

# Prediction of maximum daily ozone level using combined neural network and statistical characteristics

Wenjian Wang<sup>a</sup>, Weizhen Lu<sup>b,\*</sup>, Xiekang Wang<sup>c</sup>, Andrew Y.T. Leung<sup>b</sup>

<sup>a</sup>Department of Computer Science, Shanxi University, Taiyuan, Shanxi 030006, PR China

<sup>b</sup>Department of Building and Construction, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong, HKSAR, PR China

<sup>c</sup>State Key Laboratory Hydraulics of High Speed Flows, Sichuan University, Chengdu, Sichuan 610065, PR China

Received 8 October 2002; accepted 10 January 2003

## Abstract

Analysis and forecasting of air quality parameters are important topics of atmospheric and environmental research today due to the health impact caused by air pollution. As one of major pollutants, ozone, especially ground level ozone, is responsible for various adverse effects on both human being and foliage. Therefore, prediction of ambient ozone levels in certain environment, especially the ground ozone level in densely urban areas, is of great importance to urban air quality and city image. To date, though several ozone prediction models have been established, there is still a need for more accurate models to develop effective warning strategies. The development of such models is difficult because the meteorological variables and the photochemical reactions involved in ozone formation are very complex. The present work aims to develop an improved neural network model, which combines the adaptive radial basis function (ARBF) network with statistical characteristics of ozone in selected specific areas, and is used to predict the daily maximum ozone concentration level. The improved method is trained and testified by hourly time series data collected at three air pollutant-monitoring stations in Hong Kong during 1999 and 2000. The simulation results demonstrate the effectiveness and the reliability of the proposed method.

© 2003 Elsevier Science Ltd. All rights reserved.

**Keywords:** ARBF network; Air quality forecasting; Ozone; Statistical characteristics

## 1. Introduction

In recent years, significant efforts such as the emission controls of nitrogen oxide and volatile organic compounds have been directed to reduce the ozone concentrations in the areas where maximum acceptable ozone levels are violated frequently. For example, the current US National Ambient Air Quality Standard requires that the ozone level should not exceed 0.12 ppm (1-h average) more than once a year for a period of 3 years. To facilitate implementations of the maximum ozone concentrations and develop effective warning strategies, reliable forecast of the ambient ozone concentrations, especially in densely populated areas, is essential. Although some classical approaches, e.g., deterministic or stochastic process, have been applied to analyze ozone signals, inherent spatial and temporal variability in ozone makes the basic equations only an approximation

whose value is conditional on appropriate calibration through numerous tuning parameters. In addition, due to the influence of meteorological conditions on ozone concentrations and large uncertainty associated with input weather data, it is very difficult to obtain a good agreement between the dynamic model and the observed data, needless to say a good prediction (Seinfeld, 1988; Rao and Zurbenko, 1994; Chang and Suzio, 1995).

At present, the use of neural networks, and in particular, the multilayer feed-forward neural networks, which can be trained to approximate virtually any smooth, measurable function, have become popular in environmental science and produced promising results (Boznar et al. 1993; Comrie, 1997; Gardner and Dorling, 1996, 1998; Roadknight et al., 1997; Song and Hopke, 1996; Reich et al., 1999; Fan et al., 2000; Lu et al., 2002a,b; Yi and Prybutok, 1996). In this paper, an improved neural network model, which combines the ARBF network with the statistical characteristics of ozone, is presented to predict the daily maximum ozone level. For illustrating the effectiveness of the proposed approach, the predicting performance, robustness,

\* Corresponding author. Tel.: +852-27844316; fax: +852-27887612.

E-mail address: [bcwzlu@cityu.edu.hk](mailto:bcwzlu@cityu.edu.hk) (W. Lu).

and inherent problems existing in traditional neural networks have been particularly inspected.

## 2. Air quality and ozone pollution in Hong Kong

Hong Kong is one of the most developed metropolitans and has probably the highest population density in the world. With the continuous economic development and the population expansion, a series of severe environmental pollution problems have attracted much attention in recent years, e.g., air pollution, noise pollution, shortage of land resource, waste and sewage disposal, etc. Among these, air pollution has direct effect on human health through exposure to pollutants at high concentration levels existing outdoors, e.g.,  $O_3$  pollution. The poor air quality in Hong Kong has become a focal point for anyone concerning the environmental benign. Air pollution control is needed to prevent the situation getting worse in the long run. On the other hand, short-term prediction of air pollutant level is also required in order to take preventive and evasive action during episodes of airborne pollution.

It is known that ozone ( $O_3$ ) is an effective anti-greenhouse gas particularly in the upper troposphere, thus playing a direct role in climate change. Ozone regulates the oxidizing capacity of the atmosphere via production of the OH radical that acts as the principal cleaning agent in the atmosphere. In the lower atmosphere, elevated ozone is a pollutant and has adverse effects on both human health and foliage. For example, elevated ozone levels may cause eye irritation, cough, reduced athletic performance and possible chromosome damage, etc. (Larsen et al., 1991; Lippmann et al., 1983; Spector et al., 1991). Therefore, prediction of ambient ozone level in urban areas would provide evaluation of compliance and noncompliance with requirement of environmental protection. To date, certain studies on ozone prediction can be traced in literature (Gardner and Dorling, 1996; Chen et al., 1998; Hadjiiski and Hopke, 2000). However, rare studies can be found on prediction of ground ozone levels, e.g., space below 20-m height, which may adversely affect human health, especially in densely populated areas like Hong Kong.

According to Air Quality Objectives (AQO) stipulated by the Hong Kong Environmental Protection Department (HKEPD), the standard ozone level in ambient is 0.123 ppm (i.e.,  $240 \mu\text{g}/\text{m}^3$ ), and the hourly average  $O_3$  level should not exceed such standard more than three times in a year (Hong Kong Environmental Protection Department Annual Report, 1996). To monitor the ambient ozone level in Hong Kong, the HKEPD established three air pollution monitoring stations in Yuen Long, Central/Western, and Kwai Chung in 1995. According to the on-site measurement and preliminary statistical analysis, the AQO limit for  $O_3$  was violated more than three times in both Kwai Chung and Central/Western sites in 1996. The records in 1997 and 1998 are basically complied with the AQO limit, respectively

(Hong Kong Environmental Protection Department Annual Report, 1997, 1998). With the increasing requirement of good air quality, the ozone monitoring actions have been extended to other 11 monitoring stations in 2000 (Hong Kong Environmental Protection Department Annual Report, 2000). In severe situations, the highest 1-h average ozone levels of 335 and  $314 \mu\text{g}/\text{m}^3$  were recorded at Tung Chung station in 1999 and 2000, respectively (Hong Kong Environmental Protection Department Annual Report, 1999, 2000). According to the AQO limits, the current ozone levels in Hong Kong should not be complained through the observations in majority of monitoring stations. However, recent studies indicate that exposure to  $160 \mu\text{g}/\text{m}^3$  (0.08 ppm) for 6.6 h in a group of healthy exercising adults may lead to a decrease in lung functions from more than 10% to the most sensitive individual. Moreover, repeated exposures to a given concentration (e.g., 6.6 h to 0.08, 0.10, and 0.12 ppm, etc.) during several consecutive days may cause attenuation of functional changes but persistence of airway hyperresponsiveness (Folinsbee et al., 1994). These outcomes indicate the effect of ozone on human health, although the average 1-h ozone level is lower than the  $O_3$  AQO limit of  $240 \mu\text{g}/\text{m}^3$  (0.12 ppm) in most monitored areas. Therefore, predicting  $O_3$  levels in Hong Kong will still be important.

To better explain the proposed neural network model, it is necessary to analyze the trend of the daily maximum  $O_3$  levels in the interested areas. Tsuen Wan area, one of the three interested areas (Tsuen Wan, Kwai Chung, and Kwun Tong) in Hong Kong, is taken as an example to illustrate the tendency of the daily maximum  $O_3$  level. Fig. 1 shows the distribution of the daily maximum  $O_3$  concentration level in Tsuen Wan area during 1999 and 2000. From Fig. 1, it can be seen that in both years, the daily maximum  $O_3$  concentration levels are lower and have mild scatter in May through July. While in other months, the daily maximum  $O_3$  concentrations have larger values and higher scatters. The highest  $O_3$  concentration levels achieved  $201 \mu\text{g}/\text{m}^3$

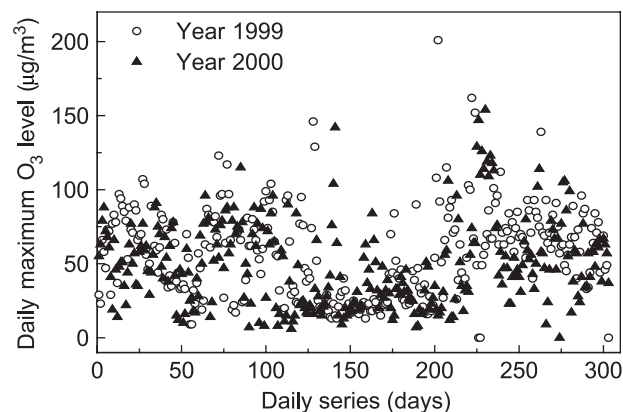


Fig. 1. The daily maximum level of  $O_3$  concentration in Tsuen Wan area in 1999 and 2000.

(the 202nd data point, July 21, 1999) and  $154 \mu\text{g}/\text{m}^3$  (the 230th data point, August 8, 2000), respectively. The annual averages of daily maximum  $\text{O}_3$  concentration levels are 56.8 and  $49.27 \mu\text{g}/\text{m}^3$ , respectively. Due to the similar seasonal changing trends in both years, the data in 1999 can be used to train the proposed neural network model, while the corresponding daily maximal  $\text{O}_3$  concentration levels in 2000 can be used as testing samples and predicted as well.

### 3. Neural network prediction and statistical analysis of ground ozone level

#### 3.1. The RBF and ARBF neural network models

Two problems have to be considered in the application of neural networks: the first one is the learning speed, and the second is the determination of network architecture. Concerning the first problem, we use the radial basis function (RBF) network, which is regarded as an effective algorithm for fast learning in feed-forward neural networks. It possesses good generalization performance, and, meanwhile, can avoid over-elaborated, lengthy computing like well-known back-propagation (BP) algorithm (Broomhead and Lowe, 1988). An RBF network is a feed-forward neural network with a single hidden layer, which can produce outputs of radial basis function directly. The input of each radial basis function within an RBF neural network is the distance between the input vector (activation) and its center (location). The RBF network model is motivated by the locally tuned responses observed in biologic neurons. The neurons with a local tuned responding characteristic can be found in several parts of the neural system, and these locally tuned neurons show responding characteristics bounded to a small range of the input space. Such processes lead to wide applications of RBF network in many disciplines, e.g., environmental science. To handle the second problem, an adaptive method is needed to decide the number of hidden nodes in network, i.e., the adaptive RBF (ARBF) model—an improved RBF network, which can determine the number of hidden nodes dynamically. It is known that the numbers of input and output nodes vary from case to case. Hence, the number of hidden nodes is usually decided by experiences. As a result, if the number is too small, the network may fail to converge to the minimum during training. Conversely, excessive nodes may result in overfitting during training, and lead to poor generalization performance as well. Generally, a trial and error method is applied in most cases at the penalty of heavy computational burden and low efficiency. Therefore, in this study, the adaptive ARBF network, an adaptive constructing method of determining the number of hidden nodes, is used. For given training error, the ARBF network will automatically add nodes and links into the network when needed until the actual error is below the given one. The

main concept of ARBF network can be briefly described as follows.

- (1) Adding hidden nodes (neurons): the number of added hidden nodes (neurons) is taken from the range  $[1, H]$ , where  $H$  is the maximum number of hidden nodes added. The hidden neurons are arranged in sequence and the location of each new neuron will not affect the existing subsequent connections. Thus, if a new neuron is added to the network, it can be inserted into the existing sequence of hidden neurons at any location.
- (2) Adding links: when a hidden neuron is added, the links between the added neuron and each input and output neuron must be inserted at the same time. There are no any links between the added neuron and other neurons in the hidden layer.
- (3) The training process of network continues until the training error is below the given error. When the training process finishes, the size of network may be larger than that of the best network, but there are fewer free parameters need to be estimated and the algorithm is easy to be implemented.

Similar to traditional RBF network, the famous Gaussian function, defined as  $f(x) = \exp(-(x - \mu)^2 / (2\sigma^2))$ , is used as the nonlinear transfer function in the hidden layer (in this study, the center and spread constants of the radial basis function are assigned as  $\mu = 0$  and  $\sigma = 0.1$ , respectively). To avoid “overfitting” training, the maximum number of hidden nodes is prescribed. If the network cannot achieve specified target even with the maximum hidden nodes, only the spread constant,  $\sigma$ , need to be widened or narrowed. The merits of ARBF network with such dynamic structure have been testified by some researchers (Wang et al., 1993; Wang, 2000; Lu et al., 2002c).

#### 3.2. The statistical characteristics of ozone

Concerning selecting independent input variables to predict the daily maximum  $\text{O}_3$  levels, factors that influence both photochemical production and atmospheric accumulation of  $\text{O}_3$  need to be considered concurrently (Robeson and Steyn, 1990). Generally, these include vehicle emissions such as nitrogen oxides ( $\text{NO}_x$ ), nitric oxide (NO), nitrogen dioxide ( $\text{NO}_2$ ), carbon monoxide (CO), etc., and meteorological parameters such as air velocity and direction, temperature and humidity, solar radiation, etc. In this study, the available original database contains seven pollutants, i.e.,  $\text{O}_3$ ,  $\text{NO}_x$ , NO,  $\text{NO}_2$ , CO, sulphur dioxide ( $\text{SO}_2$ ), and respirable suspended particles (RSP), and five meteorological factors, i.e., wind speed (WS), wind direction (WD), outdoor temperature (OT), solar radiation (SR), and indoor temperature (IT), which were hourly measured during the whole years of 1999 and 2000 (Hong Kong Environmental Protection Department Annual Report, 1999, 2000).

Based on the database, the correlation analysis has to be used to analyze the interrelations among the pollutants and the meteorological factors so as to determine major factors, i.e., input variables in network model, affecting the  $O_3$  levels in the interested areas. In general,  $O_3$  is closely related to pollutants like  $NO_2$ ,  $NO$ , and  $NO_x$  according to photochemical oxide interaction in local environment (Lu et al., 2002b). At the same time, the variation of ozone level is also affected by the meteorological conditions including wind speed, wind direction, solar radiation, temperature, etc. The proceeded statistical analyses show that the  $O_3$  levels in selected areas are negatively relevant to nitric oxide and positively to nitrogen dioxide, weakly affected by  $CO$  and hardly affected by  $SO_2$ ,  $RSP$ , and  $IT$ , increase with the increment of solar radiation and temperature and the decrement of wind speed (i.e., the range of wind direction variation may cause high-changing frequency of  $O_3$  but its concentration levels). Hence, the following equations can be obtained:

$$\begin{aligned} \text{For year 1999: } O_3 &= 12.989(NO_2/NO) + 2.0513 \\ R^2 &= 0.739 \end{aligned} \quad (1)$$

$$\begin{aligned} \text{For year 2000: } O_3 &= 12.958(NO_2/NO) + 1.2185 \\ R^2 &= 0.7616 \end{aligned} \quad (2)$$

Here,  $R$  denotes the correlation function. Eq. (1) expresses the strong linear relation between  $O_3$  level and

$NO_2/NO$  in 1999, while Eq. (2) depicts similar relation in 2000. Fig. 2 presents the regression profiles between  $O_3$  and  $NO_2/NO$  in Tsuen Wan area in both 1999 and 2000 corresponding to Eqs. (1) and (2), respectively (Lu et al., 2002b). Therefore, the parameters  $CO$ ,  $SO_2$ ,  $RSP$ , and  $IT$  will not be considered in the proposed ARBS network model due to their small contributions to the resultant  $O_3$  concentration levels such that the network architecture would be simple and the fast learning speed would be achieved as well.

### 3.3. Combination of ARBF network and statistical characters of ozone

Considering the inherent spatial and temporal variability of emission concentrations, the influence of meteorological conditions, and the uncertainties associated with initial and boundary conditions, it is very difficult to model, calibrate, and validate ozone variations from first principles. Due to the power error-tolerance ability of neural networks, the statistical characteristics of ozone must be considered and combined with the developed ARBF neural network model in this study.

According to the analysis in Section 3.2, it is necessary to consider the impacts of  $NO$ ,  $NO_2$ ,  $NO_x$ , and relevant meteorological parameters when establishing an ozone prediction model with reliable performance. By analyzing the

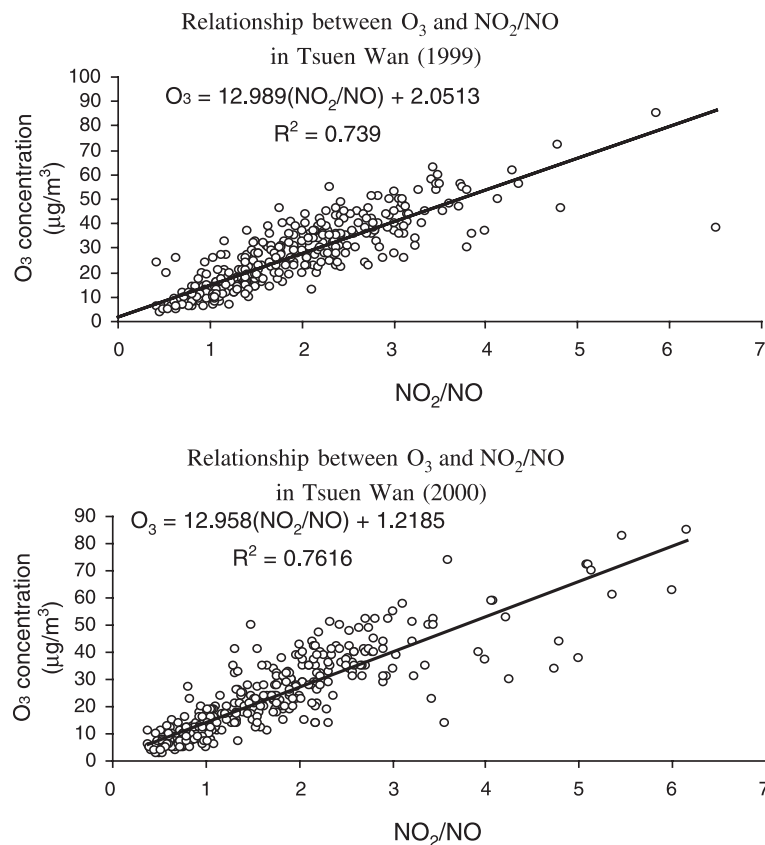


Fig. 2. Regression analysis between  $O_3$  and  $NO_2/NO$  in Tsuen Wan in 1999 and 2000 (Lu et al., 2002b).



data in Tsuen Wan area in 1999 and 2000, the corresponding linear relations between  $\text{NO}_x$  and NO can be found in both years, listed as:

$$\begin{aligned} \text{For year 1999: } \text{NO}_x &= 1.9853\text{NO} + 44.83 \\ R^2 &= 0.8293 \end{aligned} \quad (3)$$

$$\begin{aligned} \text{For year 2000: } \text{NO}_x &= 1.76\text{NO} + 48.322 \\ R^2 &= 0.9003 \end{aligned} \quad (4)$$

These two equations illustrate that the effect of NO can be analogous and substituted with that of  $\text{NO}_x$  on predicting the  $\text{O}_3$  level without compromising the predicting performance. To simplify the model (i.e., reduce the input variables and the hidden nodes as well), the input vectors will exclude

the NO variable. Therefore, in the proposed model, three pollutants (i.e.,  $\text{O}_3$ ,  $\text{NO}_2$ , and  $\text{NO}_x$ ), and three meteorological variables (i.e., WS, SR, and OT), are selected as input vectors of the ARBF network. Here, the wind direction (WD) is also omitted because it rarely changes during the same season across the Hong Kong territory, and the prediction of daily maximum  $\text{O}_3$  level will be performed in different season intervals, i.e., spring (March–May), summer (June–August), autumn (September–November), and winter (January and February at the beginning of year and December at the end of year).

To illustrate and assess the effectiveness of the proposed ARBF model, the general RBF network model is used as a benchmark. For the general RBF network, the input variables include pollutants  $\text{NO}_x$ , NO,  $\text{NO}_2$ , CO, and  $\text{O}_3$ , and meteorological factors WS, WD, SR, and OT, according to

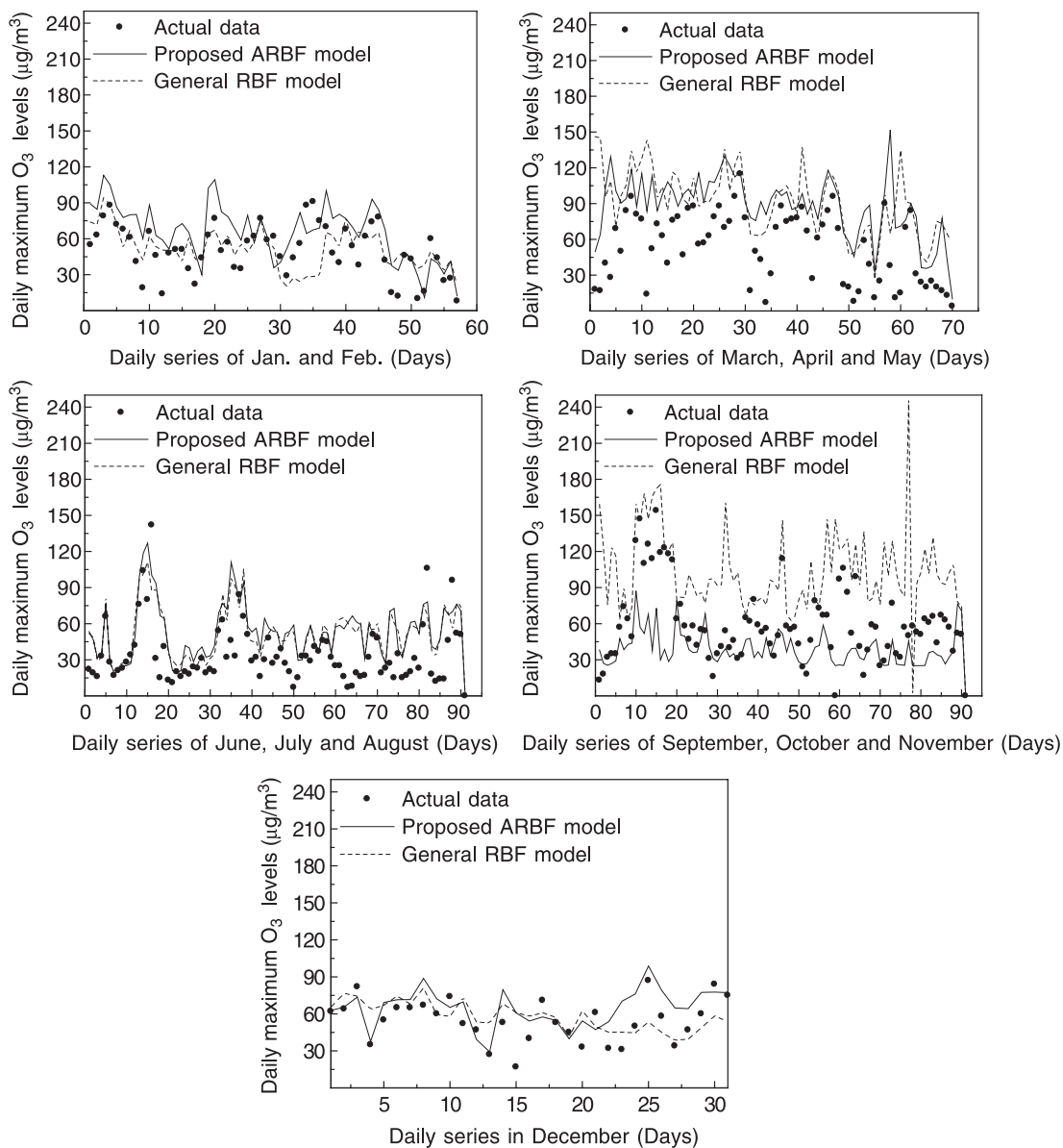


Fig. 3. Simulation results of  $\text{O}_3$  level in Tsuen Wan area in 2000.

Table 1

Comparison of predicting absolute errors between two neural network models ( $\mu\text{g}/\text{m}^3$ )

Season	The ARBF network			General RBF network		
	Max error	Min error	Average error	Max error	Min error	Average error
January and February	44.9541	0.0103	<b>19.6440</b>	62.6228	0.0206	<b>13.8243</b>
Spring	113.8809	<b>1.8839</b>	34.6000	129.321	<b>1.6895</b>	38.3682
Summer	<b>65.0142</b>	0	<b>22.2976</b>	<b>61.7273</b>	0	<b>21.8318</b>
Autumn	89.9173	0.0530	25.0636	216.0058	0.0623	45.9952
December	43.8207	<b>0.7253</b>	14.3745	44.1556	<b>0.6710</b>	14.4147
Average	71.5174	<b>0.5345</b>	23.1959	102.7665	<b>0.4887</b>	26.8868

the primary principle of the ozone formation. Both models are validated by the same database in both 1999 and 2002. The simulations include both training and prediction processes. The data in 1999 are used as training samples and the ones during the same seasons as testing samples.

#### 4. Experiments—case studies

The data collected at three monitoring stations, i.e., Tsuen Wan, Kwai Chung, and Kwun Tong, are selected as sample data, which can reveal the characteristics of ozone level across the territory of Hong Kong. These three ambient monitoring stations have similar environmental features, i.e., they are all located in densely populated residential areas mixed with commercial and industrial developments, where ozone levels are usually higher than other areas. In the study, the data at Tsuen Wan monitoring station in 1999 are used as the training set to train the proposed ARBF network model and the general RBF network model, while the data in the corresponding areas in 2000 are respectively employed as the testing set.

Fig. 3 presents the comparisons of prediction results on the testing data in Tsuen Wan area in 2000 during which the data in same seasons are as training data in 1999. From Fig. 3, it can be seen that the prediction results generated by the proposed ARBF model are getting closer to the actual data than the ones produced by the general RBF network model

during the year 2000. It is also noticed that the simulating results in summer (June, July, and August) and winter (December, January, and February) are better than those in spring (March, April, and May) and autumn (September, October, and November). The reason may be that the same values of free parameters are used in the Gaussian function for different seasons. It would be better if these parameters are adjusted individually to different seasons. Generally, the summer represents the heaviest pollution situation for  $\text{O}_3$  compliance with the respective maximum absolute errors of daily maximum  $\text{O}_3$  level prediction at the values of 65.0 and 61.7  $\mu\text{g}/\text{m}^3$  based on these two models. Winter means to the lightest occasion for  $\text{O}_3$  noncompliance corresponding to the maximum absolute errors of the daily maximum  $\text{O}_3$  levels with 44.9/62.6  $\mu\text{g}/\text{m}^3$  in January/February and 43.8/44.1  $\mu\text{g}/\text{m}^3$  in December, respectively. Besides, sunny conditions and average wind speed change frequently in spring and autumn. It is not easy to find the changing rules of daily means of  $\text{O}_3$  concentration (Lu et al., in press), needless to say the daily maximum  $\text{O}_3$  level. In this study, the proposed model can achieve satisfactory predicting results in spite of the influences of the above unstable meteorological factors. Comparing with the general RBF network, the proposed method can provide a more general approach to predict the daily maximum  $\text{O}_3$  level under certain microenvironment and perform well in different seasons due to the mild seasonal variations in Hong Kong. Although the number of input variables related to  $\text{O}_3$  is decreased in the proposed method, the forecasting performances are not worsening because the inherent factors affecting  $\text{O}_3$  level are considered here. Another advantage is that the architecture of the proposed network is, after all, simpler than that of the general RBF network due to less input variables.

Table 1 describes the comparisons of generalization performances between the proposed ARBF network and the general RBF network. By comparing the seasonal maximum, minimum, and average absolute errors of the daily maximum  $\text{O}_3$  levels, the ARBF model produces lower errors than the general RBF network does except for individual points, on which very small deviations exist between the two methods (marked in bold in Table 1). For

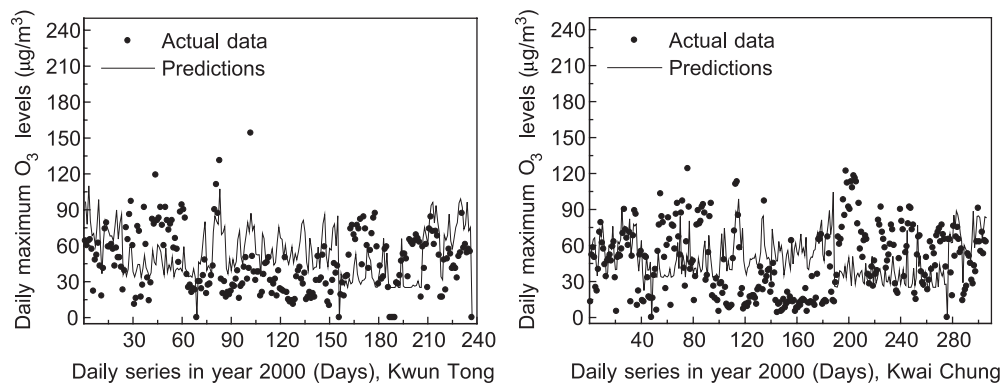
Fig. 4. Simulation of  $\text{O}_3$  levels in Kwai Chung and Kwun Tong areas in 2000.

Table 2  
Absolute prediction errors based on ARBF network in three interested areas in 2000 ( $\mu\text{g}/\text{m}^3$ )

Area	Max error	Min error	Average error
Tsuen Wan	113.8809	0	23.1959
Kwai Chung	90.5933	0	26.3465
Kwun Tong	74.7810	0	24.7912

the ARBF model, the means of maximum and average errors during the year 2000 are 71.5 and 23.2  $\mu\text{g}/\text{m}^3$ , respectively, while those produced by the general RBF network are 102.8 and 26.9  $\mu\text{g}/\text{m}^3$ , correspondingly. Therefore, it can be concluded that the ARBF network is superior to the general RBF network.

To further testify the generality and reliability of the proposed ARBF method, the data collected at Kwai Chung and Kwun Tong monitoring stations are used as additional testing sets to assess the model. Such experiments aim to verify whether the model is effective and suitable to the similar microenvironments. As mentioned above, both Kwai Chung and Kwun Tong areas are located in similar surroundings to that of Tsuen Wan area. If the model can produce satisfactory prediction results in these two areas, the robustness of the proposed model will then be proven.

Similar to Tsuen Wan area, the seasonal simulating experiments are carried on in Kwai Chung and Kwun Tong areas, and similar predicting conclusions can also be obtained. Fig. 4 shows the simulating results at these two areas during the year 2000. For these two areas, the predicting results in winter are better than in summer. The predictions for spring and autumn show less accuracy than for winter or summer. It can be seen that the ARBF method can also produce satisfactory results in these two interested areas. The maximum and average absolute errors of the daily maximum  $\text{O}_3$  take values of 90.6 and 26.3  $\mu\text{g}/\text{m}^3$  in Kwai Chung area, respectively. While in Kwun Tong area, the corresponding values are 74.8 and 24.8  $\mu\text{g}/\text{m}^3$ .

Table 2 lists the comparison of absolute errors in these three interested areas during the year 2000. It indicates that the proposed ARBF model is robust and effective due to the implementation of statistical characteristics to the model. Therefore, it can be used to the similar microenvironments, which can be easily found in Hong Kong territory.

## 5. Conclusions

This paper presents an improved neural network model, i.e., the adaptive radial basis function (ARBF) model, which combines the general RBF network with statistical characteristics of ozone  $\text{O}_3$ . It is developed and used to predict the daily maximum ozone level in relevant interested environments in Hong Kong territory. The model possesses some merits. Firstly, it can provide better predicting results and simpler network architecture than the normal RBF network.

Secondly, it is robust and less affected by seasonal factors; hence, some highly correlative pollutants, e.g.,  $\text{NO}$ ,  $\text{NO}_2$ , etc., may not be measured due to their strong correlation to  $\text{O}_3$ . Their characteristics and changing patterns are similar to that of  $\text{O}_3$  and then can be deduced from  $\text{O}_3$  records. Finally, the model is extensible and reproducible. It can be used in the areas with similar environmental features so that the expenses can be reduced as well.

Certain limitations exist in applying the developed ARBF model. For example, it may not be appropriate for rural areas due to the differences in topographical features, which can affect the air movement. The model only utilizes the correlation between  $\text{O}_3$  and three vehicle emission variables and four meteorological variables, while excludes the impact of other emissions on  $\text{O}_3$ . Although the model can produce satisfactory results, other precursors of ozone formation, e.g., nonmethane hydrocarbon, aldehydes, and reactive gases, and additional meteorological variables should also be considered corresponding to different microenvironments.

Despite of the limits existing in air quality forecasting models, the developed ARBF model is capable to predict the daily maximum ozone level successfully and is worthy of further exploration.

## Acknowledgements

The work described in this paper was partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region (HKSAR), China [Project No. CityU 1013/02E] and a Strategic Research Grant #7001244(BC) from City University of Hong Kong, HKSAR. The provision of original data from HKEPD is also appreciated.

## References

- Boznar M, Lesjak M, Mlakar P. A neural network-based method for short-term predictions of ambient  $\text{SO}_2$  concentrations in highly polluted industrial areas of complex terrain. *Atmos Environ B* 1993; 27(2):221–30.
- Broomhead D, Lowe D. Multivariable functional interpolation and adaptive networks. *Complex Syst* 1988;2:321–55.
- Chang TY, Suzio MJ. Assessing ozone–precursor relationships based on a smog production model and ambient data. *J Air Waste Manage Assoc* 1995;45:20–8.
- Chen JL, Shafiqul I, Pratim B. Nonlinear dynamics of hourly ozone concentrations nonparametric short term prediction. *Atmos Environ* 1998; 32(11):1839–48.
- Comrie AC. Comparing neural networks and regression models for ozone forecasting. *J Air Waste Manage* 1997;47:653–63.
- Fan HY, Lu WZ, Xu ZB. An empirical comparison of three novel genetic algorithms. *Eng Comput* 2000;17(8):981–1001.
- Folinsbee LJ, Hortman DH, Kehrt HR. Respiratory responses to repeated prolonged exposure to 0.12 ppm ozone. *Am J Respir Crit Care Med* 1994;149:98–105.
- Gardner MW, Dorling SR. Neural network modelling of the influence of local meteorology on surface ozone concentrations. *Proceedings 1st*

- International Conference on GeoComputation, University of Leeds; 1996. p. 359–70.
- Gardner MW, Dorling SR. Artificial neural networks (the multilayer feed-forward neural networks)—a review of applications in the atmospheric science. *Atmos Environ* University of Leeds Press, U.K. 1998;30(14/15):2627–36.
- Hadjiiski L, Hopke PK. Application of artificial neural network to modeling and prediction of ambient ozone concentrations. *J Air Waste Manage Assoc* 2000;50:894–901.
- Hong Kong Environmental Protection Department Annual Report. Environmental Protect Department, 1996.
- Hong Kong Environmental Protection Department Annual Report. Environmental Protect Department, 1997, 1998.
- Hong Kong Environmental Protection Department Annual Report. Environmental Protect Department, 1999.
- Hong Kong Environmental Protection Department Annual Report. Environmental Protect Department, 2000.
- Larsen RI, McDonnell WF, Horstman DH. An air quality data analysis system for interrelating effects standards, and needed source reductions: Part 11. A lognormal model relating human lung function decrease to ozone exposure. *J Waste Manage Assoc* 1991;41:455–9.
- Lippmann M, Liog PJ, Leikuaf G, Green KB, Baxter D, Morandi M, et al. Effects of ozone on the pulmonary function of children. *Adv Mod Environ Toxicol* 1983;5:423–46.
- Lu WZ, Fan HY, Leung AYT, Wong JCK. Analysis of pollutant levels in central Hong Kong applying neural network method with particle swarm optimization. *Environ Monit Assess* 2002a;79:217–30.
- Lu WZ, Wang XK, Wang WJ, Leung AYT, Yuen KK. A preliminary study of ozone trend and its impact on environment in Hong Kong. *Environ Int* 2002b;28(6):503–12.
- Lu WZ, Wang WJ, Fan HY, Leung AYT, Xu ZB, Wong JCK. Prediction of pollutant levels in Causeway Bay area in Hong Kong using improved neural network model. *ASCE J Environ Eng* 2002c;128(12):1146–57.
- Lu, WZ, Fan, HY, Lo, SM, in press. Application of evolutionary neural network method in predicting pollutant levels in downtown area of Hong Kong. *Neurocomputing* 50(1).
- Rao ST, Zurbenko IG. Detecting and tracking changes in air quality. *J Air Waste Manage Assoc* 1994;44:1089–92.
- Reich SL, Gomez DR, Dawidowski LE. Artificial neural network for the identification of unknown air pollution sources. *Atmos Environ* 1999;33:3045–52.
- Roadknight CM, Balls GR, Mills GE, Palmer\_Brown D. Modeling complex environmental data. *IEEE Trans Neural Netw* 1997;8(4):852–61.
- Robeson SM, Steyn DG. Evaluation and comparison of statistical forecast models for daily maximum ozone concentrations. *Atmos Environ* 1990;24B:303–12.
- Seinfeld JH. Ozone air quality models: a critical review. *J Atmos Pollut Control Assoc* 1988;38:616–47.
- Song XH, Hopke PK. Solving the chemical mass balance problem using an artificial neural network. *Environ Sci Technol* 1996;30(2):531–5.
- Spector DM, Hurston GD, Mao J, He D, Hayes C, Lippmann M. Effects of single and multiday ozone exposures on respiratory function in active normal children. *Environ Res* 1991;55:107–22.
- Wang WJ. A high efficiency algorithm of input driven BP networks. *Syst Sci Syst Eng* 2000;20:99–101 [in Chinese].
- Wang WJ, Tang XC, Li WC. A variable structure neural network model and its applications. *IEEE TENCON* 1993;2(5):799–802.
- Yi J, Prybutok V.R. A neural network model forecasting for prediction of daily maximum ozone concentration in an industrialized urban area. *Environ Pollut* 1996;92(3):349–57.