NIPES Journal of Science and Technology Research 3(3) 2021 pp.125 - 137 pISSN-2682-5821, eISSN-2682-5821

Journal of Science and Technology Research

Journal homepage: www.nipesjournals.org.ng

Investigation and Modelling of Ughelli West Gas Plant Gaseous Pollutants

Okpoko, J. S[1] and Ilaboya, I.R[2]

[1]Department of Civil Engineering, Benson Idahosa University, Benin City, Edo State
[2]Department of Civil Engineering, University of Benin, Benin City, Edo State

Article Info

Keywords:
Climatic Variables, Gaseous Pollutants, Gas plant, Modelling, mathematical models.

Received 26 July 2021
Revised 12 August 2021
Accepted 14 August 2021
Available online 31 August 2021

https://doi.org/10.37933/nipes/3.3.2021.13

https://nipesjournals.org.ng

Abstract

Gaseous contamination of environment poses great challenges to human health. Inadequate models to predict the concentration of gaseous pollutants had led to the poor control and monitoring of gas plant gaseous pollutants. This study focused on the modelling and analysis of Ughelli West gas plant gaseous pollutants located in Delta State, Nigeria. Aeroqual multi-parameter environmental monitor (series 500), was employed to monitor the concentrations of volatile organic compounds (VOCs), oxides of nitrogen (NO_2), oxides of sulphur (SO_2), ozone (O_3) and methane (CH_4). The concentrations of the particulate matter ($PM_{2.5}$ of these gases were obtained at each monitoring point on daily bases for a period of twelve weeks using Aerocet-531 SPM meter. Sky master thermo anemometer (SM-28) was used to obtain the important climatic variables (wind speed, atmospheric pressure, ambient temperature and relative humidity) which affect the dispersion of gaseous pollutants. The maximum concentration of each monitored gaseous pollutant during the twelve weeks (12) monitoring period was selected and recorded for data processing. In this study, mathematical models were developed for predicting each gaseous pollutant such as volatile organic compounds (VOCs), oxides of nitrogen (NO_2), oxides of Sulphur (SO_2), ozone (O_3) and methane (CH_4). The curve fitting tool in Matrix Laboratory {MATLAB (2016a)} was employed to select models and to model the exact mathematical relationship between the pollutant concentrations and the flare distance; then the pollutants concentrations were predicted beyond the experimental distance of 500m from the flare point, the models were validated using coefficient of determination (R^2), and root mean square error (RMSE). Based on the parameters, it was observed that the quadratic polynomial model had the lowest root mean square error value of 0.6053 and coefficient of determination R^2 value of 0.9919. The results obtained from the validation shows good predictability and adequacy of the models developed in petroleum drilling processes.

1. Introduction

Air is one of the five essentials of life (air, water, food, heat and light) for human beings. Man breathes nearly 22,000 times a day and inhales approximately 15kg of air per day [1]. The amount of air inhaled per day is far more than the quantity of food and water consumed by human per day all put together [2]. Air pollution is the contamination of the atmosphere by gaseous, liquid or solid wastes or their by-products, including noise present in the atmosphere in concentrations that can endanger human health and the health and welfare of plants and animals, or can attack materials, reduce visibility, or produce undesirable odour. The products (pollutants) from burning of fossil fuels such as gas, oil, coal and wood affect the earth, buildings, water and air, resulting in fog, smog and global warming, which deteriorate vegetation, forests, and even human health [2, 3, 4, 5]. Thousands of people in the world die each year due to heart and lung diseases that result from air pollution [6]. Air pollution which is any atmospheric condition in which certain substances (pollutants) are present in such concentrations that may produce undesirable effects on man and ecosystems, has become an extremely serious problem for the modern industrialized world [7]. Within the Niger Delta area (especially Niger Delta Western Operation), is the Oil Mining Lease (OML) 34 comprising of Utorogu field and Ughelli East and West fields. The OML 34 field which is the case study area for this research work involves large scale combustion of gas by means of horizontal and vertical flares in Ughelli East. The consequential gaseous pollutants namely methane (CH_4), ammonia (NH_3), oxides of nitrogen (NO_2), particulate matters (PM), oxides of Sulphur (SO_2) and volatile organic compounds (VOCs) which are emitted from the flare sites are hazardous to man, animals and the total environment. Gaseous pollutants contamination in the Niger Delta area of Nigeria is a serious environmental concern for the entire area in particular and for Nigeria and the world in general since air pollution is never restricted to political boundaries. The increasing population and industrial growth as well as commercial operations in the Niger Delta area due mainly to oil and gas operations and processes brings to the fore the need for air pollution monitoring through assessment of pollutants concentrations, modelling of dispersion and prediction of pollutants spread, to help regulate and manage environmental impacts. Dispersion modelling is undertaken in order to predict the concentration and spread of pollutants [8]. Douabul et al [9] carried out a study on gaseous pollutants in Basra city, Iraq. Their study was aimed at detecting the present levels and distribution of CO, CO_2, SO_2, NO_2 and total hydrocarbons gases (HC_s) produced from different industrial plants in Basra city, Iraq. Seven stations were chosen in Basra city – Al – Qurna, Al – Deer, Garmatt Ali, Al - Ashar, Abu Al-Khaseeb, Al- Seeba and Al- Faw. These stations were selected in order to monitor the concentrations of CO, CO_2, SO_2, NO_2 and HC_S in the ambient air during the winter and summer months of 2011. Grigoras et al [10] researched on air pollution dispersion modelling in a polluted industrial area of complex terrain in Romania (the surveyed North-Western part of Romania). They made use of pollutants emitted by non-ferrous metal industrial facilities existing in Baia Mare area and the emissions from other local anthropic activities such as residential heating, traffic and dump heaps. The air pollution modelling was done by using local emissions inventories drafted and validated for 2008 by WESTAGEM and by real-time monitoring results of emissions from lead smelting from lead concentrates and metallurgical residue with lead content. They performed atmospheric pollutants concentrations assessment for SO_2, particulate matters PM_{10} and lead (Pb), referencing to the limits for allowable concentration of pollutants in the ambient air by using 1-hourly and daily maximum concentration values and average concentration values for the air pollutants originated from local sources. The atmospheric emissions assessment was made considering temporal variations of the activities that lead to the time variation of emissions. Gorai et al, [11] carried out the development of Partial Least Squares Path Model (PLS-PM) to understand the role of precursors on ground level ozone concentration in Gulfport, Mississipi, USA. Their model revealed that Photochemical Reaction Catalyst (PRC) had significant direct impact on ground level ozone concentration, but very small overall effect since PRC had significant indirect effect via metheorological factor. They concluded that the direct and indirect effects have made PRC to have the weakest effect on ground level ozone. Yannarwar et al [12] had the opinion that dispersion models are used to predict the fate of pollutants after they are released into the atmosphere. The goal of air quality dispersion modelling is to estimate a pollutant's concentration at a point downwind of one or more emission sources [13]. The first step in the modelling and prediction of ground level concentration of gaseous pollutants is to understand the exact mathematical relationship between the pollutant concentrations and the distance from flow station at normal environmental stability and wind speed.

This study therefore focused on the modelling and investigation of Ughelli West gas plant gaseous pollutants.

2. Methodology

2.1 The study Area

In Niger Delta area, there are many natural gas processing plants which emit the common harmful gaseous pollutants into the atmosphere through continuous gas flaring. Example of such natural gas processing facilities for a case study is the OML 34. The OML 34 (Oil Mining Lease 34) is located in the Western Niger Delta, and covers an area of about 950km^2. The producing fields is Ughelli West, with a total flow station capacity of 33mbpd. The Ughelli West facility is located at Ughelli West Local Government Area, in Delta State of Nigeria.

OML 34 is of utmost strategic importance to Nigeria and the West African sub-region, as a major supplier of gas for electricity generation in Nigeria. It also feeds gas through the West African Gas Pipeline (WAGP) to neighbouring countries [14]. Currently most of the exploration and production activities are focused within Ughelli West field. OML 34 is key to the nation's power generation. Gas from OML 34 is produced from its field, the Ughelli West field. The Ughelli West gas processing plant have been in existence for many years came with an installed capacity of 90mmscf/d. Niger Delta Western is a major supplier of gas to the domestic market and the West African Sub-region which include Ghana, Lome and the Benin Republic via the West African Gas Pipeline [14, 15].

2.2 Procedure for Data Analysis

This study involves focused on the modelling and investigation of Ughelli West gas plant gaseous pollutants from OML 34 in the Niger Delta area of Nigeria, was carried out using the following steps: -

1. Data acquisition and processing.
2. Computation of pollution standard index (PSI).
3. Simulation of environmental condition based on the computed PSI.
4. Geo-statistical analysis of the data
5. Modelling and Data analysis concentration of pollutants
6. Model validation.

2.2.1 Data Acquisition and Processing

In this study, six (6) gaseous pollutants namely; volatile organic compounds (VOCs), methane (CH_4), nitrogen dioxide (NO_2), particulate matter ($PM_{2.5}$), ozone (O_3) and Sulphur dioxide (SO_2) were monitored on daily bases for a period of twelve (12) weeks and data were transformed into weekly maximum concentration. To select the maximum concentration of the pollutant at each sampling point within the entire period of experimentation for the modelling, extreme value statistics was carried out using the data analysis tool pack of Microsoft Excel 2010. The mandatory frequency of sampling for the gaseous point source emission monitoring shall be weekly but where appropriate, a continuous emission monitoring system approved by the Director of Petroleum Resources (DPR) shall be utilized [14]. According to the descriptive statistics of data stated in Environmental Guidelines and Standards which stated that gaseous point sources emission monitoring shall be at distances of 200m intervals away from the installation along the direction of the prevailing wind [8]. Based on this and to determine the trend, the monitoring locations were established and the range of measurement was 60m to 500m away from the flare point at each station using a spacing distance of 60m, 80m, 100m, 150m, 200m, 250m, 300m, 350m, 400m, 450m and 500m from the flare point. Standard gaseous pollutants monitoring equipment used such as Gas monitor, SPM meter, Anemometer and GPS receiver were calibrated and used as follows. Aeroqual multi-parameter environmental monitor (series 500), having different gas sensors, was employed to monitor the concentration of volatile organic compounds (VOCs), oxides of nitrogen (NO_2), oxides of sulphur (SO_2), ozone (O_3) and methane (CH_4). The concentrations of these gases were obtained at monitoring location of Ughelli West flow station on daily bases for a period of twelve weeks. Aerocet-531 SPM meter was used to monitor the concentration of particulate matter ($PM_{2.5}$) at the location on daily bases for a period of twelve weeks. Sky master thermo anemometer (SM-28) was used to obtain the important climatic variables (wind speed, atmospheric pressure, ambient temperature and relative humidity) which affect the dispersion of gaseous pollutants. An updated map of the OML 34 was sourced from Shell Petroleum Development Corporation (SPDC), Delta State, Nigeria was used for this study. The Global Positioning System (GPS) receivers and point positioning techniques were used to obtain the geographical coordinates at each monitoring location in the study area. The coordinates were converted to decimal degrees format using the Universal Traverse Mercator (UTM) software version 1.0. The maximum concentration of each monitored pollutant during the twelve weeks monitoring period was selected and recorded for data processing. The data obtained in parts per million (ppm) were processed by converting the pollutants concentrations from ppm to mg/m^3 or $\mu g/m^3$, using the model presented below (Equation 1). This is because the pollutants hourly or daily or annually concentrations are measured in $\mu g/m^3$ [14, 15].

$$\text{Concentration in } mg/m^3 \text{ or } \mu g/m^3 = \frac{\text{conc.(ppm) x MW (g/mole)}}{\text{MV (L)}} \tag{1}$$

where:

Conc. = Concentration of pollutant
mg/m^3 = milligram per cubic meter = $10^{-3} g/m^3$
$\mu g/m^3$ = microgram per cubic meter = $10^{-6} g/m^3$

3. Results And Discussion

Table 1 shows the input parameter of the flow station presented to the neural network for validation of field data. The input parameters include; the sampling distance from flow station, wind speed, atmospheric pressure, ambient temperature and relative humidity. These parameters were selected because they form the critical climatic variables that affects the dispersion of gaseous pollutants.

Table 1: Input Parameters for Ughelli West Flow Station

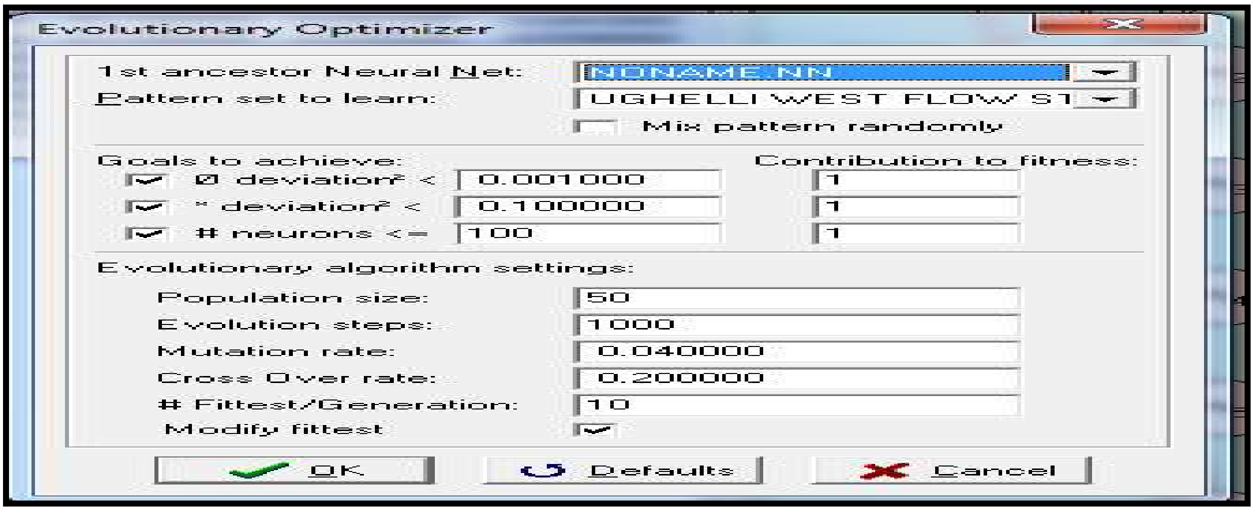
Sampling Distance (m)	Wind Speed (m/s)	Atm. Pressure (mmHg)	Ambient Temperature (deg C)	Relative Humidity (%)
60	4.1	1010	29.7	88.7
80	3.2	1010	29.4	91.2
100	3.4	1010	28.5	78.9
150	3.7	1010	29.1	88.3
200	3.9	1010	29.5	88.6
250	4.2	1010	27.4	75.8
300	3.7	1010	28.8	79.2
350	3.1	1010	26.7	88.1
400	2.8	1010	27.8	79.9
450	2.8	1010	28.6	82.5
500	2.5	1010	27.4	78.9

.

3.1 Validation of Field Data using Pythia (Neural Net Software)

The field data from the study locations were validated using artificial neural network designer software (Pythia). The field data were validated to evaluate the adequacy of the field data for use in air pollution modelling. The artificial neural network is trained to prevent it from memorizing data presented before it. Using the training data which is the field data collected, Pythia employed the evolutionary optimizer to search the neural network topology that best understand the input and output data presented for training, and the back propagation algorithm to produce the network. The reason for training is for the model to understand the data obtained and the condition under which they were obtained, so as to be able to make accurate correction. The criteria for selecting the best topology are the square of the deviation between the observed output and the predicted output coupled with the fitness accuracy. Topology with the least square deviation having 100% fitness was regarded as the best for the task. Pythia employs the back propagation algorithm to produce the network. During the training phase, the actual output of the network was compared with the experimental output and the error propagated back towards the input of the network. The network parameter which is the input is also called the "weight".

Table 2: Condition for best performance



Based on the conditions of Table 2, the evolutionary optimization was performed to obtain the best neural network topology that best fits the input and output data for Ughelli West. The evolutionary optimization is presented in Table 2

Table 3: Evolutionary Optimization for Selecting Best Network Topology

Evolutionary Optimization (Generation 0)

Ancestor Net: (5,6,6), 'NONAME.NN' Pattern Set: '(no name)'
Goals: (Ø deviation² < 0.001000, 33.33%) AND (* deviation² < 0.100000, 33.33%) AND (#Neurons < 100, 33.33%)
GA settings: 1000 gen max, pop size 50, mutation rate 0.04, crossover rate 0.20, keep best 10 (modif)

No	Topology	Neurons	Ø dev²	* dev²	Fitness
31	5,6,7,6,6,6	31	0.000305	0.000868	100.00000
32	5,6,6,6,6,6	30	0.000524	0.001431	100.00000
33	5,6,6,6,6,6	30	0.000479	0.000941	100.00000
34	5,6,6,6,6	24	0.000588	0.002221	100.00000
35	5,6,6,6	18	0.000765	0.002121	100.00000
36	5,6,5,6	17	0.001469	0.004796	89.36437
37	5,7,6	13	0.000805	0.001918	100.00000
38	5,6,6	12	0.002072	0.007394	82.75606
39	5,6,5,6	17	0.000609	0.001953	100.00000
40	5,6,6	12	0.001338	0.005047	91.58726
41	5,6,5,6	17	0.001399	0.004289	90.48680
42	5,5,5,5,6	21	0.000546	0.001897	100.00000
43	5,5,5,4,6	20	0.001216	0.004123	94.07202
44	5,5,5,4,4,6	24	0.000751	0.002169	100.00000
45	5,5,5,4,4,6	24	0.014059	0.141193	59.31273
46	5,5,4,3,6	18	0.001790	0.004660	85.29146
47	5,5,4,3,6	18	0.003841	0.011044	75.34487
48	5,4,3,6	13	0.003315	0.015837	76.72220
49	5,4,4,6	14	0.003555	0.015841	76.04195

Ok Stop Cancel

Table 4: Evolutionary Optimization for Selecting Best Network Topology

Evolutionary Optimization (Generation 0)

Ancestor Net: (5,6,6), 'NONAME.NN' Pattern Set: 'C:\Users\User\Desktop\UGHELLI WEST FLOW STATION.PAT'
Goals: (Ø deviation² < 0.001000, 33.33%) AND (* deviation² < 0.100000, 33.33%) AND (#Neurons < 100, 33.33%)
GA settings: 1000 gen max, pop size 50, mutation rate 0.04, crossover rate 0.20, keep best 10 (modif)

No	Topology	Neurons	Ø dev²	* dev²	Fitness
11	5,6,6	12	0.000556	0.002202	100.00000
12	5,6,6	12	0.001569	0.007560	87.91357
13	5,6,6	12	0.000341	0.001055	100.00000
14	5,6,6	12	0.000742	0.002033	100.00000
15	5,6	6	0.012439	0.041270	69.34650
16	5,6	6	0.012439	0.041270	69.34650
17	5,6	6	0.012263	0.030285	69.38497
18	5,4,6	10	0.003074	0.014690	77.50957
19	5,6,3,6	15	0.001958	0.008304	83.69118
20	5,6,3,6	15	0.002033	0.005777	83.06266
21	5,6,3,6	15	0.004217	0.014985	74.57049
22	5,6,7,3,6	22	0.001040	0.002574	98.70691
23	5,7,3,6	16	0.002606	0.014515	79.45930
24	5,3,6	9	0.003782	0.028671	75.48071
25	5,4,3,6	13	0.002669	0.014925	79.15539
26	5,3,6	9	0.012618	0.049759	69.30848
27	5,3,2,6	11	0.011970	0.058453	69.45139
28	5,3,2,6	11	0.004850	0.023265	73.54003
29	5,2,6	8	0.013061	0.067779	69.21887
30	5,3,6	9	0.011080	0.062930	69.67502

Ok Stop Cancel

From Table 4, it was observed that a minimum of 12 neurons was needed to obtain an optimum topology of 5, 6, 6 having a fitness of 100% with a square deviation of 0.001055 for Ughelli West. Based on these parameters, optimum neural network architecture were produced as shown in Table 5.

Using the network of Table 5, the Repro Pattern Set function of the Pythia program was then activated to predict the pollutants concentrations based on the input and output data from Ughelli West flow stations. Reliability plots of the field data and the ANN predicted data were obtained to test the correlation between the field data and the ANN predicted data.

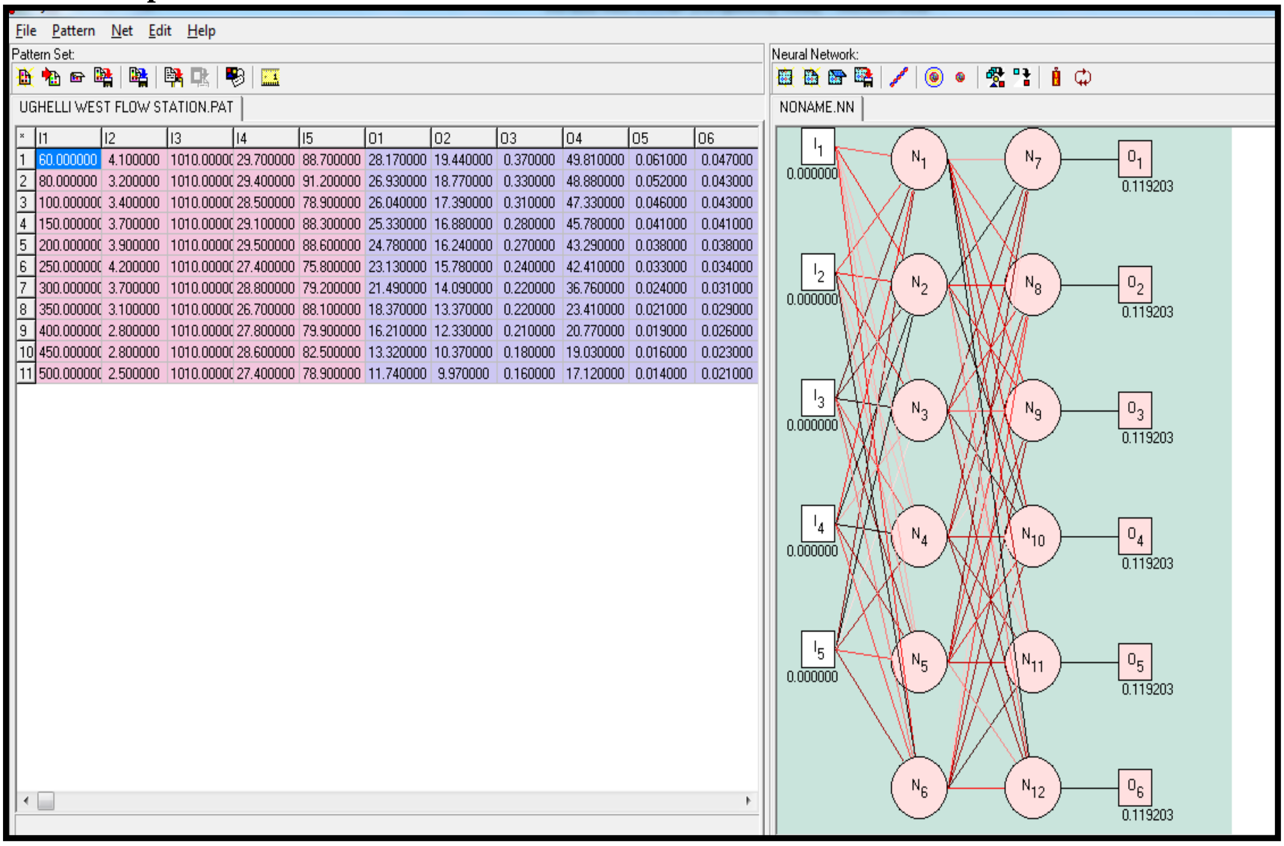
3.2 Sensitivity analysis of ANN

A sensitivity analysis was performed to allow the network assign weight to each input variable on the bases of their significant contribution so as to determine the input parameter that contributes mostly to variation in the pollutant concentrations around the study locations.

3.3 Selection of Trend Analysis Model using Normality Test

In order to check the distribution of the field data, normality test was also employed to select the most appropriate model for trend detection and estimation. If data are linearly distributed then parametric model (linear regression model) such as least square linear regression will be most appropriate for trend detection and estimation otherwise non-parametric model such as Mann-Kendall and Thiel Sen's slope estimation will be employed for detection and estimation of trend in the data. To test if the field data followed a normal distribution, histogram plot and normal Q-Q plot was employed. For normality, the histogram must be assumed to have a bell shape configuration and the data points on the normal Q-Q plot must follow the 45^0 center line else it will be concluded that the data are not normally distributed. Consequently, non-parametric analysis will be needed to determine the occurrence of trend in the data [6]. The practical implication of trend detection is to know exactly what is happening around the study location.

Table 5: Optimum Neural Network Architecture



*	I1	I2	I3	I4	I5	O1	O2	O3	O4	O5	O6
1	60.000000	4.100000	1010.00000	29.700000	88.700000	28.170000	19.440000	0.370000	49.810000	0.061000	0.047000
2	80.000000	3.200000	1010.00000	29.400000	91.200000	26.930000	18.770000	0.330000	48.880000	0.052000	0.043000
3	100.000000	3.400000	1010.00000	28.500000	78.900000	26.040000	17.390000	0.310000	47.330000	0.046000	0.043000
4	150.000000	3.700000	1010.00000	29.100000	88.300000	25.330000	16.880000	0.280000	45.780000	0.041000	0.041000
5	200.000000	3.900000	1010.00000	29.500000	88.600000	24.780000	16.240000	0.270000	43.290000	0.038000	0.038000
6	250.000000	4.200000	1010.00000	27.400000	75.800000	23.130000	15.780000	0.240000	42.410000	0.033000	0.034000
7	300.000000	3.700000	1010.00000	28.800000	79.200000	21.490000	14.090000	0.220000	36.760000	0.024000	0.031000
8	350.000000	3.100000	1010.00000	26.700000	88.100000	18.370000	13.370000	0.220000	23.410000	0.021000	0.029000
9	400.000000	2.800000	1010.00000	27.800000	79.900000	16.210000	12.330000	0.210000	20.770000	0.019000	0.026000
10	450.000000	2.800000	1010.00000	28.600000	82.500000	13.320000	10.370000	0.180000	19.030000	0.016000	0.023000
11	500.000000	2.500000	1010.00000	27.400000	78.900000	11.740000	9.970000	0.160000	17.120000	0.014000	0.021000

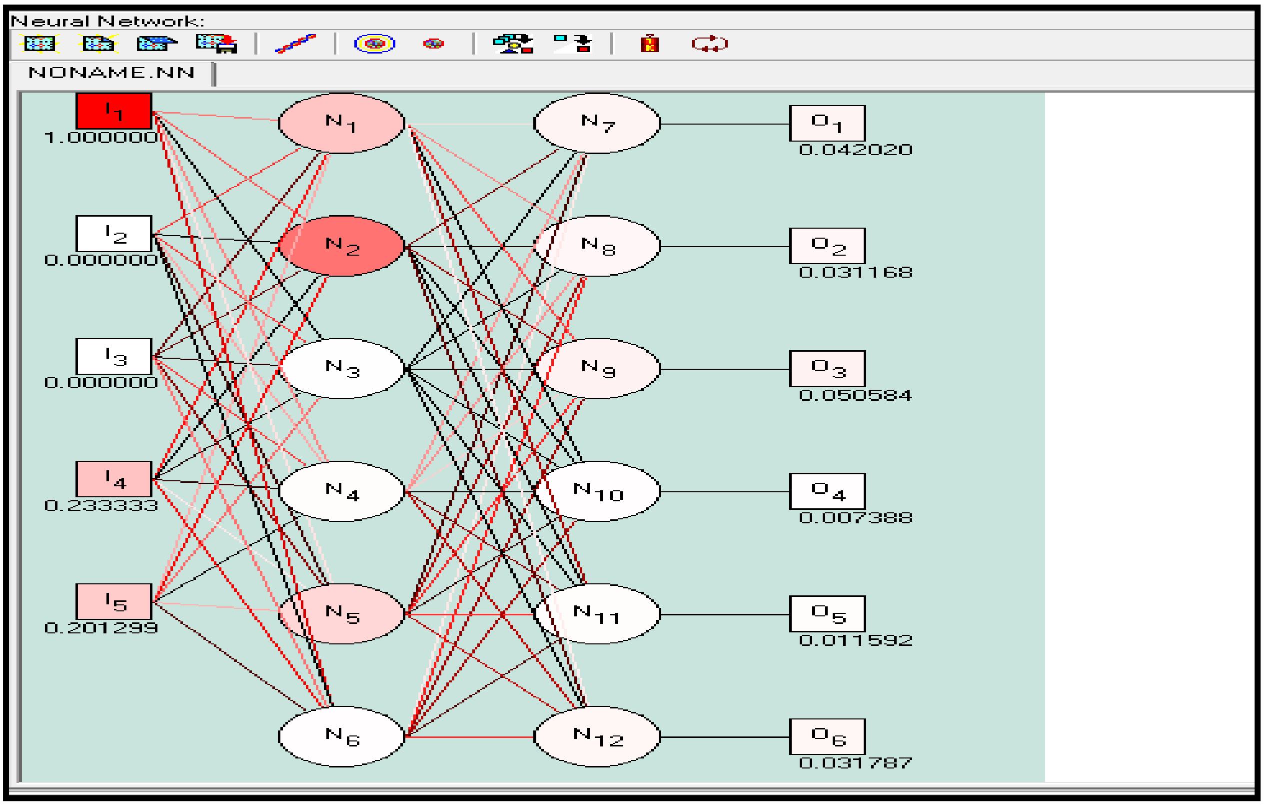


Figure 3: Sensitivity Analysis of ANN (Ughelli West)

Table 6: Calculated model parameters for VOC (Ughelli West)

Model	R-Square	Adj. R-Square	SSE	RMSE
Linear polynomial	0.9670	0.9633	10.64	1.087
Quadratic Polynomial	0.9919	0.9886	2.932	0.6053
Exponential	0.9282	0.9202	23.16	1.604
Gaussian	0.9907	0.9884	2.992	0.6116
Fourier	0.9915	0.9879	2.728	0.6243

Based on the parameters of Table 6, it was observed that the quadratic polynomial model had the lowest root mean square error value of 0.6053 and coefficient of determination R^2 value of 0.9919.

Table 7: Quadratic polynomial function for VOCs versus sampling distance

```
Linear model Poly2:
     f(x) = p1*x^2 + p2*x + p3
Coefficients (with 95% confidence bounds):
     p1 = -4.813e-005   (-7.233e-005, -2.393e-005)
     p2 =     -0.01021  (-0.02364, 0.003209)
     p3 =        28.29  (26.78, 29.81)
```

where;
f(x): VOCs concentration ($\mu g/m^3$)
x: Sampling distance (m)

Using the quadratic polynomial function of Table 7, the concentration of VOCs was projected to a sampling distance of 1500m at 95% confidence level.

Table 8: Calculated Model Parameters for $PM_{2.5}$ (Ughelli West)

Model	R-Square	Adj. R-Square	SSE	RMSE
Linear polynomial	0.9356	0.9284	110.1	3.498
Quadratic Polynomial	0.9456	0.9320	93.01	3.41
Exponential	0.8913	0.8792	185.9	4.545
Gaussian	0.9616	0.9519	65.75	2.867
Fourier	0.9717	0.9596	48.35	2.628

Based on the parameters of Table 8, it was observed that the Fourier function model had the lowest root mean square error value of 2.628 and coefficient of determination R^2 value of 0.9717.

3.4 Analysis of Seasonal Variability using Autocorrelation Function

Accurate analysis of data collected over time requires that seasonality analysis to check for the presence of seasonal variability be performed [6]. To estimate the degree of seasonality present in the field data collected from the three flow stations, autocorrelation plot was employed. A statistical software EViews version 9.0 was employed to generate the correlogram.

Table 9: Fourier function model for $PM_{2.5}$ versus sampling distance

```
General model Fourier1:
       f(x) =  a0 + a1*cos(x*w) + b1*sin(x*w)
Coefficients (with 95% confidence bounds):
       a0 =       32.83  (29.35, 36.3)
       a1 =       11.11  (1.023, 21.19)
       b1 =        11.6  (3.661, 19.54)
       w =    0.007821  (0.004846, 0.0108)
```

where;
f(x): particulate matter concentration ($\mu g/m^3$)
x: Sampling distance (m)

Using the Fourier function model of Table 9, the concentration of particulate matter was projected to a sampling distance of 1500m at 95% confidence level.

3.5 Trend Estimation using Non-Parametric Analysis

Since it was established that the field data collected from Ughelli West flow station are not normally distributed by exhibiting some characteristics that is occasioned by the presence of trend and seasonal variability, then non parametric analysis became the most suitable method to estimate the presence of trend in the data. Mann- Kendall trend test was therefore carried out by plotting the pollutants concentration against generated index.

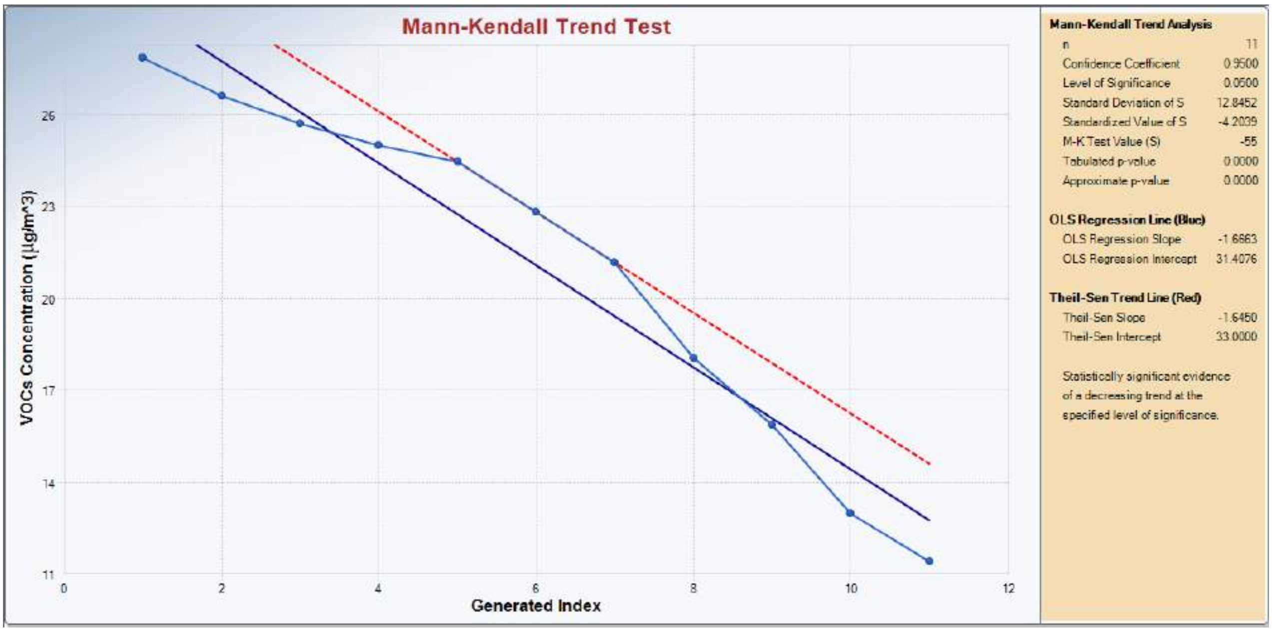


Figure 4: Mann-Kendall Trend Test of VOCs Data from Ughelli West Flow Station

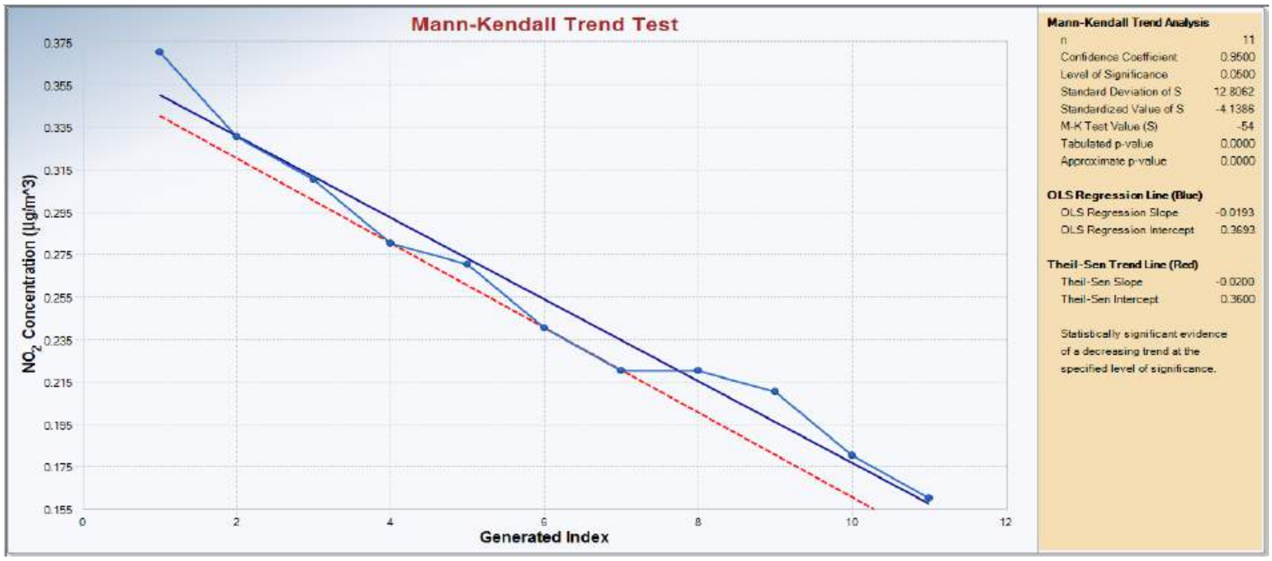


Figure 5: Mann-Kendall trend Test of NO_2 Data from Ughelli West Flow Station

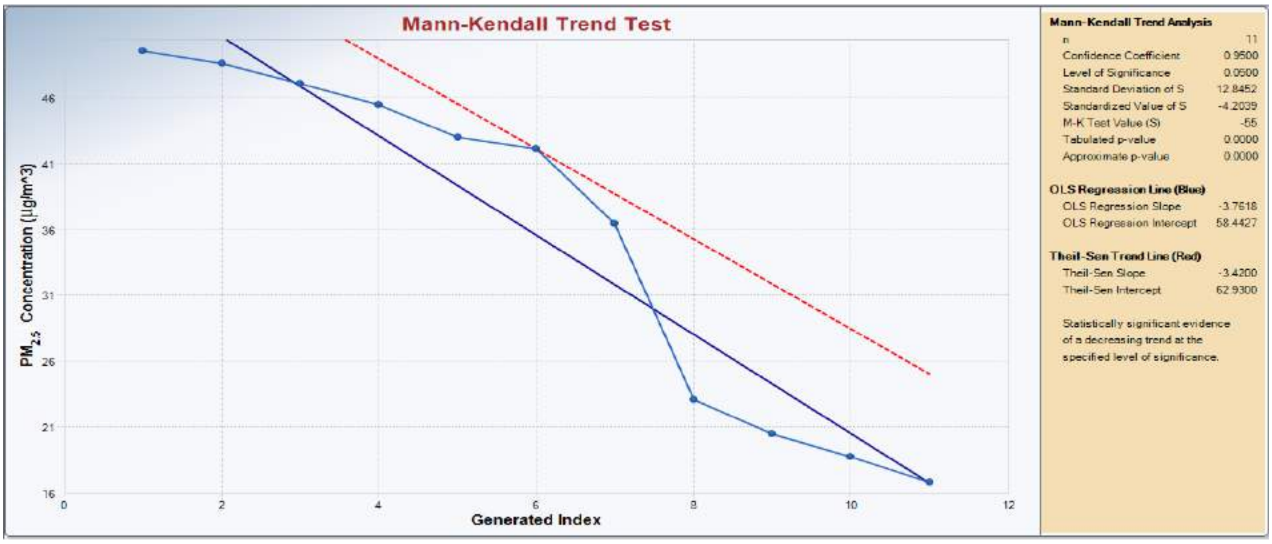


Figure 6: Mann-Kendall Trend Test of $PM_{2.5}$ Data from Ughelli West Flow Station

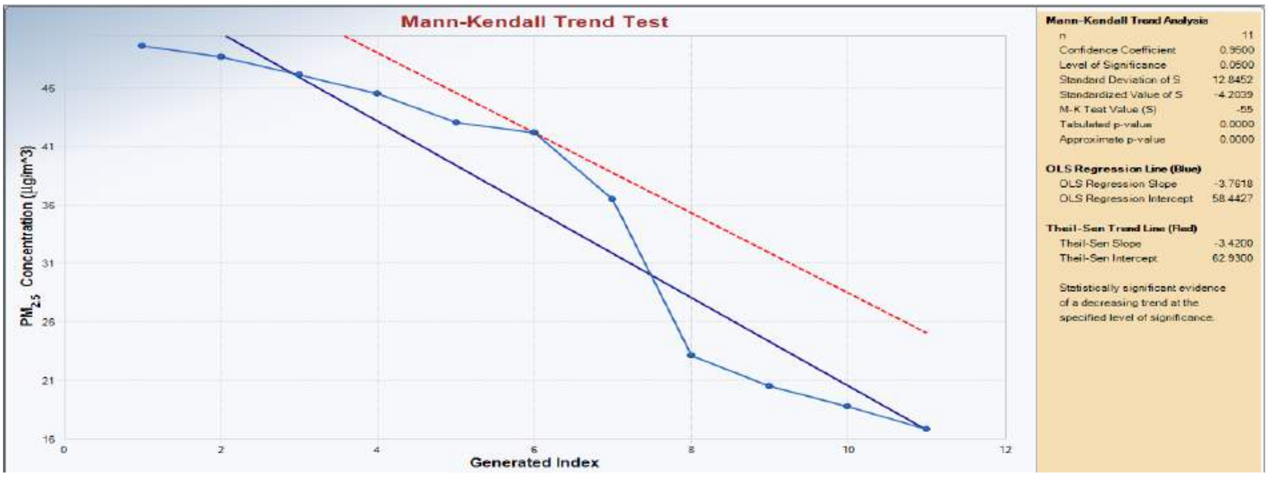


Figure 7: Mann-Kendall Trend Test of Ozone Data from Ughelli West Flow Station

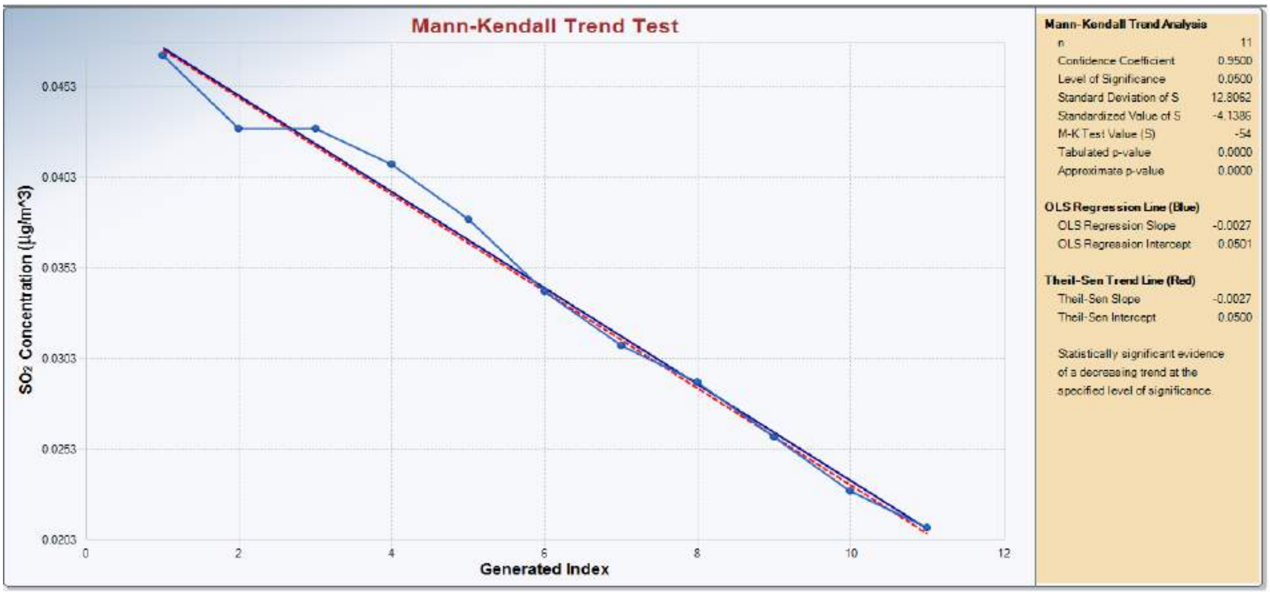


Figure 8: Mann-Kendall Trend Test of SO_2 Data from Ughelli West Flow Station

3.6 Geo-statistical Analysis of the Data

Detailed geo-statistical analysis of the data acquired from the three locations was carried out using ArcGIS 10.4.1. Geospatial modelling was employed to generate the prediction maps which show the distribution of the pollutants around the study locations. The prediction map is a pictorial presentation of the pollutant concentrations in space. The input parameters for the modelling were the rectangular coordinates of each sampling point and the concentration of the different pollutants measured, namely: volatile organic compounds (VOCs), methane (CH_4), nitrogen dioxide (NO_2), particulate matter ($PM_{2.5}$), ozone (O_3) and Sulphur dioxide (SO_2). Several geostatistical methods exist for geospatial modelling but the kriging method for point source pollution was employed.

3.7 Network Testing/Validation

Cross validation data representing 25% of the total input data was introduced to monitor the training process and to prevent the network from memorizing the input data while the remaining 15% was employed to test the performance of the trained network [4].

Table 10: Ughelli West Network Testing using Different Sets of Input Parameters

Sampling Distance (m)	Wind Speed (m/s)	Atm. Pressure (mmHg)	Ambient Temperature (deg C)	Relative Humidity (%)	$PM_{2.5}$
600	3.44	1255	31.4	89.6	36.9428
650	3.76	1023	32.3	87.3	15.55204
700	4.89	1015	33.8	99.2	21.76609
750	5.6	1116	28.2	94.2	36.9428
800	4.33	1222	29.7	83.8	36.9428
850	6.43	1033	34.8	90.1	15.24561
900	5.89	1035	36.3	75.4	18.66871
950	2.33	1010	31.8	78.8	15.9446
1000	4.44	1056	30.9	88.2	19.8917

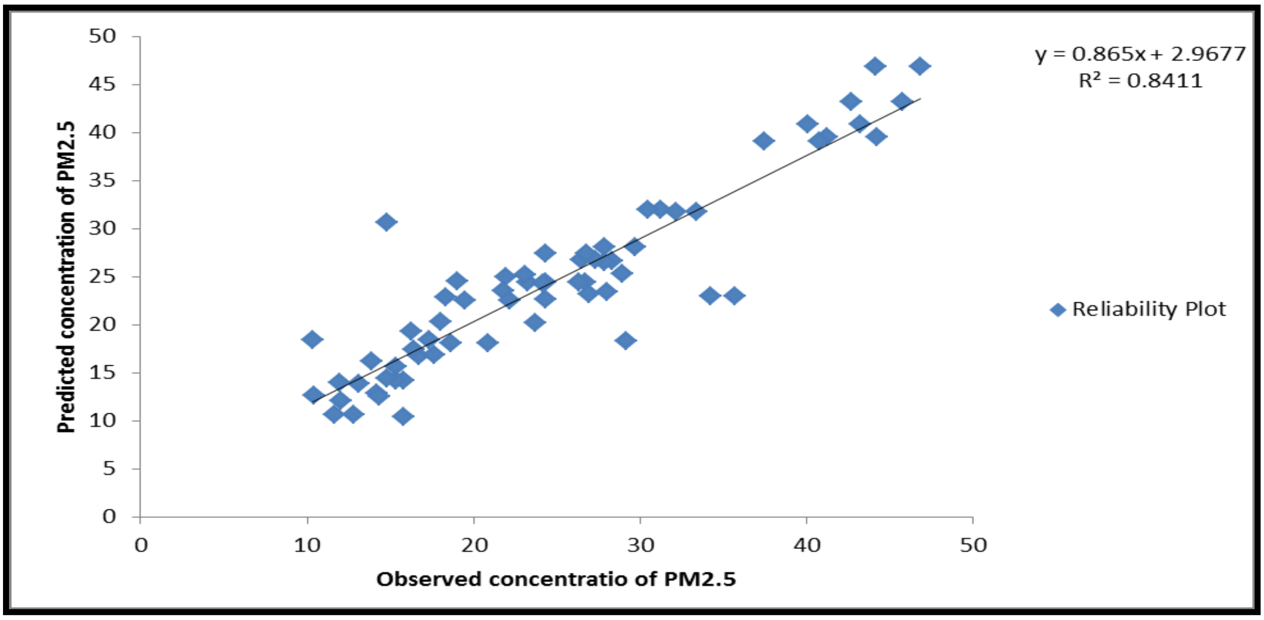


Figure 9: Regression plot of observed versus predicted concentration of PM$_{2.5}$

Based on the calculated coefficient of determination (R^2) as observed in Figure 9, the trained network was then employed to predict the concentration of PM$_{2.5}$ around Ughelli West flow station as presented in Table 8.

3.8 Computation of Pollution Standard Index (PSI):

The PSI values for PM, O_3 and SO_2 were computed and checked against the units in the standard Table. Computation of Pollution standard index (PSI) was done to determine the status of the environment within and around the study location. PSI gives reliable information on whether the environment is good, moderate, healthy, unhealthy, very unhealthy or hazardous [12]. The computed PSI values of the gaseous pollutants were compared and the pollutant with the highest PSI value was used to describe the air quality and environmental status as shown in Figure 4-8.

4. Conclusion

The modelling and investigation of Ughelli West gas plant gaseous pollutants parameters show that the quadratic polynomial model had the lowest root mean square error value of 0.6053 and coefficient of determination R^2 value of 0.9919. The results obtained from the validation shows good predictability and adequacy of the models developed in petroleum drilling processes. The data obtained from this research and developed models can find practical applications to reducing the gas flaring in oil and gas companies across the world.

References

[1] Khitoliya, R.K. (2007). Environmental Pollution. Management & Control for Sustainable Development. S. Chand & Company Ltd. Ram Nagar, New Delhi. Revised Edition.

[2] Agaja, S.A. (2014). Measurement of greenhouse gases in the environment of a petroleum refinery and petrochemical company. *Journal of Environmental Science, Toxocology and Food Tecnology*, Vol. 8, Issue 2 ver.11, pp. 53-56.

[3] Obahiagbon, K. (2002). Air Pollution. Hand book on Environmental Pollution. Edited by A.O. Ibhadode, Faculty of Engineering, University of Benin, Nigeria. 206 pp.2-10.

[4] Ajugwo, A.O. (2013). Negative effects of gas flaring. The Nigerian Experience. *Journal of Environment, pollution and Human health.* Vol. 1, pp 6-8.

[5] Hamed, K.H. and Rav, A.R. (2008). A Modified Mann-Kendall Trend Test for Autocorrelated Data, *Journal of Hydrology*, Vol. 204 (1-4), pp 182-196.
[6] Najjar, Y.S.H. (2011). Gaseous Pollutants Formation and their Harmful Effects on Health and Environment. Ashdin Publishing, *Innovative Energy Policies*. Vol. 1, Article IDE101203, doi:10.4303/iep/E101203.
[7] Rai, R., Rajput, M., Agrawal, M, and Angrawal, S. B (2011). Gaseous Air Pollutant: a Review on Current and Future Trends of Emissions and Impact on Agriculture. *Journals of Scientific Research. Banaras Hindu University, Varanasi, India*. Vol. 55 pp 77-102
[8] Anjaneyuhu, Y., Venkata, B.R.D. and Sudha, Y. (2011). Air Pollution, Modeling and GIS based Decision Support Systems for Air Quality Risk Assessment. ISBN: 978-953-307-511.2, In Tech, http://www.intechopen.com/books/advanced - Air -pollution/ air - pollution modelling - and – GIS
[9] Douabul, A.A.Z., AL-Maarofin, S. S., AL-Saad, H.T and AL-Hassan, S. (2013). Gaseous Pollutants in Basera City, Iraq. Libertas Academics Ltd Iraq. http://www.la-press.com
[10] Grigoras, G., Cuculeanu, V., Ene, G., Moeioaca, G. and Deneanu, A. (2012). Air Pollution Dispersion Modelling in a Polluted Industrial Area of Complex Terrain from Romania. *Romania Reports in Physics*, Vol. 64, No. 1, pp 173-186. WESTAGEM Ltd, Bucharest Romania
[11] Gorai, A.K., Tuluri, F. and Tchounwou, P.B. (2015). Development of Partial Least Squares Path Model (PLS-PM) for understanding the role of precursors on ground level ozone concentration in Gulfport, Mississipi, USA. *Atmospheric Pollution Research*, 6 (2015) 389-397, doi:10. 5094/APR.2015.043.(GLO)
[12] Yannawar, V.,Bhosle, A. and Yannawar, S. (2014). Prediction of air pollution concentration using a fixed box model. *International Journal of Science.* Vol. 6, No5, pp 89-92. ISSN:1553 -9865, http://www.sciencepub.net/researcher
[13] Narayanan, P. (2009). Environmental Pollution: Principles, Analysis and control CBS Publishers and Distributors PVT. Ltd. New Delhi. Reprint.
[14] Mgbamena,O.O., and Uzoma, A. (2015). Evaluation of Some Oil Companies in the Niger Delta Region of Nigeria: An Environmental Approach. *International Journal of Environment and Pollution Research*. Vol. 3, No.2, pp13-31.
[15] Grigoras, G., Cuculeanu, V., Ene, G., Moeioaca, G. and Deneanu, A. (2012). Air Pollution Dispersion Modelling in a Polluted Industrial Area of Complex Terrain from Romania. *Romania Reports in Physics*, Vol. 64, No. 1, pp 173-186. WESTAGEM Ltd, Bucharest Romania