The 5th PSU-UNS International Conference on Engineering and Technology (ICET-2011), Phuket, May 2-3, 2011
Prince of Songkla University, Faculty of Engineering
Hat Yai, Songkhla, Thailand 90112

# Application of Legendre Neural Network for Air Quality Prediction

**Santosh Kumar Nanda[1*], Debi Prasad Tripathy[2], S.S. Mahapatra[3]**
[1]Department of Computer Science and Engineering, Eastern Academy of Science and Technology, Bhubaneswar, Orissa, India, 754001
[2]Department of Mining Engineering, National Institute of Technology, Rourkela, Orissa, India, 769008
[3]Department of Mechanical Engineering, National Institute of Technology, Rourkela, Orissa, India, 769008
* email: santoshnanda@live.in

**Abstract:** *In this present time, air pollution is treated as an environmental predicament for developed and developing countries. Numerous models which could be useful to estimate pollutant concentrations as a function of the emission distribution and the attendant meteorological conditions have been investigated. So far a lot of them have been based up on physical and chemical principles. In this paper an intelligence system approach, in air pollution modelling has been proposed. The main target of this approach is the prediction, on the basis of meteorological prevision. The system was developed by using air pollution concentration as a function of the presence of the pollutant in it and the meteorological parameters. From this present investigation, it is found that legendre neural based intelligence system has a good prediction capability.*

**Key Words:** ***Air Quality Models, Artificial Neural Network, Functional Link Artificial Network (FLAN), Legendre Neural Network***

## 1. INTRODUCTION

In this present time, the environmental problems were treated as a headache for all developing or developed countries. Increased mechanization, transportation, populations etc. causes air pollution which was a major and effective environmental problem. It was generally a mixture of chemicals, particulate matter, or biological materials that cause harm or discomfort to humans or other living organisms, or damages the natural environment into the atmosphere. The impact of air pollution was depending upon its parameters present in it. Ozone (O3), Carbon Monoxide (CO), sulfur dioxide (SO2), nitrogen monoxide (NO), and nitrogen dioxide (NO2) are all the primary pollutant. These can take part in further chemical reactions once they are in the atmosphere, forming smog and acid rain. To prevent air pollution, it is essential need to predict the presence of influence parameters with meteorological condition [1,2]. Earlier the predictions of air pollutant parameters were conducted by using statistical or regression models. The statistical models or regression equations are subject to the assumptions and cautions inherent in the analyses. The major postulation with statistical models is that the data represent an underlying "reality" which can be uttered by an algebraic equation. Since air pollution parameters prediction involves a number of input and output variables, the statistical analysis ensure a group of intricate, difficult-to-read, and often difficult-to-employ mathematical expressions. The prediction correctness of statistical models is usually represented by the associated

Earlier the predictions of air pollutant parameters were conducted by using statistical or regression models. The statistical models or regression equations are subject to the assumptions and cautions inherent in the analyses. The major postulation with statistical models is that the data represent an underlying "reality" which can be uttered by an algebraic equation. Since air pollution parameters prediction involves a number of input and output variables, the statistical analysis ensure a group of intricate, difficult-to-read, and often difficult-to-employ mathematical expressions. The prediction correctness of statistical models is usually represented by the associated regression coefficient $R^2$. Once a model has been formed, the user cannot improve the accuracy of fit. In the case where inputs and outputs are intrinsically variable, a system that is intended only to be predictive is more appropriate. A system which can "learn" from consecutive experimental trials is ideal because it will become more consistent through time and will be able to adapt to unpredicted changes in the data. In comparison to statistical models, an intelligence system can approximate the input and output relationship with maximum number of input and output data pattern. Furthermore, in intelligence system, user can extend its degree of accuracy at any time. Therefore intelligent systems are more economic than classic methods [3–6].

In this paper, an attempt has been made to develop a Legendre artificial neural network model

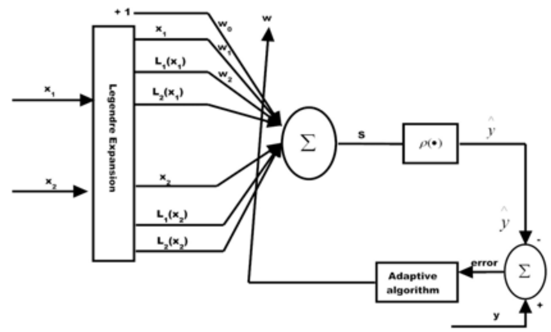
(LeNN) for air pollution parameter prediction with its depended variables. The data assembled through surveys, measurement or knowledge to predict air pollutants is often imprecise or speculative. Since neural network based systems are good predicted tools for imprecise and uncertainty information, therefore this proposed approach would be the most appropriate technique for modeling the prediction of air pollution parameters.

## 2. INTRODUCTION TO FUNCTIONAL LINK BASED ARTIFICIAL NEURAL NETWORK

Neural Network (NN) represents an important paradigm for classifying patterns or approximating complex non-linear process dynamics. These properties clearly indicate that NN exhibit some intelligent behavior, and are good candidate models for non-linear processes, for which no perfect mathematical model is available. Neural networks have been a powerful tool for their applications for more than last two decades [7–8]. Multilayer Perceptron (MLP), Radial Basis Function (RBF), Support vector machine (SVM)etc. are the types of Neural Network Model, where these models have better prediction competence with high computational cost. Generally, these models have high computational cost due to the availability of hidden layer. To minimize the computational cost, structures like, legendre neural network (LeNN) [9,10] were proposed. In this paper, legendre neural network have been applied to predict air pollution parameter. In general the functional link based neural network models were single-layer ANN structure possessing higher rate of convergence and lesser computational load than those of a MLP structure. The mathematical expression and computational calculation is evaluated as per MLP. Structure of the Legendre Neural Network [9,10] (LeNN) (shown in Fig. 1) is similar to FLANN. In contrast to FLANN, in which trigonometric functions are used in the functional expansion, LeNN uses Legendre orthogonal functions. LeNN offers faster training compared to FLANN. The performance of this model may vary from problem to problem. The Legendre polynomials are denoted by $L_n(X)$, where n is the order and $-1 < x < 1$is the argument of the polynomial. The zero and the first order Legendre polynomials are, respectively, given by $L_0(x) = 1$and $L_1(x) = x$. The higher order polynomials are given by $L_2(x) = 1/2(3x^2 - 1)$, $L_3(x) = 1/2(5x^3 - 3x)$ etc. The structure of LeNN was similar to FLANN except the functional block. The functional expansion block make use of a functional model comprising of a subset of orthogonal sin and cos basis functions and the original pattern along with its outer products. For example, considering a two-dimensional input pattern $X = [x^1 x^2]^T$. The enhanced pattern is obtained by using the trigonometric functions as $X^* = [x_1 \cos(\pi x_1) \sin(\pi x1) ... x_2 \cos(\pi x_2) \sin(\pi x_2) \ . \ . \ . \ x_1 x_2]^T$ which is then used by the network for the equalization purpose. The BP algorithm, which is used to train the network, becomes very simple because of absence of any hidden layer. Polynomials are generated by using the following mathematical expression:

$$L_{n+1}(x) = \frac{1}{n+1}\left[(2n+1)x\, L_n(x) - n\, L_{n-1}(x)\right] \quad (1)$$

As similar to FLANN, the two dimensional input pattern $X = [x_1 x_2]^T$ is enhanced to a seven dimensional pattern by Legendre functional expansion $X^e = [1, L_1(x_1), L_2(x_1), L_3(x_1), L_1(x_2), L_2(x_2), L_3(x_2)]$.For legendre neural network, the training is carried out in the same manner as




FLANN.

Fig1. System architecture of legendre neural network

## 3. DEVELOPMENT OF LEGENDRE NEURAL NETWORK MODEL FOR PREDICTING AIR POLLUTION PARAMETERS

As the prediction of the air pollution parameters e.g. Carbon dioxide (*CO*), Nitrogen oxide (*NO*2), Sulphur dioxide (*SO*2), Ozone (*O*3) is a very complex task as this can be treated as MIMO System (Multi input and multi output). The data was collected from CPCB as data is limited, here a new procedure was followed to design the model. In first process statistical relationship or regression models were developed and after that LeNN model was applied. The MIMO system is presented in Fig 2 and where, temperature and humidity as the inputs and $CO$, $NO_2$, $SO_2$, $O_3$ are the output of the system.

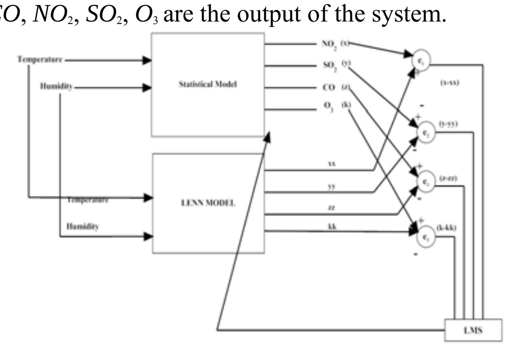
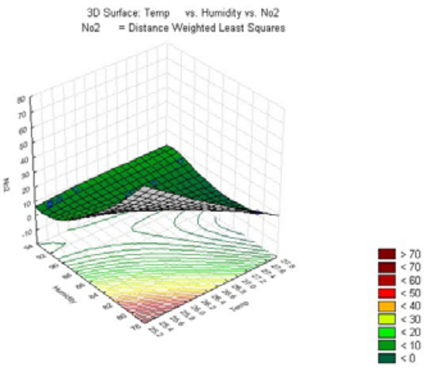


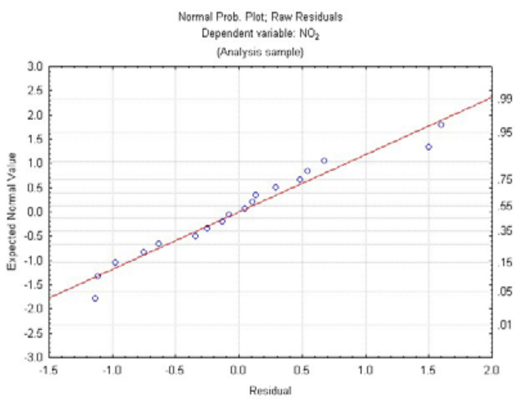

Fig.2. Block diagram of the proposed model

### 3.1 Development of Non-Linear Regression Model

The first procedure was to develop a non-linear statistical relationship between the input parameters (temperature and Humidity) and the output (Nitrogen Oxide, Sulphur Oxide, Ozone and Carbon Monoxide).Table 1 represented the data set. Preliminary this data set was used to developed non-linear regression model with a good $R^2$ value and after that this data set was used as a validation data for testing the prediction capability of the LeNN model. Figure 3 represents the regression analysis between Temperature, Humidity and Nitrogen Oxide, where expressions 2 represent the non-linear regression equation. Similar results were obtained for Ozone, Sulphur dioxide and Carbon Monoxide and their regression models were represented in expressions

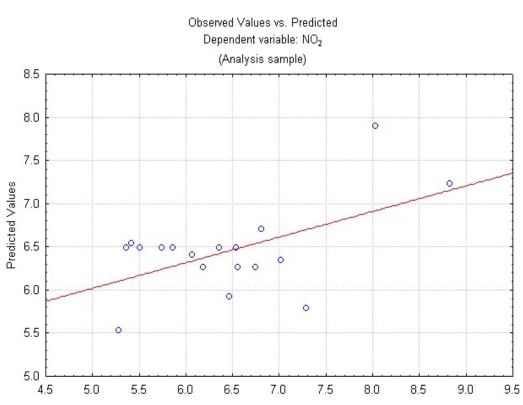



3, 4 and 5 respectively.

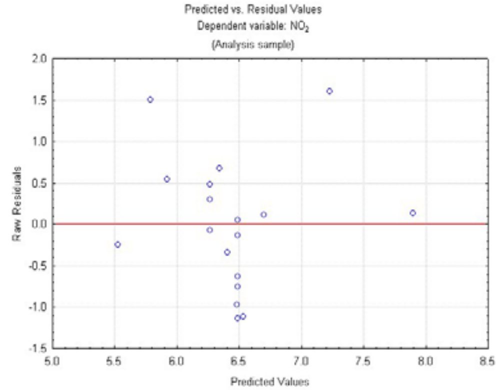



(a)

(b)




(c)




(d)

Fig.3 Regression analysis for Nitrogen Oxide, (a) Surface Plot between the three parameters (b) Normal Plot (c) Comparison between observed and predicted values (d) Comparison between predicted and residual values

$NO_2 = 5610.61705 - 171.83913 \times Temp + 0.678598061 \times Temp^2 - 73.211098 \times Humidity + 0.188224955 \times Humidity^2$ (1.2) $1.48314019 \times Temp \times Humidity$; ($Residual$ :0.856) (2)

$O_3 = -3408.9528 + 100.464137 \times Temp - 0.14391973 \times Temp^2 + 45.1729881 \times Humidity - 0.10328934 \times Humidity^2$ (1.3) $-0.99731255 \times Temp \times Humidity$; ($Residual$ :0.855) (3)

$SO_2 = 819.356296 - 25.545045 \times Temp + 0.153040565 \times Temp^2 - 10.690252 \times Humidity + 0.031119699 \times Humidity^2$ (1.4) $+0.193505042 \times Temp \times Humidity$; ($Residual$ :0.858) (4)

$CO = 387.657347 - 13.425931 \times Temp + 0.104830655 \times Temp^2 - 4.6491564 \times Humidity + 0.013182283 \times Humidity^2$ (1.5) $+0.086685822 \times Temp \times Humidity$; ($Residual$ :0.886) (5)

### 3.1 Development Legendre based air pollution parameters prediction models

According to Fig.1, the proposed system was designed. The legendre neural network based air pollutant parameters prediction model consists of two input constituting temperature ($^1xk$) and humidity ($^2xk$). The inputs patterns are $^1x1(k)$, $^1x2(k)$, $x3(k)$ ... $^1xn(k)$ $*$ R, $^2x1(k)$, $^2x2(k)$, $^2x3(k)$ ... $^2xn(k)$) and the desired output patterns are:d1(k),d2(k),d3(k)...d$n$(k) $\in$ R. In this proposed system, $NO_2$, $CO$, $SO_2$ and $O_3$ are the deisred output patterns. During training period the desired network output was calculated with the regression models (expression 2 to 5). The algorithm of legndre neural network based model.

Step1: Initialize the inputs distance = $x1,i$ , (i=1...n), sound power level = $x2,j$ (j=1...m), where n and m are the number of input pattern and an error tolerance parameter $\varepsilon > 0$. The dimension of m and n should be same.

Step2: Randomly select the initial values of the weight vectors $w_i$ , for i=1, 2,...l, where 'i' is the number of functional elements.

Step3: Initialization All the weights $w_i$ were initialized to random number and given as $w_i$ (0). $w_i \rightarrow w_i(0)$

Step 4: Produce functional blocks

For FLANN the functional block is made as follows:
$Xi=[1,x1,sin(\pi x1),cos(\pi x1),x2, sin(\pi x2), cos(\pi x2)..]$ For PPN the functional block is made as follows:
$Xi=[1,x1,x_1^2,x_1,x2,x_2^2,x_2,....]$
For LeNN the functional block is made as follows:
$Xi=[1,x1,L1(x1),x2,L1(x2),....]$
where $L1(x) = 1/2(3x^2-1)$ $L2(x)=1/2(5x^3-3x)$ etc.

Step 5: Calculation of the system outputs

$$O_i = \tanh\left(\sum_{i=1}^{N} w_i \times X_i\right)$$

Step 6: Calculation of the output error
The error was calculated as $e_i = d_i - O_i$. It may be seen that the network produces a scalar output.

Step 7: Updating the weight vectors
The weight matrixes are updated next using the following relationship $w_i(k+1) = w_i(k)+ \alpha e_i(k)X_i(k)$ where k is the time index and α is the momentum parameter.

Step 8: If error $\leq \varepsilon$ (0.01) then go to Step 8, otherwise go to Step 3.

Step 9: After the learning is complete, the weights were fixed and the network can be used for testing.

## 4. SIMULATION RESULT AND DISCUSSION

The proposed system models for noise prediction were validated using simulation studies. The studies were carried out by using MATLAB simulation environment. For validation of the model the data was collected from Central Pollution and Control Board, NewDelhi.Table1 represents the data set. In this propose system, $NO_2$, $CO$, $SO_2$ and $O3$ are all the dependent parameters, where temperature and humidity are the independent parameters. In this research, due to availability of less data, here a procedure was follows. At first, using the data set in table1, statistical relationship or regression models with good $R^2$ were developed. Expressions 1.2 to 1.5 are represents the regression models. After generating regression models, the system can design by suing these models. Training and testing data sets were generated using the regression models. The system network was trained with this dataset. In this proposed system, total number of 3200 data set were generated, where 3000 data set were used for training and rest of others were used as testing data set. Supervised training method was applied here to train the intelligent system. Fig.4 represents the performance of the statistical model with LeNN model. It was seen that the performance of the proposed system was good prediction capacity as compared to regression models.

Table 1. *Experimental data Source [11]*

| Date | Hour | NO2(PPB) | SO2(PPB) | O3(PPB) | CO(PPM) | TMPo*C* | HUM(%) |
|---|---|---|---|---|---|---|---|
| 7/27/2006 | 0100 hrs | 6.750 | 1.410 | 2.360 | 0.821 | 25.400 | 93.400 |
| 7/27/2006 | 0200 hrs | 6.190 | 1.350 | 2.090 | 0.713 | 25.400 | 93.400 |
| 7/27/2006 | 0300 hrs | 6.560 | 1.320 | 1.830 | 0.629 | 25.400 | 93.400 |
| 7/27/2006 | 0400 hrs | 6.540 | 1.310 | 2.410 | 0.619 | 25.500 | 93.400 |
| 7/27/2006 | 0500 hrs | 6.360 | 1.350 | 2.010 | 0.657 | 25.500 | 93.400 |
| 7/27/2006 | 0600 hrs | 5.360 | 1.270 | 2.110 | 0.581 | 25.500 | 93.400 |
| 7/27/2006 | 0700 hrs | 5.860 | 1.280 | 2.500 | 0.643 | 25.500 | 93.400 |
| 7/27/2006 | 0800 hrs | 6.810 | 1.300 | 1.570 | 0.737 | 25.600 | 93.400 |
| 7/27/2006 | 0900 hrs | 8.830 | 1.700 | 1.610 | 0.890 | 25.900 | 93.400 |
| 7/27/2006 | 1000 hrs | 8.030 | 1.840 | 2.000 | 0.770 | 26.600 | 93.400 |
| 7/27/2006 | 1100 hrs | 7.020 | 1.930 | 4.750 | 0.840 | 27.700 | 91.200 |
| 7/27/2006 | 1200 hrs | 5.280 | 1.870 | 6.520 | 0.690 | 27.700 | 80.200 |
| 7/27/2006 | 1300 hrs | 7.290 | 1.340 | 1.720 | 0.636 | 26.700 | 87.100 |
| 7/27/2006 | 1400 hrs | 6.470 | 1.410 | 2.530 | 0.560 | 26.600 | 88.000 |
| 7/27/2006 | 1500 hrs | 6.070 | 1.360 | 3.160 | 0.574 | 26.500 | 90.000 |
| 7/27/2006 | 1600 hrs | 5.740 | 1.310 | 4.020 | 0.527 | 26.600 | 90.100 |
| 7/27/2006 | 1700 hrs | 5.420 | 1.360 | 3.250 | 0.532 | 26.800 | 90.100 |
| 7/27/2006 | 1800 hrs | 5.510 | 1.350 | 3.560 | 0.618 | 27.000 | 90.000 |

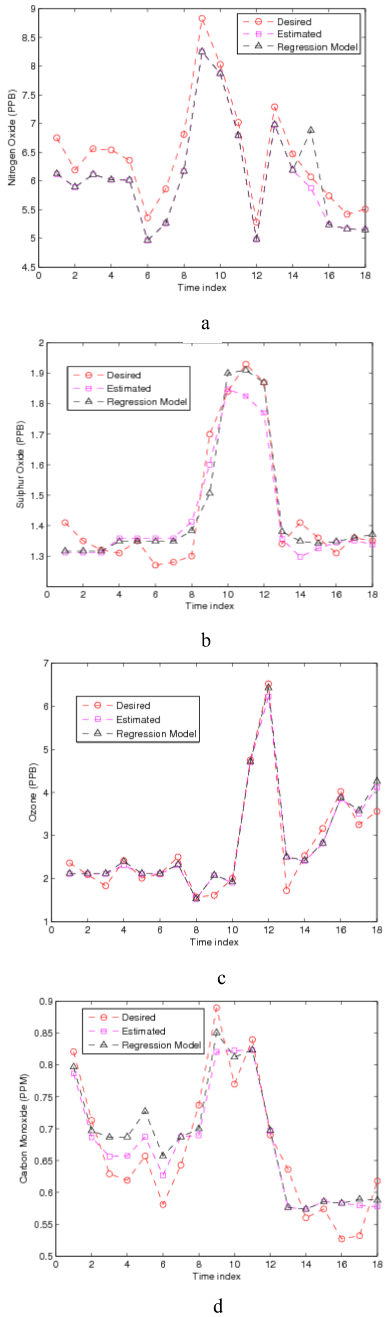



Fig. 4. Performance study of the LeNN system (a) $NO_2$ (b) $SO_2$ (c) Ozone (d) CO

## 5. CONCLUSION

This paper introduced the idea of designing the Legendre neural network based air quality parameter prediction for environmental engineering application. In this research contribution, one new idea has been proposed as designed an intelligent system with regression models. From simulation result, this proposed model has good performance as compared to regression models. Hence it may consider as a useful tool for the environmental engineers to predict pollution parameters with intelligent system and less computational cost.

## ACKNOWLEDGEMENT

Authors are thankful to ASME and Central Pollution Control Board, Delhi, India to provide the data in their website to free use for research work.

## 5. REFERENCES

[1] N.P. Cheremisinoff, "Handbook of Air Pollution Prevention and Control", Butterworth-Heinemann, USA, 2002.

[2] D.A Vallero, "Fundamentals of Air Pollution", 4th edn. Academic Press, USA, 2008.

[3] X. Dai, H. Shi, Y. Li, Z. Ouyang1 and Z. Huo, "Artificial neural network models for estimating regional referenceevapotranspiration based on climate factors. *Hydrological Processes*, Wiley InterScience, 2009, Vol. 23, pp. 442-450

[4] H.K. Elminir and H.A. Galil, "Estimation of air pollutant concentrations from meteorological parameters using artificial neural network", *Journal of Electrical Engineering*, 2006, vol. 57, No.2, pp. 105–110.

[5] C.R. Chen, H.S. Ramaswamy and I. Alli, "Prediction of quality changes during osmo-convective drying of blueberries using neural network models for process optimization", *DryingTechnology*, Talyor and Francis, 2001, vol. 19, No.3&4, pp. 507–523.

[6] Saral and F.Erturk, "Prediction of ground level $so_2$ concentration using artificial neural net-works", *Water, Air, and Soil Pollution: Focus*, Kluwer Academic Publishers, 2003, vol. 3, pp. 297–306

[7] S. Haykin, "Neural Networks: A omprehensive Foundation", Prentice-Hall, Reading, MA, 1994.

[8] J.S. Jang, C.T. Sun and E. Mizutan, "Neuro-Fuzzy and Soft Computing", Prentice Hall of India Private Limited, New Delhi, 2005.

[9] J. Patra,W.Chin, P. Meher, and G. Chakraborty, "Legendre-flann-based nonlinear channel equal-ization in wireless communication systems". *Proceedings of IEEE International Conference on Systems, Man, Cybernetics (SMC2008)*, 2008, pp. 1826–1831.

[10] J. Patra, P. K. Meher, and G. Chakraborty , "Nonlinear channel equalization for wireless communication systems using legendre neural networks". *Signal Processing*, 2009, vol. 89, pp.2251–2262

[11] CPCB:Air quality data acquisition system, http://www.asemindia.com/main Project.aspx?SubSectionID=10&SectionID=2&RowID=11&Projid=18