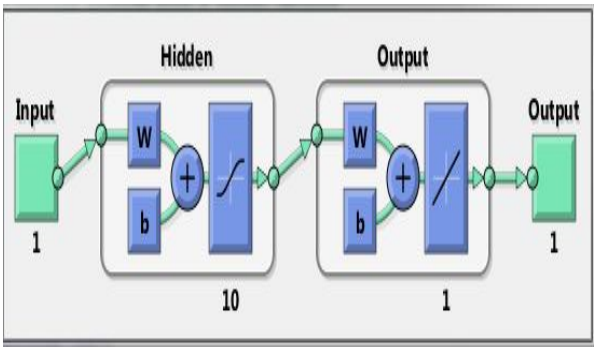

Fig. 1. Neural network structure

The prediction is performed with 300 hidden units which provides root mean square error (RMSE) and the prediction plots. The closed-loop structure for the analysis is indicated in Fig. 1.

C. Graphical User Interface

MATLAB Guide is used to develop the graphical user interface (GUI). The different UI controls used for the applications are pushbuttons, UI editable text, UI static text, Frames, and axis. The 'Guide' is composed of the Guide Control Panel and supporting tool, guide property editor. The GUI figure window is created using Guide and the various UI objects are fine-tuned and are automatically connected to the callback handling function when saved.
The five parameter plots are analyzed in a designed graphical user interface (GUI) and the analysis push buttons will analyze the parameters.
The carbon monoxide (CO), Sulfur dioxide (SO2), PM10, PM2.5, and Ozone data are analyzed for the data of Singrauli of Madhya Pradesh state.
Another GUI is developed to analyze the data and to check if the data value exceeds the standards and how many times it crosses the threshold. In this GUI, the push button callback is connected to the different analysis programs.

D. Performance Evaluation of Situational Awareness Platform in Mining Industry

The software data analytics is used to analyze the five parameters. The sensor data for the five parameters CO, SO2, PM10, PM2.5, and Ozone were collected from the Central Pollution Control Board (CPCB) website (CPCB, 2018) making use of an advanced search method for Singrauli where open cast mine is located For Singrauli, data were collected both for the year 2018 and 2019. The data from Bangalore city is taken as the target data as the CO and SO2 are within the limit set by Indian standards.
The standard set by CPCB is used to compare the measured data with the standards.

The five parameters namely CO, SO2, PM10, PM2.5, and Ozone are plotted using the designed GUI.
Then parameters were analyzed using the curve fitting method which gives the curve fitting error and time series analysis gives regression analysis result. By this, the data clustering, as well as regression analysis, are performed and abnormalities observed. Finally, the GUI is designed for five parameter analysis which gives the number of samples that are exceeding the standards.

IV. Performance Evaluation of Situational Awareness Platform in Mining Industry

The neural network curve fitting is used to check the abnormalities in the data. Using this method, the curve fitting analysis is performed on the different data.
The curve fitting analysis for Ozone produces the plot as in Fig.2. This plot is taken as a sample of curve fitting analysis.

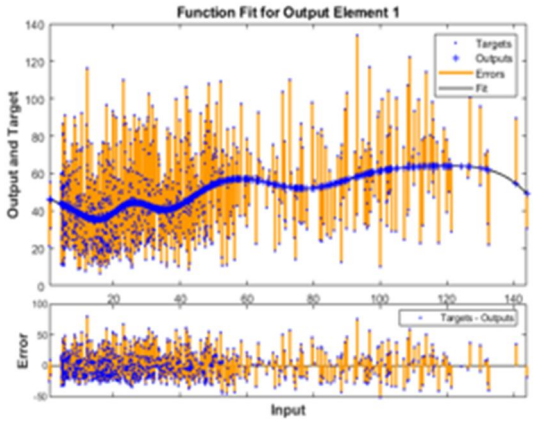

Fig.2. Ozone curve fit

From this figure, one can observe that the Ozone is gradually increasing over time. All the instances of parameters exceeding the standard values are listed in Table I.
The comparison between the 2018 and 2019 situation for Singrauli is indicated in Table.1. There is an increase in SO2 level from the year 2018 to the year 2019. The PM10 parameter remained high during both these years, while PM2.5 remained marginally high during these two years.

TABLE I. AIR QUALITY AS MEASURED FOR SINGRAULI

S N	Year	2018	2019
1	Co Exceedance	1	1
2	So_2 Exceedance	4	98
3	Pm10 Exceedance	283	240
4	Pm2.5 Exceedance	26	25
5	Ozone Exceedance	0	1

A. Regression Analysis

From the regression graph in Fig.2 one can see the errors and deviation of the data from normal data like measured data for Bangalore which gives the value of 0.46 for ‘m’, the value of 0.57 for ‘b’, and the value of 0.6729 for correlation coefficient ‘R’. This plot is shown as one sample of the regression plots obtained. For all the parameters, the regression values are listed in Table 1.2 for the year 2018-19. If the regression value is higher, the data will be even closer to the target value.
From the Table, the variation in CO and Ozone are within normal variation, but error points are higher for SO2, PM10, and PM2.5.

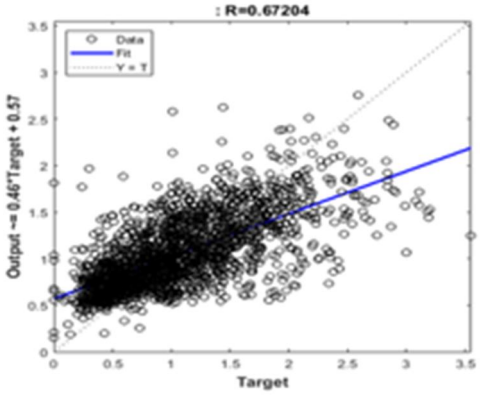

Fig.3. Regression plot

B. Result Table for Singrauli

The correction coefficient or regression R is listed in Table II for the year 2018 to 2019. Comparing the data with data from Bengaluru, the table of regression Table III is obtained.

TABLE II. REGRESSION R FOR SINGRAULI FOR 2018-19

SN	Parameter	m	b	r
1	CO	0.46	0.57	0.67
2	SO2	0.36	13	0.6005
3	PM10	0.57	86	0.7575
4	PM2.5	0.67	29	0.8188
5	Ozone	0.36	17	0.589

When the regression R values are compared from the two tables, one can observe that the second table has higher regression values. This will indicate that some of the parameters are increasing both in the mine area and city area of the country.

TABLE III. REGRESSION R FOR SINGRAULI FOR 2019

SN	Parameter	m	b	r
1	CO	0.51	0.23	0.7139
2	SO2	0.6	18	0.769
3	PM10	0.8	17	0.8925
4	PM2.5	0.7	8.9	0.85
5	Ozone	0.58	19	0.759

C. Situation Awareness for Singrauli

The GUI gives output specifying that the parameter has exceeded the standard and how many times it has crossed the standard. The situation analysis performed on the Singrauli data using the developed GUI results in Fig. 4 for the year 2019.

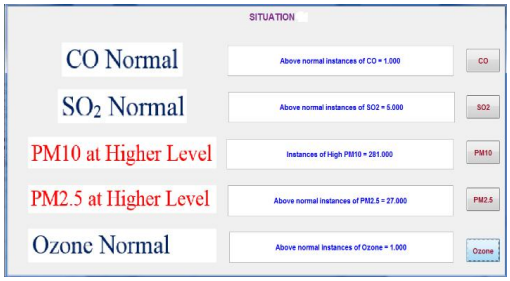

Fig.4.Situation analysis performed on the Singrauli

The neural network-based data analytics is used to perform the analysis on the environmental parameters. The different techniques like Curve fitting, Time-series, and Regression analysis are used to perform the analysis on the environmental parameters like CO, SO2, PM10, PM2.5, and Ozone. The curve fitting method will provide the errors and the trend. One trend observed is that the Ozone is gradually increasing. All the parameters are listed in the table form. The time-series analysis is used for the regression.

The regression analysis by the time series has provided better results for the correlation coefficient or regression R. All the regression values are listed in the table form.

The Graphical user interface is developed. A GUI developed analyzes the data vector and gives the situation as output by comparing it with the standards. The real data from the CPCB website is used to analyze the situation of the air quality. This GUI can be used to analyze any other parameters related to mining.

From the data analytics used, there is a pollution of both SO2 and PM surrounding the Singrauli area. One could observe that there is worsening of SO2 pollution level from the year 2018 to 2019 in the Singrauli area. There is no much difference in pollution levels of CO, PM2.5, and Ozone from the year 2018 to the year 2019.

Surprisingly, the regression fit is better when Singrauli data is compared with city data for the same year compared to the previous year's data. This will conclude that some of the parameters are increasing both in the city and in the mining site.

ACKNOWLEDGEMENTS

The authors acknowledge the resources provided by the parent institutions.

REFERENCES

[1] B.Barret, et al., “Global carbon monoxide vertical distributions from spaceborne high-resolution FTIR nadir measurements,” Journal of Atmosheric. Chemistry and Physics, 5, pp. 2901–2914, 2015.

[2] T.Borsdorff et al. , “Mapping carbon monoxide pollution from space down to city scales with daily global coverage,” Atmospheric Measurement Techniques. 11, pp. 5507–5518, 2018.

[3] Y.Bruinen et al., “Simulation of working population exposures to carbon monoxide using Expolis-Milan microenvironment concentration and time-activity data,” Journal of Exposure Analysis and Environmental Epidemiology, 14, 154–163, 2018.

[4] Central Pollution Control Board (CPCB), "Northern Coalfields Limited", Retrieved from http://nclcil.in, 2019.

[5] M.Hamzeloa, A.Gharagozloub, B.S.Sadeghian, S.H.Baikpoura, C.A.Rajabi,. “Modelling of Carbon Monoxide air pollution in large cities by evaluation of spectral lands at images,” The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-1/W5, pp. 281-285, 2015.

[6] K.S.John, K.Feyisayo, “Air Pollution by Carbon Monoxide (CO) Poisonous Gas in Lagos Area of Southwestern Nigeria”, Atmospheric and Climate Sciences, 3, pp. 510-514, 2013.

[7] V.W.J.H.Kirchhoff, C.B.Aires, and P.C.Alvala,. “An experiment to determine atmospheric CO concentrations of tropical South Atlantic air samples,” Q. J. R. Meteorol. Soc., 129, pp. 1891–1902, 2003.

[8] A.Lashgari and V.Kecojevic , “Assessment of Environmental Impact of Drilling Equipment,” J Powder Metallurgy & Mining, S:1, pp. 1-3, 2013.

[9] P.Kasibhatla , and A.Arellano ,” Top-down estimate of a large source of atmospheric carbon monoxide associated with fuel combustion in Asia,” Geophysical Research Letters, 29(19), pp. 6.1-6.4, 2002.

[10] G.Sarwa et al., “Spirometric Evaluation of Lung Function of Coal Workers working at Mach (Bolan District),” Indo American Journal of Pharmaceutical Sciences, 4 (11), pp. 3850-3855, 2017.

The name of the corresponding author: Dr. Ramesh Bantwal.

First R.Bantwal is with the KLS Gogte Institute of Technology, Belagavi (e-mail: rbantg5@gmail.com).

Second K. Panduranga Vittal, is with National Institute of Technology Karnataka, India. Surathkal, India. He is senior member of IEEE. (e-mail: vittal.nitk@gmail.com).