Transactions on Ecology and the Environment vol 21, © 1998 WIT Press, www.witpress.com, ISSN 1743-3541

Adaptive fuzzy system for prediction of SO_2 concentration in Macau

S. C. Tam[1], K. M. Mok[2], C. Sam[3]
Faculty of Science and Technology, University of Macau, PO Box 3001, Macau
Email: [1]fstsct@umac.mo, [2]fstkmm@umac.mo, [3]fstsc@umac.mo

Abstract

Due to the small size and the complexity of the effects from the neighboring metropolises, traditional modeling such as the wind flow and dispersion modeling of the air pollution problem in Macau may not be economical or feasible. Investigation on the applicability of an adaptive neuro-fuzzy inference system (ANFIS) for predicting the daily average sulfur dioxide concentration in Macau is carried out. The 5-day ahead SO_2 concentration values predicted by the model are within 14% of the measured values which indicates that the ANFIS could be used to develop efficient air-quality prediction models.

1 Introduction

Macau, a 21.4 km^2 territory, sits at the west side of the Pearl River Delta with Hong Kong and Guangzhou situated at 60 km to its East-NorthEast and 105 km to its North, respectively. With such a geographical location, Macau is experiencing rapid economic growth as its neighbor cities (e.g. Hong Kong, Shen Zhen, and Guangzhou) in the Delta area in recent years. The adverse consequence for the rapid economic growth is the increase of SO_2 emission. Therefore, prediction of SO_2 concentration in the ambient air is increasingly gaining importance as there is growing concern about the effects of it on the Macau's ecosystems.

Macau comprises a Macau peninsula (7.8 km^2) and two islands of Taipa (5.8 km^2) and Coloane (7.8 km^2) as shown on figure 1. The total population in 1996 is 414,128 with 94.4% of them residing on the Macau peninsula[2]. Hence the comprehension of the air-concentration condition in the Macau peninsula is of high priority due to its direct effects on the major inhabitants of Macau.

To model the SO_2 condition in Macau, the usual approach is to develop mathematical meteorological and dispersion models which can describe the relationship between pollutant emission, transmission and ambient air concentrations of the interested air pollutant as a function of space and time in a mathematical way. The calculation quality of the said models depends on precise information on the pollutant amounts emitted, wind field and the individual turbulence condition of the atmosphere. However, there is currently no systemic study on estimating the SO_2 emission inventory in Macau. Therefore, the essential SO_2 emission data in Macau required by the traditional modeling is not well defined. Furthermore, the air quality of Macau could be influenced by the conditions of its neighbors, especially China. Studies indicate that the total projected SO_2 emission in Asia in the year of 2010 is 79 million tones, of which more than 60% is contributed by China[3]. Macau is geographically connected to Guangdong, which is one of the Chinese provinces with the fastest economic growth. The large amount of SO_2 emitted in China could eventually contribute in adverse ways to the air quality of Macau through long range transport processes. If the models developed in tend to include these foreign effects from the closer emission sources in the southern China, the spatial scales of such simulation models could range from a few kilometers to tens of kilometers. Then the small size Macau would appear in the simulated area more of less like a point. With such concern as well as the ill defined SO_2 emission inventory of Macau, development of the traditional dispersion model may not be economical (if only the situation in Macau is concerned) or feasible at the present time. In this study an alternate approach, an adaptive neuro-fuzzy inference system, is tried for the prediction of future SO_2 concentration variation at a representing site in Macau with minimum input information; i.e. only the history of the SO_2 concentration.

Transactions on Ecology and the Environment vol 21, © 1998 WIT Press, www.witpress.com, ISSN 1743-3541

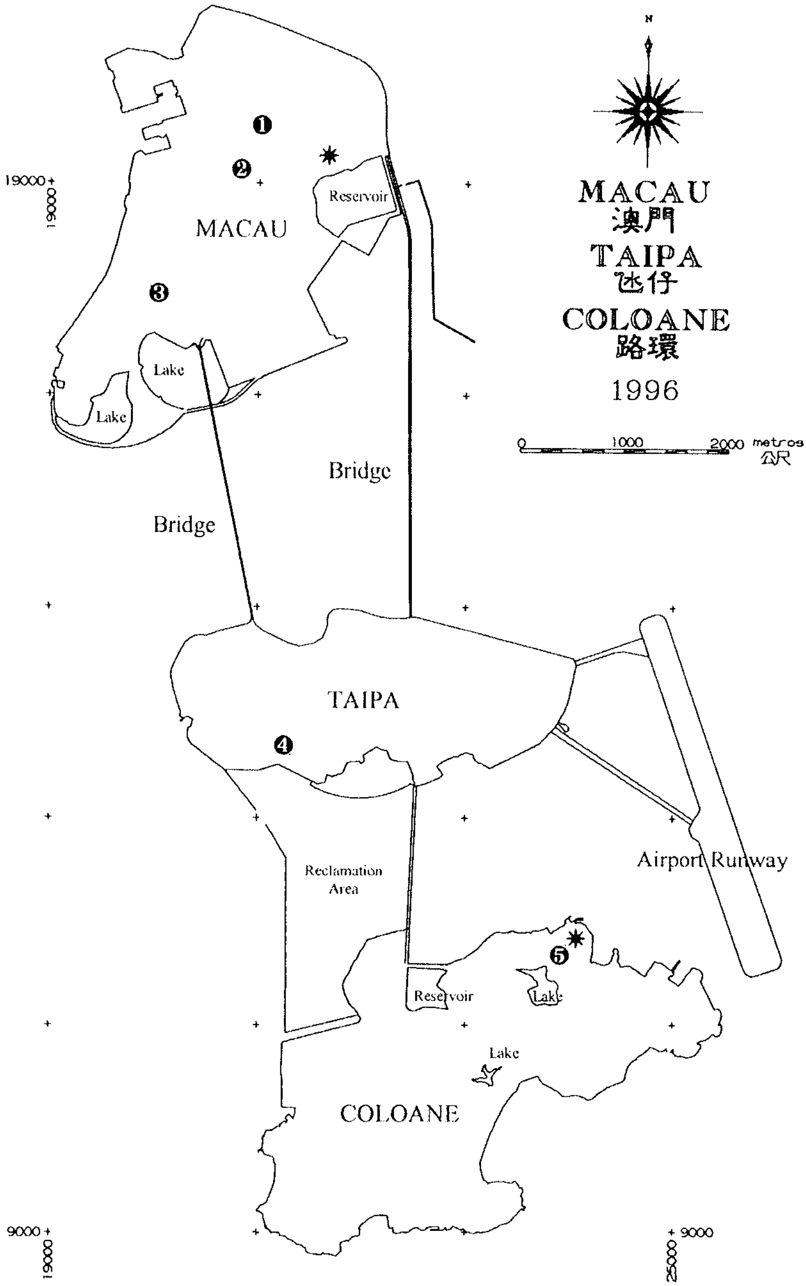

Figure 1: Map of Macau [1]; ❶ A. Preta, ❷ H. Costa, ❸ Fortaleza, ❹ CMI, ❺ Ka-Hó are the air quality monitoring stations; ✷ represents the electric power plants.

2 Adaptive Neuro-fuzzy Inference System

Assuming that the time series x_t of SO_2 concentration is stationary and by the theorem of Taken, the future value of the time series can be predicted. Precisely, there exists a function F such that

$$x_t = F(x_{t-\tau_1}, x_{t-\tau_2}, \ldots, x_{t-\tau_d}) \tag{1}$$

where τ_k is a none negative increasing finite sequence and d is called the embedding dimension. It is noted that the functional form of F is unknown but it is usually a nonlinear function. In this paper, an estimator called adaptive neuro-fuzzy inference system (ANFIS) is used to approximate F. The ANFIS was introduced by Jang and Sun [4] and it is a fuzzy inference system with learning capability.

2.1 The architecture

The fuzzy system under consideration in ANFIS is the first order Sugeno type fuzzy model[5]. A common rule set with two fuzzy if-then rules is the following:

1. Rule: If x is A_1 and y is B_1, then $f_1=p_1 x+ q_1 y +r_1$,
2. Rule: If x is A_2 and y is B_2, then $f_2=p_2 x+ q_2 y +r_2$.

Figure 2 illustrates the reasoning mechanism for this Sugeno model. The Sugeno Model could also be represented in a different but equivalent way as shown in figure 3. The architecture of it is summarized below:

Layer 1 Every node i in this layer is a square node with a node function

$$O_i^1 = \mu_{A_i}(x), \tag{2}$$

where $\mu_{A_i}(x)$ is the membership function of the fuzzy concept A_i. The usual choice of the membership function is of bell-shaped with maximum equal to 1 and minimum equal to 0, such as the generalized bell function

$$\mu_{A_i}(x) = \frac{1}{1+[(\frac{x-c_i}{a})^2]^{b_i}} \tag{3}$$

or the Gaussian function

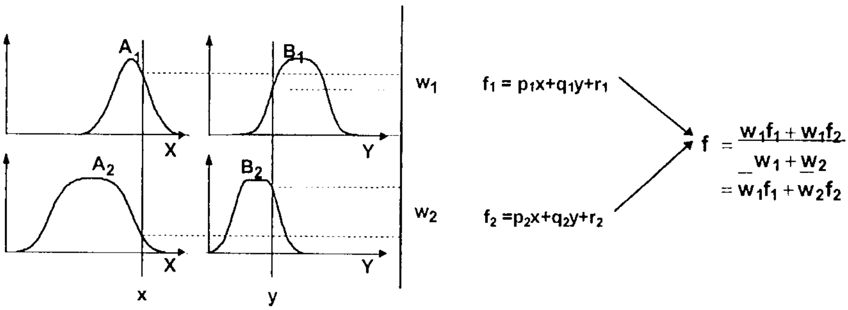

Figure 2: Sugeno type fuzzy reasoning.

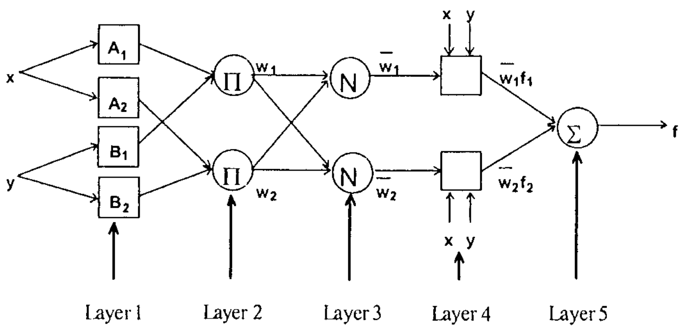

Figure 3: Corresponding equivalent ANFIS architecture.

$$\mu_{A_i}(x) = \exp[-(\frac{x-c_i}{a})^2] \tag{4}$$

where $\{a_i, b_i, c_i\}$ (or $\{a_i, c_i\}$ in the latter case) is called the premise parameter set. The generalized bell function type (3) is chosen for the present application.

Layer 2 Every node in this layer is a circle node labeled Π which multiplies the incoming signals and sends the product out. For example,

$$w_i = \mu_{A_i}(x) \times \mu_{B_i}(x), \qquad i = 1, 2. \tag{5}$$

This action corresponds to the fuzzy operation "and" in the fuzzy system. Actually, other operators that perform generalized "and" can be used. (e.g. min($\mu_{A_i}(x)$, $\mu_{B_i}(x)$).)

Layer 3 Every node in this layer is a circle node labeled *N*. The *i*-th node calculates the ratio of the *i*-th rule's firing strength to the sum of all rules' firing strengths:

$$\overline{w_i} = \frac{w_i}{w_1 + w_2}, \qquad i = 1, 2. \tag{6}$$

Layer 4 Every node i in this layer is a square node with a node function

$$O_1^4 = \overline{w}_i f_i = \overline{w}_i (p_i x + q_i y + r_i)\,, \tag{7}$$

where $\overline{w}_i$ is the output of layer 3 and $\{p_i,\ q_i,\ r_i\}$ is called the consequent parameter set.

Layer 5 The single node in this layer is a circle node labeled Σ. It computes the overall output as the summation of all incoming signals, i.e.

$$O_1^5 = \sum_i \overline{w}_i f_i\,. \tag{8}$$

Therefore, in this case,

$$f = (\overline{w}_1 x)p_1 + (\overline{w}_1 y)q_1 + (\overline{w}_1)r_1 + (\overline{w}_2 x)p_2 + (\overline{w}_2 y)q_2 + (\overline{w}_2)r_2 \tag{9}$$

2.2 The hybrid learning method

The learning method of ANFIS is a hybrid learning algorithm which consist of two passes, namely the forward pass and the backward pass. In the forward pass (the first pass), the premise parameters will be fixed and the consequent parameters are identified by the least square estimate. In the backward pass (the second pass), the consequent parameters will be fixed and the premise parameters will be updated by the gradient descent which is a supervised learning.

2.3 The Approximation Capabilities

It was showed by the Stone-Weiestrass Theorem that any continuous function defined on a closed and bounded subset of $\boldsymbol{R}^n$ can be approximated by an ANFIS. Therefore, ANFIS can be considered as an universal estimator for functions. Also ANFIS has been tested in [6] for predicting chaotic times series generated by the chaotic Mackey-Glass differential delay equation and it has been used to predict the finance time series. Considering the Chaotic property of the SO_2 time series due to the influences of the local meteorological and topographic conditions, ANFIS would be an ideal tool for this application.

3 The SO_2 Data

The *Serviços Meteorológicos e Geofísicos de Macau* is monitoring the SO_2 concentration in the ambient air of Macau at five locations which are A. Preta, H. Costa, Fortaleza, Camara Municipal das Ilhas (CMI) and Ka-Hó (figure 1). Note that three of the five monitoring stations are located at the Macau peninsula due to the presence of the major population and the concern of their well being there. Adopting from the results of Mok & Tam[7], who found that A. Preta is the area that has the highest potential of being polluted, it is identified as the representing site for the investigation of sulfur dioxide pollution in Macau. The present study uses the daily average concentration recorded at this site during the last three months (October to December) of the year 1995 for analysis. The thick line in figure 4 displays the evolution of the measured SO_2 concentration during this period. There are 36 days (39% of the period) with SO_2 concentration exceeding the 50 $\mu g/m^3$ Chinese primary standard. The high frequency of polluted days in A. Preta is probably due to the large amount of SO_2 emission caused by the nearby electric power plant and the concentrated industrial, commercial and residential activities in the areas.

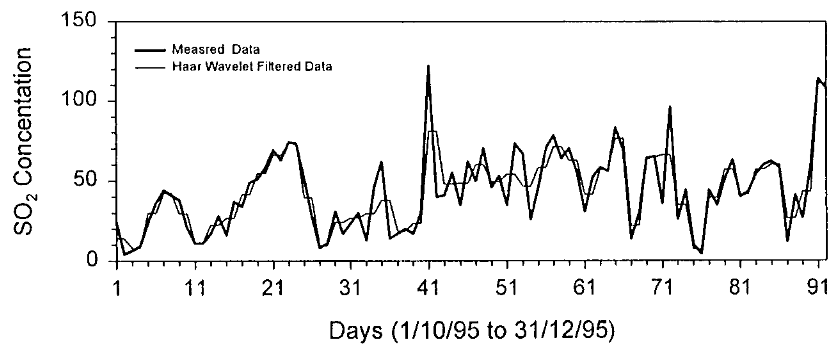

Figure 4: Measured and filtered daily average SO_2 concentration at A. Preta of Macau during the period of October to December of 1995; the units for SO_2 concentration is in $\mu g/m^3$.

4 The ANFIS results and discussion

In order to model the general temporal evolution trend of the SO_2 concentration, the recorded data is pre-processed by feeding it through a Haar wavelet filter to remove the influence of noises due to probable local emissions. The thin line in figure 4 shows that the filtered profile can capture the main trend of the original data with the removal of some relatively large fluctuations caused by local emissions which do not represent the general situation at that location. The subsequent analysis will utilize the de-noised profile.

It is known that the evolution of ambient air quality could be influenced by local weather conditions. The temporal evolution of SO_2 concentration is expected to be chaotic as the changing meteorological factors. Prediction of the future daily mean SO_2 concentration based on the knowledge of long history record may not be meaningful and computationally inefficient. In the present study, it is assumed that the future SO_2 concentration is only related to the evolution patterns occurred in the past thirty days. Based on this assumption and the size of the selected SO_2 data, two separate tests are set to examine the prediction ability of the ANFIS. The first test is to use the data of October to train the networks to predict the first five-day SO_2 concentrations of November (Case I), and the second test uses the November data to predict the first five-day values of December (Case II).

For the five-day ahead prediction of the SO_2 concentration, the ANFIS model assumes the relation of (1) with d = 3 and a time lag τ_d of 4 days; i.e. to predict the SO_2 concentration at day t, the daily averaged values at days t - 4, t - 8, and t - 12 are used. This arrangement is based on a trial and error basis and it is suitable for both Case I and II. The present ANFIS utilizes eight fuzzy rules and the training data set for both case I and II consisted 19 and 18 samples of the daily average SO_2 concentration measured in October and November, respectively. The software used to develop the present ANFIS model is the Fuzzy Logic Toolbox of MATLAB[8].

The learning and predicting capability of the ANFIS model for case I are shown in figure 5. The learning process shows that the values calculated by the model fit the measurements well and the differences fall within - 4.40×10^{-4} % to 1.52×10^{-4} % of the measured values. The five days ahead values are slightly over-predicted by the ANFIS model with a maximum difference of 13.9 %. For case II, figure 6 shows that the model learns well with the differences falling within - 2.40×10^{-4} % to

1.60×10^{-4} % of the measured values. The model under-predicts the five days ahead values with a maximum difference of 11.2 %. The promising results indicate that the ANFIS could be used to develop efficient air-quality prediction models in the future.

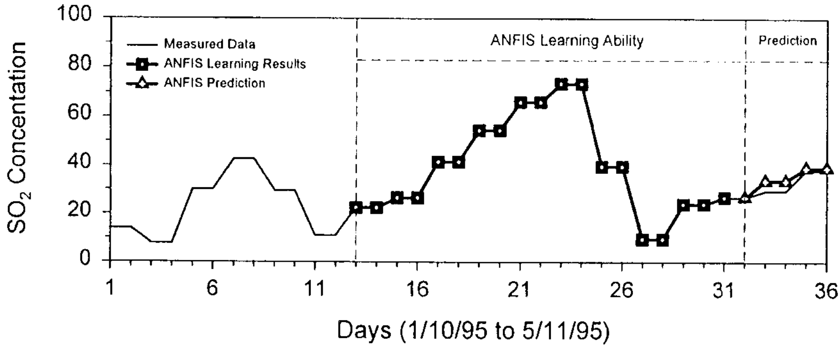

Figure 5: ANFIS calculated and measured daily average SO_2 concentration at A. Preta of Macau during the period of October 1st to November 5th of 1995; the units for SO_2 concentration is in $\mu g/m^3$.

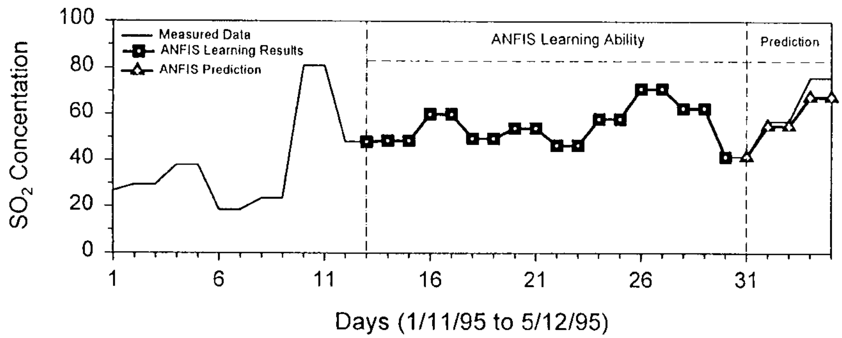

Figure 6: ANFIS calculated and measured daily average SO_2 concentration at A. Preta of Macau during the period of November 1st to December 5th of 1995; the units for SO_2 concentration is in $\mu g/m^3$.

Acknowledgments

This study is funded by the Research Committee of the University of Macau (RG002/97-98S/MKM/FST). The *Serviços Meteorológicos e Geofísicos de Macau* is thanked for supplying the SO_2 data and the *Direcção dos Serviços de Cartografia e Cadastro de Macau* is thanked for granting permission for the usage of the map of Macau.

References

[1] DSCC, *Digital Base Map of Macau's Territory (CD-ROM),* Direcção dos Serviços de Cartografia e Cadastro de Macau, 1996.

[2] DSEC, *Anuário Estatístico,* Serviços de Estatística e Censos de Macau, 1996.

[3] Arndt, R., Xu, Y., Carmichael, G.R., Sunwoo, Y. & Zhang, Y., Long Range Transport of Sox in Asia, *Air Pollution II,* Vol. 1, Computational Mechanics Publications, Southampton and Boston, pp. 245-250, 1994.

[4] Jang, J-S. R. & Sun, C.-T., Neuro-fuzzy Modeling and Control, *Proceedings of the IEEE*, pp. 378-406, 1995.

[5] Takagi, T. & Sugeno, M., Derivation of Fuzzy Control Rules from Human Operator's Control Actions, *Proc. of the IFAC Symp. on Fuzzy Information, Knowledge Representation and Decision Analysis*, pp. 55-60, 1983.

[6] Jang , J.-S. R., Neuro-Fuzzy Modeling: Architectures, Analyses, and Applications, Ph.D. Dissertation, EECS Department, Univ. of California at Berkeley, 1992.

[7] Mok, K.M. & Tam, S.C., Short-Term Prediction of SO_2 Concentration in Macau with Artificial Neural Networks, *Energy and Buildings*, (Accepted for printing), 1998.

[8] Jang, J.-S. R., & Gulley, N., *Fuzzy Logic Toolbox for Use with MATLAB*, The Math-Works, Inc., Natick, MA, 1995.