Proceedings of the International Conference on Robotics, Vision, Information and Signal Processing
January 22 – 24, 2003, Penang, Malaysia

Carbon Monoxide Concentration Forecasting Using Hybrid Radial Basis Function Network

Mazlina Mamat[1] **Mohd. Yusof Mashor[1]** **Abdul Rahman Saad[2]**
Ahmad Farhan Sadullah[3]

[1]*Centre For Electronic Intelligent System (CELIS), School of Electrical & Electronic Engineering,*
University Sains Malaysia, Engineering Campus Nibong Tebal, Seberang Perai Selatan, Pulau Pinang, Malaysia
Tel: +604-5937788, Fax: 604-5941023, E-mail: yusof@eng.usm.my

[2]*School of Electrical & Electronic Engineering,*
Kolej Universti Kejuruteraan Utara Malaysia (KUKUM), Kampus Sementara, Taman JKKK,
02600 Arau, Perlis, Malaysia, Tel: +604-9778000, Fax: 604-9778011

[3]*School of Civil Engineering*
University Sains Malaysia, Engineering Campus, Nibong Tebal, Seberang Perai Selatan, Pulau Pinang, Malaysia
Tel: +604-5937788x6209, Fax: 604-5941009, E-mail: cefrhn@pd.jaring.my

Abstract

Neural network has been renown for its applications in many fields of research especially related to pattern recognition. In this paper, Hybrid Radial Basis Function (HRBF) Neural Network will be exploited to carry out the Carbon Monoxide forecasting. This research utilized the past Carbon Monoxide data to forecast Carbon Monoxide concentrations for several hours in advance. Instead of performing the off-line forecasting technique, this research tries to forecast the Carbon Monoxide concentrations by on-line technique. To establish this requirement, the HRBF network is trained by using Adaptive Fuzzy C-Means Clustering Algorithm and Exponential Weighted Recursive Least Square Algorithm. For evaluation purpose, we use Carbon Monoxide concentration time series from three air quality monitoring stations situated at Sekolah Menengah Victoria, Wilayah Persekutuan Kuala Lumpur; Sekolah Kebangsaan Raja Muda, Selangor; and Institut Latihan Perindustrian, Penang. The performance of each model is indicated in the terms of coefficient of determination (R^2) between observed and forecast values. The results showed that the HRBF Network has the credibility to be used as a Carbon Monoxide forecaster and its performance dependent on the complexity of the data.

Keywords

Carbon Monoxide Forecasting, Neural Networks, On-line Forecasting, HRBF Network.

Introduction

Carbon Monoxide (CO) are the leading cause of poisonous death in the United States [1] & [2]. Based on the Journal of the American Medical Association (JAMA), around 1500 deaths every year due to CO poisoning [1]. Recently a research has reported that continuous exposure to a thin amount of Carbon Monoxide can lead to the debilitating residual effects that may continue for days, months and even years [2].

Nearly all the Carbon Monoxide poisonous cases were contributed by the vehicle's exhaust. This statement are based on the Centers for Disease Control and Prevention's record (1996) which stated that between 1979 to 1988, 57% from 11547 deaths of Carbon Monoxide poisoning were caused by vehicle's exhaust production [3].

In Malaysia, Carbon Monoxide are emitted into the urban atmosphere mainly from vehicle exhausts [4]. CO is a colorless, odorless but poisonous gas, a product of incomplete burning of hydrocarbon based fuels. CO consists of a Carbon atom and an Oxygen atom linked together.

A lot of researches have been carried out to determine the factors which control CO concentrations in order to enable the development of tools for forecasting the resulting pollutant concentrations. One approach is to predict future concentrations by using statistical model which attempt to determine the underlying relationship between input data and targets. An example of statistical approach is regression analysis. It has been applied to CO modeling and prediction in a number of studies [5],[6].

Another method in statistical modeling is Artificial Neural Network (ANN). It is well known that ANN can model non-linear systems and it has been used to model Sulphur Dioxide concentrations in Slovenia [7] and to model $PM_{2.5}$ concentrations in Santiago, Chile [8]. In this paper, ANN was used to model and predict hourly CO concentrations from readily observable CO data.

Approach and Methods

Data

The data for this investigation were obtained from Alam Sekitar Malaysia Sdn. Bhd. (ASMA) Malaysia. These data contain the average hourly CO measurement for variables such as temperature, wind speed and wind direction. The first data set is a Traffic Data from 1st January 2001 to 5th May 2001 (3000 data) were selected from a site operating at Sekolah Menengah Victoria, Wilayah Persekutuan Kuala Lumpur (Traffic Data). These data are classified by ASMA as CO concentration data in the traffic area. For these data, the average concentration is 2.73 ppm, the maximum is 13.85 ppm and the standard deviation is 1.64. The second set of data was obtained from a site operating at Sekolah Kebangsaan Raja Muda, Selangor and was classified by ASMA as CO concentration data for residential area (Residential Data) whereas the third data was obtained from Institut Latihan Perindustrian, Penang and are classified as Industrial Data. The average concentration, maximum concentration and standard deviation for Residential Data are 0.859 ppm, 5.82 ppm and 0.802. For Industrial Data, the average concentration, maximum concentration and standard deviation are 0.579 ppm, 3.42 ppm 0.433. respectively. The rationale to evaluate the selected ANN by using three different categories of CO data is to confirm the ANN's capability in performing the on-line CO concentration forecasting.

Neural Network Predictors

The standard neural network method of performing time series prediction is to induce the function f in a standard feed forward neural network architecture, using a set of N-tuples as inputs and a single output as the target value of the network. By using this method, the on-line forecasting on CO concentrations by using neural network was made. In the on-line technique, the network parameters are always updated whenever they receive new input. These make on-line technique yields better performance compared to off-line technique.

There exists a lot of neural network architectures. However majority of the neural network based forecasters use the feed forward Multilayer Perceptron neural network [9]. In this paper, Hybrid Radial Basis Function (HRBF) neural network with Adaptive Fuzzy C-Means Clustering Algorithm [10] and Exponential Weighted Recursive Least Square Algorithm [11] have been used to model the CO concentrations time series. The HRBF network and the algorithms were chosen because it can be implemented in the on-line technique. Details about HRBF neural network can be found in [12].

In this study, the number of steps ahead to be forecasted has been restricted to eight. There are three different architectures that can be used to determine multiple steps ahead forecasting. The first architecture is to use a single HRBF to forecast CO value at time t+1 and the forecast value is then used as a new input to forecast CO value t+2, t+3 and so on. The second architecture is to use single HRBF with as many outputs as values to forecast. The third architecture is to train a number of n HRBF, on for each value to forecast. In this paper, we have decided confer the performance of the first architecture (Sing Model) as CO predictors.

In general, the HRBF architecture for forecasting CO concentration is shown in Figure 1. Input to the HRBF consists of several numbers of previous CO concentration which was arranged in a specific form based on researcher selection. However the numbers of input lags to the HRBF is limited to 60 lags only. The same restriction is also applied to the number of hidden nodes used. Yet the maximum numbers of hidden nodes has been restricted to 80. These limitations can be explained by the importance of reducing the processing times and load in the system itself. The output node for Single Model is taken as one and it will represents the forecasted CO concentration.

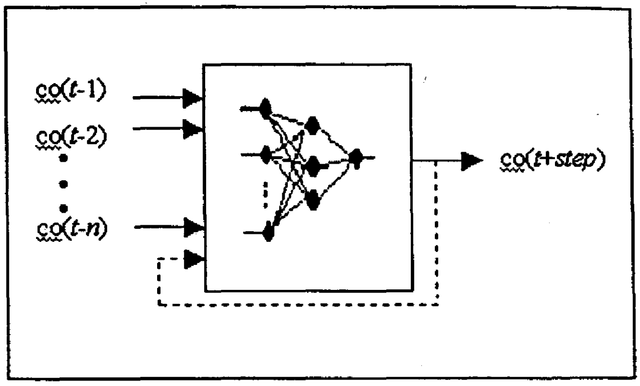

Figure 1: The HRBF architecture for forecasting CO concentration

Forecasting Performance

Ding, Canu and Denceux [13] stress that the selection of input lag and the structure of neural network have strong impact on the forecaster performance. In parallel to this, an analysis on input selection and network structure must be made. In this research the Single Model have gone through two analyses, first to determine the best input lag and second to determine the correct number of HRBF centers. Both analyses were made by replacing other HRBF parameters with the typical values [9]. The model were trained and tested on the Traffic Data, Residential Data and Industrial Data. The performance of the model is indicated in terms of coefficient of determination (R^2) given by

$$R^2 = 1 - \left(\frac{\sum_{t=n_a}^{n_t} (\hat{\varepsilon}(t))^2}{\sum_{t=n_a}^{n_t} (y(t) - \bar{y})^2} \right) \qquad (1)$$

where $\hat{\varepsilon}(t)$ and $y(t)$ are estimated error and observed value at time t, $\bar{y}$ is the average observed value, n_d and n_t are the first and the last test data respectively.

Traffic Data

The analysis to find out the best input lags for Traffic Data gave results as in Figure 2. From Figure 2, it can be concluded that the performance of Single Model depends on the selection of input lags. It can be noted that the R^2 value change when the input lags were changed. Apart from that the performance of the model in multiple steps ahead forecasting was faded when the numbers of input lags used is smaller than 24. However model seems to perform well especially in multiple steps ahead prediction while operating in the range of input lags from 27 to 44. Further increase in the number of lags greater than 46 do not give much benefit to the model, in fact it deterioration the model performance. Three input lags that provided the highest R^2 value for every step ahead prediction were selected and shown in Table 1.

Table 1: Selected Input Lags For Traffic Data

Number	Input Lags
1	$(t-1)(t-2)(t-3) \ldots (t-27)(t-28)$
2	$(t-1)(t-2)(t-3) \ldots (t-38)(t-39)$
3	$(t-1)(t-2)(t-3) \ldots (t-51)(t-52)$

The selected input lags was then used in the analysis to find out the correct number of HRBF centers to represent the Traffic Data. This analysis yields the results shown in the Figure 3 below.

From the result, it can be concluded that the HRBF network needs number of centers in the range of 4 to 20 in order to perform well. However, the use of large number of centers must be avoided because it can degrade the model performance in terms of time consuming. Table 2 shows the highest R^2 value achieved for each lag. From this result, the best model performance can be achieved by using input lags $(t-1)(t-2)(t-3) \ldots (t-27)(t-28)$ and the number of center 47.

Table 2: The Highest R^2 Value Achieved For Each Lag

Lag	[redacted]	2	3
Number of Center	[redacted]	63	10
Step 1	[redacted]	0.70	0.72
Step 2	[redacted]	0.47	0.47
Step 3	[redacted]	0.46	0.46
Step 4	[redacted]	0.44	0.45
Step 5	[redacted]	0.45	0.45
Step 6	[redacted]	0.46	0.44
Step 7	[redacted]	0.46	0.45
Step 8	[redacted]	0.46	0.45

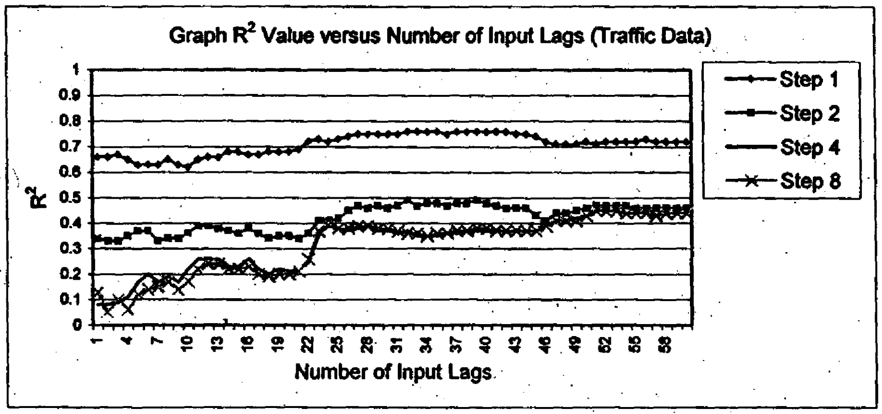

Figure 2: Graph R^2 Value versus Number of Input Lag (Traffic Data)

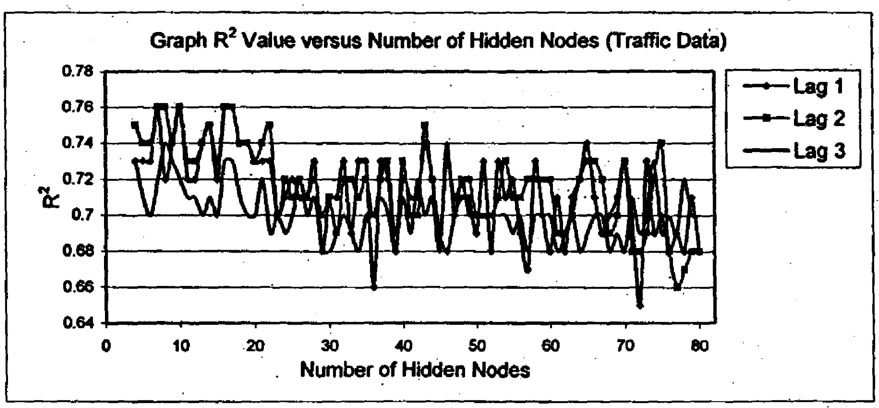

Figure 3: Graph R^2 Value versus Number of HRBF Center (Traffic Data)

Residential Data

The analysis to find out the best input lags for Residential Data yields results shown in Figure 4. By considering the graph plot in Figure 4, it can be noted that Single Model is capable to perform one step ahead prediction by just using four input lags. Nevertheless, the model fail to perform well in multiple steps ahead prediction and the R^2 value for four and eight steps ahead are around 0.3 to 0.4. However the performance of the model increases when the number of input lags increases. It can be observed that the model gives good performance at the number of input lags 25 and above.

From the graph in Figure 4, a total of three input lags have been identified to be used in the analysis to determine the correct number of HRBF center. Table 3 shows the selected input lags for Residential Data.

The result of analysis to find out the HRBF center for Single Model was given in Figure 5 below. Table 4 shows the highest R^2 value achieved for each lag. From the result, the best performance for Single Model can be achieved by using input lags $(t-1)(t-2)(t-3) \ldots (t-58)(t-59)$ and number of center of 36.

Table 3: Selected Input Lags for Residential Data

Number	Input Lags
1	$(t-1)(t-2)(t-3)$
2	$(t-1)(t-2)(t-3) \ldots (t-27)(t-28)$
3	$(t-1)(t-2)(t-3) \ldots (t-58)(t-59)$

Table 4: The Highest R^2 Value Achieved For Each Lag

Lag	1	2	[redacted]
Number of Center	18	17	[redacted]
Step 1	0.88	0.87	[redacted]
Step 2	0.66	0.68	[redacted]
Step 3	0.46	0.61	[redacted]
Step 4	0.28	0.59	[redacted]
Step 5	0.16	0.57	[redacted]
Step 6	0.23	0.57	[redacted]
Step 7	0.23	0.55	[redacted]
Step 8	0.17	0.55	[redacted]

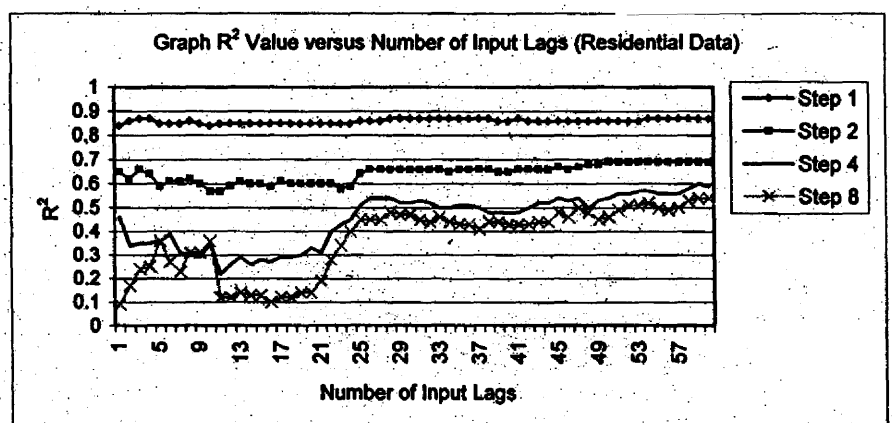

Figure 4: Graph R^2 Value versus Number of Input Lags (Residential Data)

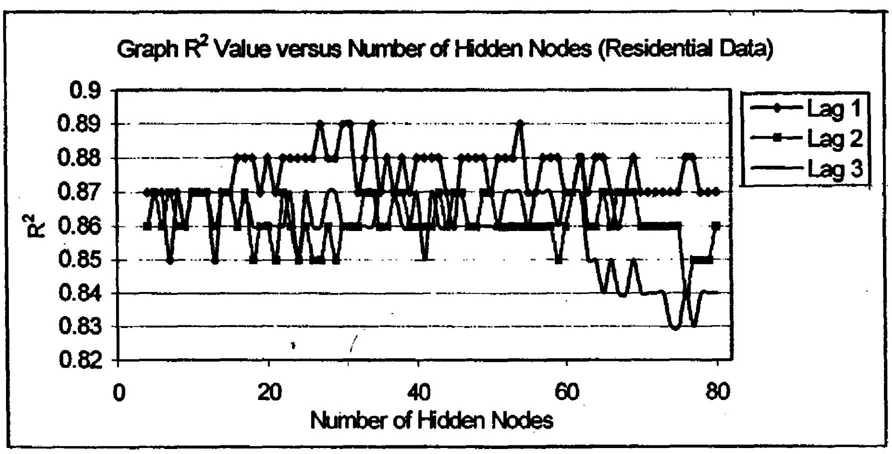

Figure 5: Graph R^2 Value versus Number of HRBF Center (Residential Data)

Industrial Data

The analysis to determine the best input lags for Industrial Data by using the Single Model produces results as shown in Figure 6. By observing the graph, the performance of Single Model appears to be superior by simply using input lag (t-1). Whenever the number of input lags used increases, the value of R^2 for multiple step ahead prediction also boosted. The model begin to perform pleasantly at the input lags starting from 25 above. Alike two previous data, a total of three most excellent input lags were selected to be used in the analysis to determine the optimize number of hidden nodes for Industrial Data. Table 5 shows the selected input lags for Industrial Data.

Table 5: Selected Input Lags for Industrial Data

Number	Input Lags
1	(t-1)(t-2)(t-3) ... (t-8)(t-9)
2	(t-1)(t-2)(t-3) ... (t-23)(t-24)
3	(t-1)(t-2)(t-3) ... (t-46)(t-47)

In general, every selected input lags listed in Table 5 produced a considerable good performance. The values of R^2 test obtained by using these input lags go above 0.4 for eight steps ahead prediction. Despite that, the second input lags that is (t-1)(t-2)(t-3) ... (t-23)(t-24) produce the most accurate prediction with the R^2 value exceeding or reaching 0.8 for every step ahead prediction. The outcome from the analysis to determine the optimized structure of Single Model for Industrial Data was shown in Figure 6 and Figure 7. Table 6 shows the optimized structure of Single Model for Industrial Data.

Table 6: The Highest R^2 Value Achieved For Each Lag

Lag	1	2	3
Number of Center	45	[illegible]	35
Step 1	0.84	[illegible]	0.85
Step 2	0.64	0.80	0.75
Step 3	0.56	[illegible]	0.67
Step 4	0.54	0.76	0.64
Step 5	0.53	[illegible]	0.61
Step 6	0.50	[illegible]	0.59
Step 7	0.48	[illegible]	0.59
Step 8	0.45	0.78	0.58

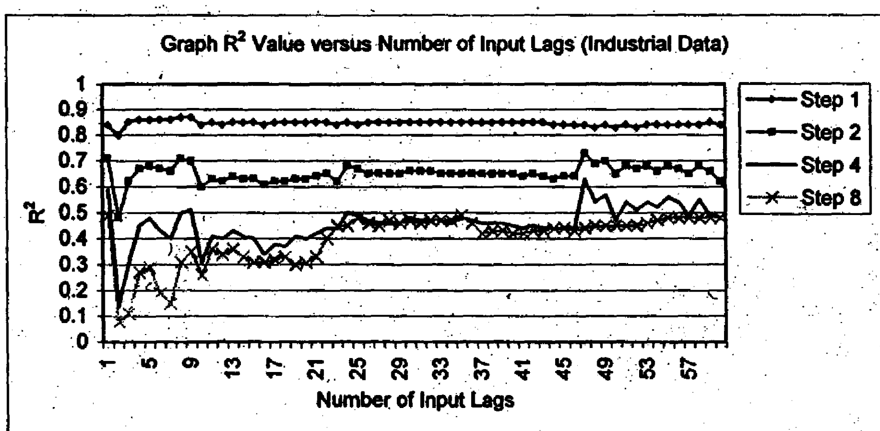

Figure 6: Graph R^2 Value versus Number of Input Lags (Industrial Data)

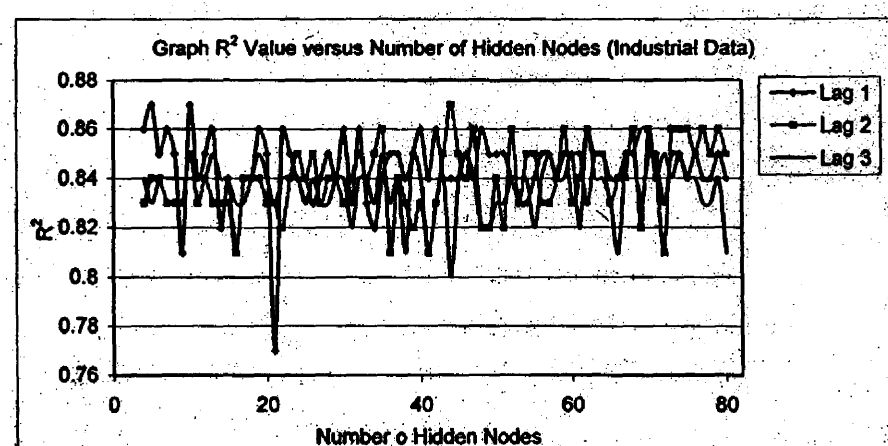

Figure 7: Graph R^2 Value versus Number of HRBF Center (Industrial Data)

Additional Analysis

By referring to the results obtained from the optimization analysis of the Single Model's structures, it can be concluded that the Single Model needs at least 24 average data of previous CO concentration to perform CO forecasting. Besides using the suitable input lags, the Single Model also need number of hidden nodes greater than 30 to produce the stable and accurate forecasting.

Although Single Model seems to perform well in forecasting three different categories of CO data, the capability of the Single Model to forecast unsteady data must been examined. To achieve this, a set of combined CO data will be introduced. The combined data consists of 6000 CO data by combining 2000 Residential Data, 2000 Traffic Data and 2000 Industrial Data sequentially. The shape of the combined CO data was shown in Figure 8 below.

The performance of the Single Model to predict the combined CO data was evaluated in the form of MSE test. Since the analysis was completed by using the combined CO data, the data used to calculate the MSE values were arranged by following the order of the combined data. An amount of 500 data from each category will be used to measure the MSE value in this analysis.

The MSE test was carried out by selecting the input lags of the Single Model as (t-1)(t-2)(t-3) ... (t-29)(t-30) and the number of hidden nodes as 35 while the value of $\mu(0)$ and q were taken as 0.95 and 0.30 respectively. The MSE values obtained were plotted versus training data as shown in Figure 9.

Referring to the MSE plot in Figure 9, we can observed that the Single Model will continuously updates the weights parameters to suite with every variation in the data. It can be observed that the MSE values start to converge after 100 training data and reach the good values after 1000 data. However when the pattern and values of the combined data changed drastically at point 2001, the weights of the Single Model again being revised to suite the new unseen data. The modifications of the weights caused the prediction value diverged from the actual values. However, after some times, the Single Model seems to be able to update the weights to match with the subsequent variation of data.

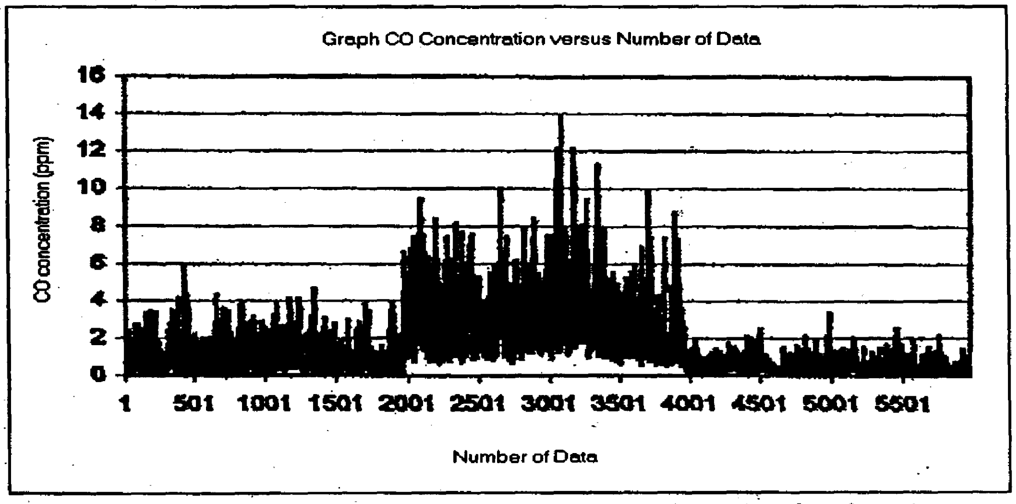

Figure 8: Graph CO Concentration versus Number of Data (Combined Data)

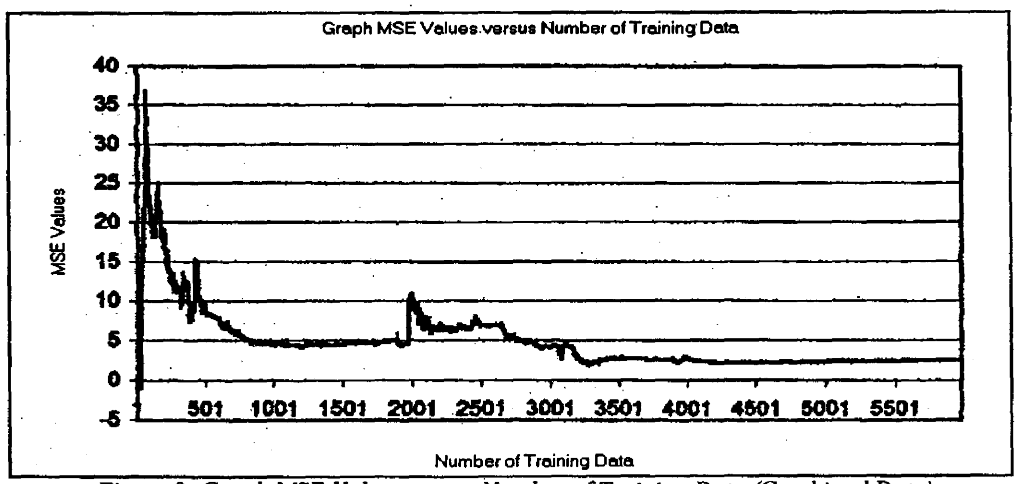

Figure 9: Graph MSE Values versus Number of Training Data (Combined Data)

Conclusion

The results have proved that the Hybrid Radial Basis Function (HRBF) network trained by on-line algorithms (Adaptive Fuzzy C-Means Clustering and Exponential Weighted Recursive Least Square Algorithm) can produce overwhelming forecasting results. However the accuracy of forecasted value depend on the complexity of the analyzed data and the selection of the optimized HRBF structure. Generally, the HRBF neural network has been shown to be a useful tool for CO prediction. This work proved that the HRBF network can model the relationship between past CO values with the present value in a time series without any external guidance. Consequently, this enables the model to be easily constructed.

It is known that Carbon Monoxide gas concentration tremendously depended on the meteorology circumstances in the surrounding environment. Thus, considering the meteorology data such as wind speed and relative humidity can be a useful attempt to improve the forecasted CO value.

Acknowledgements

We are grateful to Alam Sekitar Malaysia Sdn. Bhd. (ASMA) especially to Cik Haslina Mansor for making suitable data freely available and also the University Sains Malaysia for supporting this work in form of research grants.

References

[1] Cobb, N.; Etzel, R. 1991. Unintentionally Carbon Monoxide Related Deaths in The United States, 1979 through 1988. *JAMA*. 226, 659-663.

[2] Penney, D. G. 2002. A Website on Carbon Monoxide Toxicology.*www.phymac.med.wayne.edu/FacultyProfile/penney/COHQ/co1.htm.*

[3] Centers for Disease Control and Prevention. 1996. Deaths from Motor-Vehicle Related Unintentional Carbon Monoxide Poisoning Colorado, 1996, New Mexico, 1980 1995, and United States, 1979 1992. *http://www.cdc.gov/mmwr/preview/mmwrhtml/00044617.htm.*

[4] Jabatan Alam Sekitar Malaysia, Kementerian Sains, Teknologi dan Alam Sekitar. (1998). Laporan Kualiti Alam Sekeliling. 1998.*www.jas.sains.my/doe/bmlkas98.htm*

[5] Comrie, A. C.; and Diem, J. E. 2000. Climatology and Forecast Modeling of Ambient Carbon Monoxide in Phoenix, Arizona. *Atmospheric Environment* 4:24-40.

[6] Zickus, M., Kvietkus, K., Maroalka, A., and Auguline, V. 1996. An Investigation of Meteorological Effects on Urban Air Quality Using Carbon Monoxide Measurement Results in The Vilnius City (Lithuania). *Atmospheric Physics* 2.

[7] Boznar, M., Lesjak, M., and Mlakar, P. 1993. A Neural Network – Based Method for Short – Term Predictions of Ambient SO_2 Concentrations in Highly Polluted Industrial Areas of Complex Terrain. *Atmospheric Environment* 27B(2): 221-230.

[8] Perez, P., Trier, A., and Reyes, J. 2000. Prediction of $PM_{2.5}$ Concentrations Several Hours In Advance Using Neural Networks In Santiago, Chile. *Atmospheric Environment* 34: 1189-1196.

[9] Rumelhart, D. E., and McClelland, J. L. eds. 1986. *Parallel Distributed Processing*. MIT Press.

[10] Mashor, M. Y. 2001. Adaptive Fuzzy C-Means Clustering Algorithm for Radial Basis Function Network. *International Journal of Systems Science*, 32: 53-63.

[11] Karl J, A. ed. 1997. *Computer-Controlled Systems: Theory & Design*. New Jersey: Prentice Hall.

[12] Mashor, M.Y. 2000. Modification of the RBF Network Architecture. *ASEAN J. Sci. Technol. Dev.* 17(1), 87-108

[13] Ding, X., Canu, S., and Denceux, T. 1995. Neural Network Based Models for Forecasting. In Proceeding of ADT'95, 243-252.