Proceedings of the International Conference on Artificial Intelligence in Engineering & Technology
June 17 – 18 2002, Kota Kinabalu, Sabah, Malaysia

Carbon Monoxide Level Forecasting Using Neural 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,
University Sains Malaysia, Engineering Campus Nibong Tebal, Seberang Perai Selatan, Pulau Pinang, Malaysia
Tel: +604-5937788, Fax: 604-5941023, E-mail: rahman@eng.usm.my

[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

The neural network approach on forecasting Carbon Monoxide concentration is one topic of air quality research due to the health effects caused by Carbon Monoxide gasses in urban area. The neural network approach on forecasting time series is not new. However most of the existing are based on offline techniques. This paper proposes an online neural network approach to forecast Carbon Monoxide concentration. In this research, we investigate the performance of two different Radial Basis Function (RBF) neural network architectures (Single Model and Multiple Model) in forecasting multiple steps ahead Carbon Monoxide concentration. The RBF 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 and meteorological data from air quality monitoring station at Sekolah Menengah Victoria, Wilayah Persekutuan Kuala Lumpur given by Alam Sekitar Malaysia Sdn. Bhd. 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 best forecast can be achieved using the Single Model.

Keywords

Carbon Monoxide Forecasting, Neural Networks, Online Technique.

Introduction

Carbon Monoxide (CO) are emitted into the urban atmosphere mainly from vehicle exhausts. 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

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 [2],[8].

Another method in statistical modeling is Artificial Neural Network (ANN). It is well known that ANN can model nonlinear systems and it has been used to model Sulphur Dioxide concentrations in Slovenia [1] and to model $PM_{2.5}$ concentrations in Santiago, Chile [6]. In this paper, ANN was used to model and predict hourly CO concentrations from readily observable CO data and local meteorological data.

Approach and Methods

Data

The data for our 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 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 (SMV 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.

It is an established fact that atmospheric pollution depends strongly on meteorological conditions such as wind speed and temperature. In addition to this, we also investigate the effect of these meteorological parameters to the CO concentrations by calculating the cross correlation coefficient between CO concentrations and wind speed time series and between CO concentrations and temperature time series. The cross correlation coefficient, C_{ab} between series S_a and S_b is defined as [6]:

$$C_{ab} = \frac{\langle S_a S_b \rangle - \langle S_a \rangle \langle S_b \rangle}{\sqrt{\left(\langle S_a^2 \rangle - \langle S_a \rangle^2\right)\left(\langle S_b^2 \rangle - \langle S_b \rangle^2\right)}} \quad (1)$$

where $$\langle S_a \rangle = \frac{1}{N}\sum_{n=1}^{N} S_a(n), \quad (2)$$

$$\langle S_a^2 \rangle = \frac{1}{N}\sum_{n=0}^{N} S_a^2(n), \quad (3)$$

$$\langle S_a S_b \rangle = \frac{1}{N}\sum_{n=0}^{N} S_a(n) S_b(n) \quad (4)$$

and N is the number of data.

We found that the cross correlation coefficient between CO and temperature is -0.004 and between CO and wind speed, there is an anti correlation of –0.128. From the result, we can conclude that there was a very small correlation between CO concentrations and temperature but there was strong correlation between CO and wind speed. There exists an anti correlation between CO and wind speed because of strong winds will imply unfavorable conditions for particle accumulation in a given region.

Neural Network Predictors

The standard neural network method of performing time series prediction is to induce the function f in a standard eed forward neural network architecture, using a set of N-uples as inputs and a single output as the target value of the network. By using this method, the online forecasting on CO concentrations by using neural network was made. In the online technique, the network parameters are always updated whenever they receive new input. These make online technique yields better performance compared to offline 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 [7]. In this paper, Radial Basis Function (RBF) neural network with Adaptive Fuzzy C-Means Clustering Algorithm [5] and Exponential Weighted Recursive Least Square Algorithm [4] have been used to model the CO concentrations time series. The RBF network and the algorithms were chosen because it can be implemented in the online technique. Details about RBF neural network and the algorithms mentioned can be found in [5].

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 RBF to forecast CO value at time t+1 and the forecast value is then used as a new input to forecast CO value at t+2, t+3 and so on. The second architecture is to use a single RBF with as many outputs as values to forecast. The third architecture is to train a number of n RBF, one for each value to forecast. In this research, we have decided to use the first architecture (Single Model) and the third architecture (Multiple Model) as CO predictors.

Forecasting Performance

Ding, Canu and Denceux [3] 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 both models (Single and Multiple Model) have gone through two analyses, first to determine the best input lag and second to determine the correct number of RBF centers. Both analyses were made by replacing other RBF parameters with the typical values [5]. Both models were trained and tested on the SMV Data. The performance of each model is indicated in terms of coefficient of determination (R^2) given by

$$R^2 = 1 - \left(\frac{\sum_{t=n_a}^{n_t} (\hat{\varepsilon}(t))^2}{\sum_{t=n_a}^{n_t} (y(t) - \bar{y})^2} \right) \quad (5)$$

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

Single Model

The analysis to find out the best input lags for Single Model gave results as in Figure 1.

From Figure 1, 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 change. Apart from that the performance of model in multiple steps ahead forecasting was faded when the numbers of input lags used is smaller than 25. However model seems to perform well while operating in the range of input lags from 25 to 40. Further increase in the number of lags greater than 40 do not give much benefit to the model, in fact it deterioration the model performance. Three input lags which results the highest R^2 value for every step ahead prediction were selected and shown in Table 1.

Table 1: Selected Input Lags For Single Model

Number	Input Lags
1	$(t\text{-}1)(t\text{-}2)(t\text{-}3) \ldots (t\text{-}26)(t\text{-}27)$
2	$(t\text{-}1)(t\text{-}2)(t\text{-}3) \ldots (t\text{-}29)(t\text{-}30)$
3	$(t\text{-}1)(t\text{-}2)(t\text{-}3) \ldots (t\text{-}38)(t\text{-}39)$

The selected input lags was then used in the analysis to find out the correct number of RBF centers. This analysis yields the results shown in the Figure 2 below.

Figure 2 shows that every lag has similar plot. Thus it can be concluded that the RBF network needs number of centers greater than 25 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 r each lag. From this result, the best model performance can be achieved by using input lags $(t\text{-}1)(t\text{-}2)(t\text{-}3) \ldots (t\text{-}29)(t\text{-}30)$ and the number of center 47.

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

Lag	1	2	3
Number of Center	47	47	64
Step 1	0.70	0.72	0.76
Step 2	0.49	0.51	0.48
Step 3	0.51	0.54	0.43
Step 4	0.51	0.53	0.41
Step 5	0.50	0.53	0.42
Step 6	0.51	0.53	0.42
Step 7	0.52	0.54	0.44
Step 8	0.52	0.53	0.43

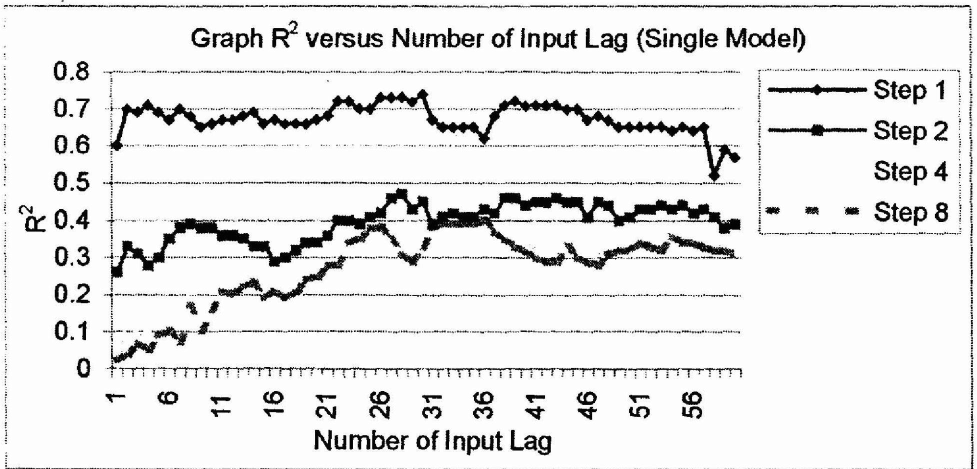

Figure 1: Graph R^2 Value versus Number of Input Lag (Single Model)

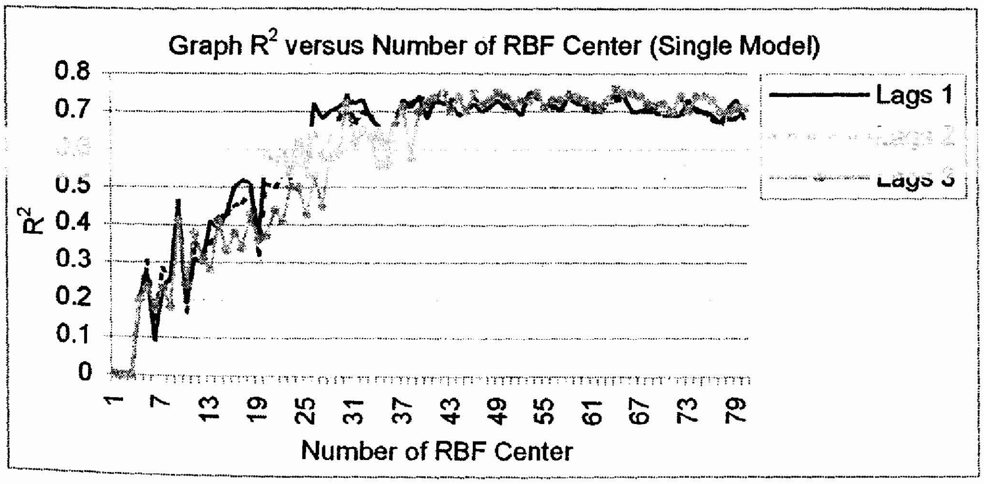

Figure 2: Graph R^2 Value versus Number of RBF Center (Single Model)

Multiple Model

The analysis to find out the best input lags by using Multiple Model yields results shown in Figure 3. By considering the graph plot in Figure 3, it can be noted that Multiple Model is capable to performing one step ahead prediction by just using two 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 0.1. However the performance of the model increases when the number of input lags increases. Overall it can be concluded that the model gives excellent performance at the number of input lags 20 to 38. Further increase in the number of input lags will degrade the model performance.

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

The result of analysis to find out the RBF center for Multiple Model was given in Figure 4 below. Table 4 shows the highest R^2 value achieved for each lag. From the result, the best performance for Multiple Model can be achieved by using input lags (t-1)(t-2)(t-3) ... (t-21)(t-22) and number of center 47.

Table 3: Selected Input Lags for Multiple Model

Number	Input Lags
1	(*t*-1)(*t*-2)(*t*-3) ... (*t*-21)(*t*-22)
2	(*t*-1)(*t*-2)(*t*-3) ... (*t*-26)(*t*-27)
3	(*t*-1)(*t*-2)(*t*-3) ... (*t*-32)(*t*-33)

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

Lag	1	2	3
Number of Center	47	32	44
Step 1	0.74	0.73	0.75
Step 2	0.41	0.43	0.49
Step 3	0.47	0.39	0.46
Step 4	[illegible]	0.38	0.42
Step 5	[illegible]	0.39	0.41
Step 6	[illegible]	0.41	0.41
Step 7	[illegible]	0.41	0.40
Step 8	[illegible]	0.4	0.40

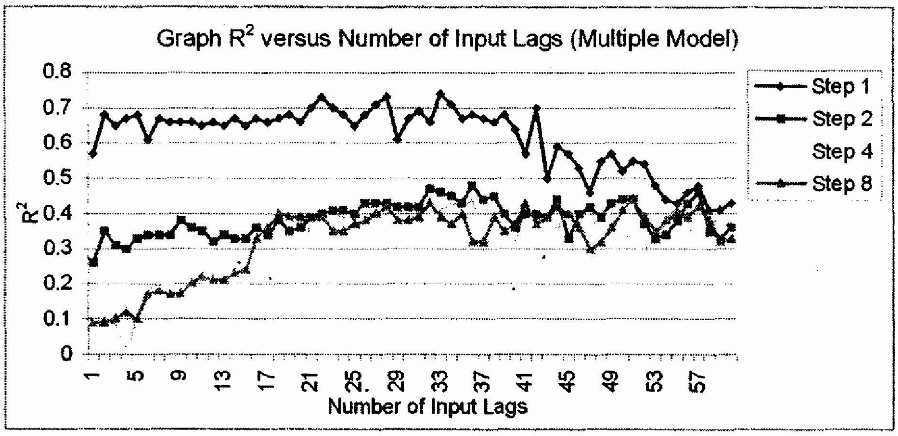

Figure 3: Graph R2 Value versus Number of Input Lags (Multiple Model)

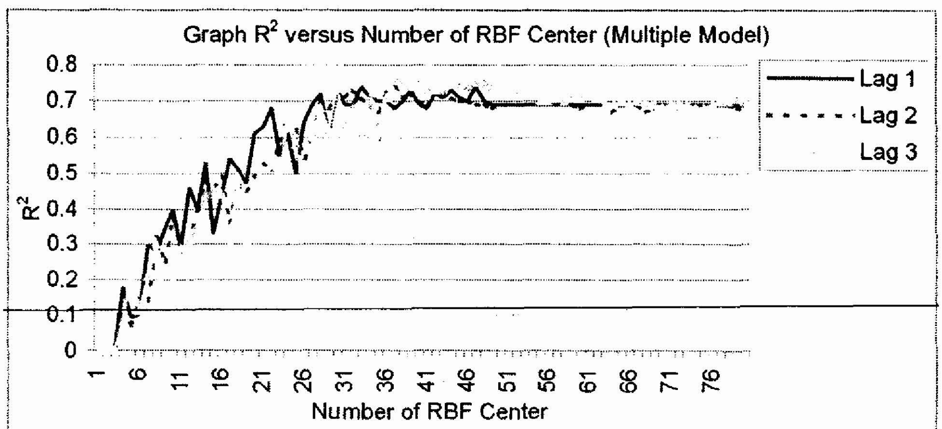

Figure 4: Graph R Value versus Number of RBF Center (Multiple Model)

By observing the results obtained from the Single Model and the Multiple Model, it can be concluded that the Single Model gives better performance as compared to the Multiple Model. Therefore the Single Model will be used to test the effect of considering the meteorological variables to the model performance. This research aims to study the effects of including wind speed data as an additional input to the model. The effect of other meteorological parameters such as wind direction and temperature was to be neglected because there were no strong correlation showed between them and the CO concentration. Because our aims is to predict CO for eight steps ahead, the wind speed data for eight steps ahead must be ready. To achieve this, the wind speed value for eight steps ahead must be forecasted. We have used the RBF network with the same algorithms used to forecast CO concentration to forecast wind speed.

As in CO forecaster, the analysis to determine the best input lags and the best RBF network structure for wind speed forecaster have also been carried out. Table 5 shows the R^2 value achieved for wind speed forecasting. The result shows that the optimized RBF structure to represent wind speed forecasting can be achieved by using input lags of (t-1)(t-2)(t-3) ... (t-25)(t-26) with the number of 30 centers.

The wind speed forecaster was then combined with the Single CO forecaster to perform two stages forecasting. The input lags and the number of centers in both forecasters have been set to the best value achieved in the two analysis described earlier. In general, the combined model can be described as in Figure 5. The results obtained from the combined forecaster in R^2 value is shown in Table 6.

Table 5: R^2 Value for Wind Speed Forecaster

Input Lags	(t-1)(t-2)(t-3) ... (t-25)(t-26)
Number of Center	47
Step 1	0.87
Step 2	0.80
Step 3	0.75
Step 4	0.71
Step 5	0.70
Step 6	0.70
Step 7	0.70
Step 8	0.70

Table 6: R^2 Value for Combined Forecaster

Input Lags	R^2 Value
Step 1	0.76
Step 2	0.56
Step 3	0.53
Step 4	0.53
Step 5	0.52
Step 6	0.51
Step 7	0.51
Step 8	0.51

When using the Combined Forecaster, the R^2 value obtained for one and two steps ahead prediction are higher than the R^2 value obtained by the Single Model. Nevertheless the R^2 value obtained for three to eight steps ahead prediction are slightly lower compared to the result achieved by the Single Model. These can be explained by the degree of exactness of wind speed forecasting. Because of the strong correlation between wind speed and CO, a small error in the forecasted wind speed can alter the performance of the CO forecaster.

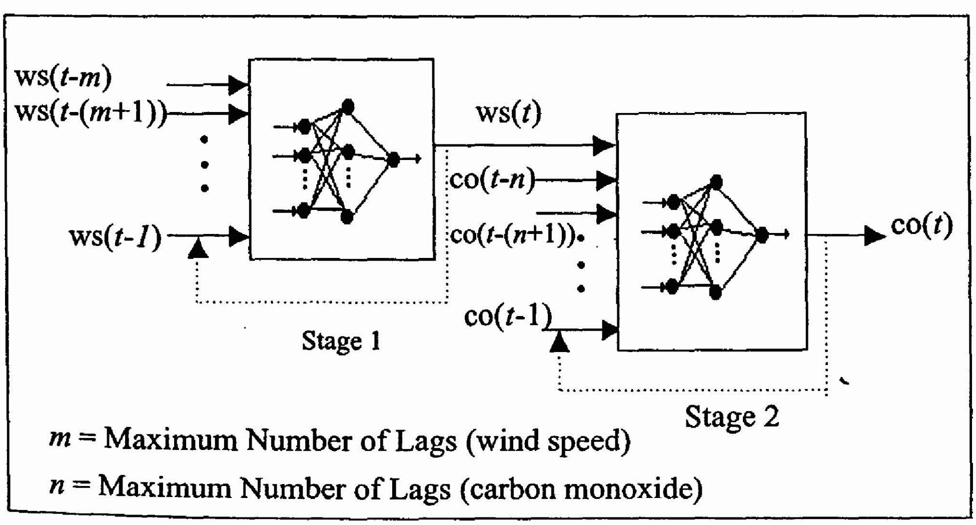

Figure 5: The Combined Wind Speed Forecaster and Carbon Monoxide Forecaster

Conclusion

From the results, it can be concluded that the Single Model and the Multiple Model can perform well in one step ahead prediction but the single model give better performance in multiple step ahead prediction. The performance of the Single Model can be improved by including wind speed data as additional input to forecast CO concentrations.

Overall the Radial Basis Function neural network has been shown to be a useful tool for CO prediction. This work has proved that the Radial Basis Function 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.

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] 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.

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

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

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

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

[6] 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.

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

[8] 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.