A preliminary study of respirable suspended particulate level in taxi transport interchanges in Hong Kong

Lu Wei-Zhen[a,*], Wang Xie-Kang[b]

[a]*Department of Building and Construction, City University of Hong Kong, Kowloon, Hong Kong;* [b]*State Key Laboratory of Hydraulics of High Speed Flows, Sichuan University, Chengdu 610065, People's Republic of China*

ABSTRACT
A study on pollutant dispersion and distribution inside public taxi transfer interchanges (TTIs) is reported. The pollutant levels inside TTIs are affected by many factors, for example, taxi data, climatic conditions, human activities and geometrical layout of TTIs. A site measurement of respirable suspended particulate (RSP) level is carried out in a typical TTI in Hong Kong. After analysing the effect of the above factors on RSP level, we propose to use artificial neural networks (ANNs) to study such phenomena. The recorded data within different time periods inside the selected TTIs are used as the test data set to train the proposed neural network model. The recovery performance of the ANN model is analysed and justified. In this study, we compare the forecasting results and the measured data of RSP in morning and afternoon sessions, respectively. The results show the feasibility and reliability of the proposed approach for forecasting pollutant levels inside TTIs.

INDEX TERMS
Artificial neural networks; Taxi transport interchanges; Hong Kong

INTRODUCTION
As there is continuous economic development and population increase in Hong Kong, a series of severe problems relating to environmental benefits and protection has attracted much attention than ever before, for example, air pollution, shortage of land resources, waste and sewage disposal, etc. (Qin *et al.*, 1997; Chan *et al.*, 1999; Environmental Protection Department Report, 2000; Chan and Kwok, 2001). Among these, air pollution has a direct effect on human health through exposure to high ambient concentration levels of pollutants. The respirable suspended particulate (RSP) levels have frequently exceeded the air quality objective (AQO) of Hong Kong during the past 6 years (Environmental Protection Department Report, 2000). The main sources of air pollution in Hong Kong come from industry and vehicles. In recent years, with better land-use planning, stricter controls on certain industrial processes, and reduction in sulphur content and viscosity of industrial fuel oil, air pollution from industrial operations has been put under effective control. Nevertheless, as a result of increasing number of vehicles, vehicular emissions have become the main source of air pollution. Unfortunately, in Hong Kong, almost all commercial vehicles such as taxis, buses and goods vehicles run on diesel and therefore emit high levels of RSP. The situation is particularly serious in public transfer locations, for example, public taxi transfer interchanges (TTIs). These locations are normally built at ground level under large building complexes, surrounded by solid walls and structural columns, and poorly ventilated. As a result, the exhaust gas from vehicles could be trapped inside the TTI and the air inside these locations is smoky, filthy, choking and harmful to passengers.

Concerning the high demand of good air quality in public places, a study on pollutant dispersion and distribution inside public TTIs is required and is necessary for the ultimate

*Corresponding author. E-mail: bcwzlu@cityu.edu.hk

purpose of diluting pollutant levels and improving air quality inside these locations. The concentration levels of pollutants inside TTIs are affected by many factors, for example, the taxi flow-rate, the continuous emission rates and periods of taxies, the ventilation rate, the environmental conditions (temperature, humidity, etc.), the geometrical layout inside TTIs, and the opening size and location of TTIs. The relationships between pollutant levels and their affected factors are basically non-linear, highly coupled or even not known yet. Full-scale measurement and computational fluid dynamics (CFD) simulation, for example, would be very expensive and time-consuming by using a conventional engineering approach. Artificial neural network (ANN) is good at dealing with such problems (Karayiannis and Venetsanopoulos, 1993; Perez and Reyes, 2001). This paper aims to analyse the factors affecting RSP level and the variable characters of RSP in TTIs based on the collected historical data in typical TTI sites, to investigate the probability and its effectiveness using an ANN approach and to predict RSP time series tendency in selected TTIs in Hong Kong.

MEASUREMENT AND ANALYSIS OF FIELD DATA

Field Data Acquisition

In order to study the RSP levels in TTIs, a site measurement of RSP level was carried out in a typical TTI in Hong Kong (see Figure 1). The RSP level and its affected factors were monitored using relevant equipment. The measured data include air velocity (Anemomaster Model 6112 Kanomax) at a location 1.5 m above the ground within the passenger waiting area, temperature/humidity (No. 8016-00 Model SK-L200TH Sato Keiryoki Mfg. Co., Ltd.), and RSP levels (Airborne Particulates Monitor Model No. PDR-1000AN) at the same location within 10 s intervals. All parameters are measured within 1 h duration during off-peak hour at 09:30–10:30 a.m. and peak hour at 13:30–14:30 p.m. for a whole week. In this study, the taxi flow-rate is counted in queue every minute and then the average taken for 5-min intervals (see Figure 2). The air velocity data are listed in Table 1. The parameters of temperature, humidity and RSP level are collected within the interval of every 10 s simultaneously. Figure 3 shows the variations of temperature and humidity during the collected hours on Wednesday, which represents typical tendency of these parameters during a week.

Table 1 Average air velocity at different times during a week

Time	Air Velocity (m/s)						
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
09:30	0.01	0.01	0	0.02	0	0.02	0.03
10:30	0	0.01	0.03	0	0.02	0.01	0.02
13:30	0.04	0	0	0.02	0.2	0.04	0
14:30	0.04	0.03	0	0.02	0.02	0.03	0.02

Remark The ventilation system operates between 10:00 a.m. and 24:00 p.m.

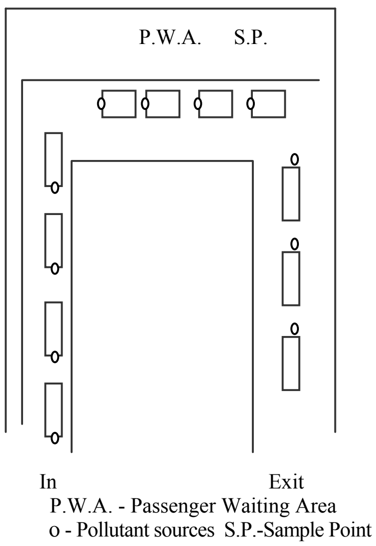


Figure 1 Scheme of a simplified TTI.

Data Analysis

Before establishing the RSP forecasting model, it is necessary to analyse the characteristics of RSP and the properties of its affected factors based on the original measured data. For the selected time durations in the study, the differences of the RSP levels between the morning and afternoon sessions, in particular, on Tuesday and Wednesday, can be observed (see Figure 4). However, the changes of some affected factors are not obvious, for example, the temperature and air velocity in the same time periods, shown in Table 1 and Figure 3. In general, it is believed that the taxi flow-rate is the main affecting factor because the RSP originates from its exhaust emission. From Figure 4, it is also of interest to observe that the maximum RSP levels are not in par with the maximum taxi occupancy during off-peak or peak hours as shown in Figure 2. Hence, other parameters may have greater influence on the RSP levels than the taxi flow-rate.

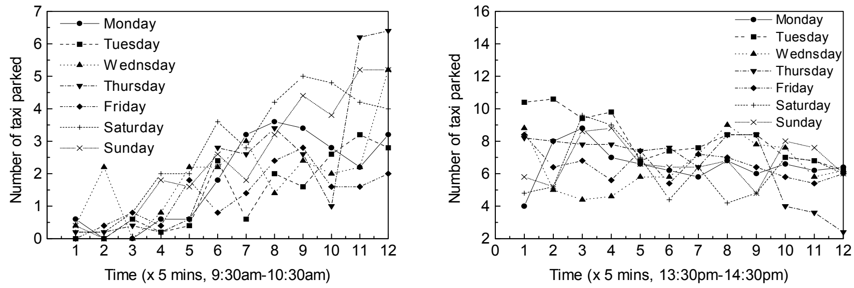


Figure 2 Taxi flow rate during off-peak and peak hours for a week in a typical TTI.

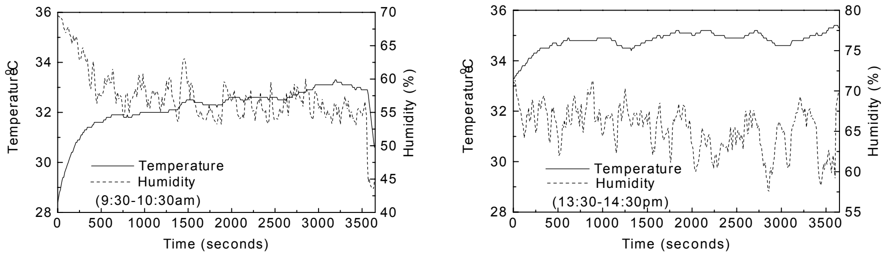


Figure 3 Profiles of temperature and humidity on Wednesday in a typical TTI.

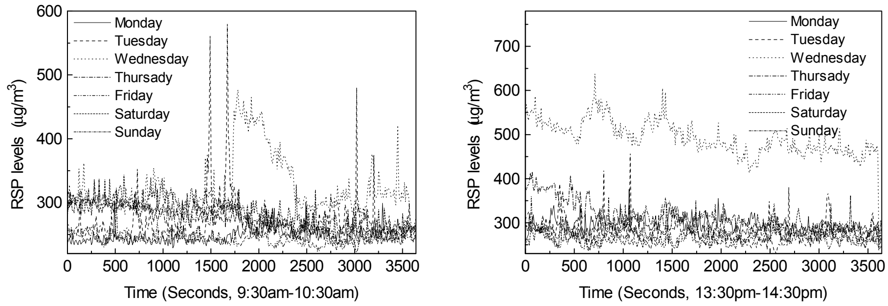


Figure 4 RSP profiles during off-peak and peak hours for a week.

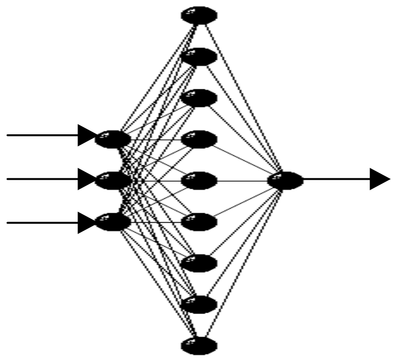
According to the humidity curves during the measured periods shown in Figure 3, it is reasonable to take the humidity as a major factor to influence the RSP levels (Neubauer *et al.*, 1998). In the mean time, we take an additive factor produced by stochastic function to produce a complex vector instead of other factors such as temperature, and short sequence, for example, taxi flow-rate and air velocity, etc. Considering the time differences between taxi emission and pollutant monitor, some delay time series must be considered in the forecasting model. A function model for predicting RSP level is founded as below:

$$\text{RSP}(t) = f(\text{Add}, \text{Humidity}(t), \cdots \text{Humidity}(t-\tau 1); \text{RSP}(t-1), \ldots, \text{RSP}(t-\tau 2)) \quad (1)$$

Here Add, Humidity() and RSP() represent additive factor, humidity and RSP time series, while, τ_1 and τ_2 represent the delay time, respectively.

APPLICATION OF ARTIFICIAL NEURAL NETWORK MODEL

The training process of a three-layer, feed-forward neural network, which will be used to predict the RSP variations in public TTIs, is presented in this section. Figure 5 presents the typical topology structure of the network. It consists of an input layer, a single hidden layer and an output layer, every neuron or node in the topology represents a non-linear mapping between input and output vectors (Karayiannis and Venetsanopoulos, 1993). It is denoted that the input vector is $X(x_1, x_2, \ldots, x_n)$, and the output vector is $Y(y_1, y_2, \ldots, y_m)$. The s denotes number of hidden nodes. The $w_{ij}, w_{jk} (w_{ij} \in R^{n \times s}, w_{jk} \in R^{s \times m})$ are, respectively, the connection weighting matrix between input and hidden layer, and that between hidden and output layer. The $\theta_j, \theta_k (\theta \in R^s, \theta_k \in R^m)$ are threshold vectors of the hidden layer and the output layer separately. Therefore, these interrelationships form a non-linear map from the input space R^n to the output space R^m.



Input layer Hidden layer Output layer

Figure 5 Topology of three layer feed-forward neural networks.

Here, the output vector of the hidden layer is:

$$O_p = f(w_{ij} \times x_p + \theta_j) \quad (2)$$

The output vector of the output layer is:

$$y_p = f(w_{jk} \times O_p + \theta_k) \quad (3)$$

In this study, function $f(\cdot)$ is a bounded, non-decreasing, non-linear function, for example, the sigmoid function, expressed as:

$$f(\mu) = \frac{1}{1+\mathrm{e}^{-\mu}} \quad (4)$$

The network is trained by using a gradient-descent technique (Perez and Reyes, 2001). The root mean square (RMS) error is defined as below:

$$E = \sqrt{\frac{1}{N}\sum_{p=1}^{N}(y_p - d_p)^2} = \min(E) \quad (5)$$

Here, N is the number of the training samples and d_p is the desired output. The normal back-propagation algorithm is used to train the neural network by modifying the learning parameters w_{ik} and θ_{jk}. Once the training is completed successfully, the RMS achieves the minimum value for all training samples and the objective of the investigation can be matched as:

$$\Delta w_{ik}(t) = -\eta \frac{\partial E(t-1)}{\partial w_{ij}(t-1)} + \alpha \Delta w_{ik}(t-1) \quad (6)$$

In Eqn (6), $\Delta w_{ik}(t) = w_{ik}(t) - w_{ik}(t-1)$; $\eta \in (0,1)$ is the learning rate, and $\alpha \in (0,1)$ is the momentum factor. The training requires a set of training data, which normally consists of a series of input and associated output vectors according to supervised rules. During the training, the model will continuously adjust the weights and threshold in the neural network using certain kinds of training algorithms so that the desired input–output mapping could be obtained. After the training is completed, the multi-layer feed-forward neural network has the ability to forecast unknown variables based on an input vector similar to the input samples in training.

FORECASTING OF THE RSP CONCENTRATIONS

A multi-layer feed-forward neural network has the ability to model an unknown non-linear function and to predict its tendency through training. In this study, an implicit function of RSP, that is Eqn (1), has been proposed. According to the RSP distributions shown in Figure 4, the RSP levels on Tuesday and Wednesday are taken as samples, all measured data are divided into two sets, that is, the training set (200 samples) and the testing set (150 samples). In order to be compatible with the multi-layer feed-forward neural network, all data are normalized into the range [0,1] by sigmoid function. The normalization is carried out according to Eqn (7), and two parameters, α and $\beta \in (0,1)$, are adopted in order to prevent the nodes becoming over-fitting, in general, $\beta = 0.9$, $\alpha = (1-\beta)/2$:

$$x_{\text{norm}} = \frac{(x - x_{\min})}{x_{\max} - x_{\min}} \beta + \alpha \quad (7)$$

Thus, first of all, we need to consider the values of those delaying parameters and the number of hidden nodes in the forecasting model. Having analysed the distributions of measured variables and compared the training precision by increasing or decreasing input nodes during the training, the values of τ_1 and τ_2 are both assigned as 4. The number of hidden nodes is selected as 18 according to the general rule of input number plus ten. In case studies, the proposed multi-layer feed-forward neural network has been structured with the form of '9-19-1'. Figures 6 and 7 present a comparison between simulation and measurement on Tuesday and Wednesday by the above established neural network model. The recovery performance of the ANN model is satisfactory and reliable.

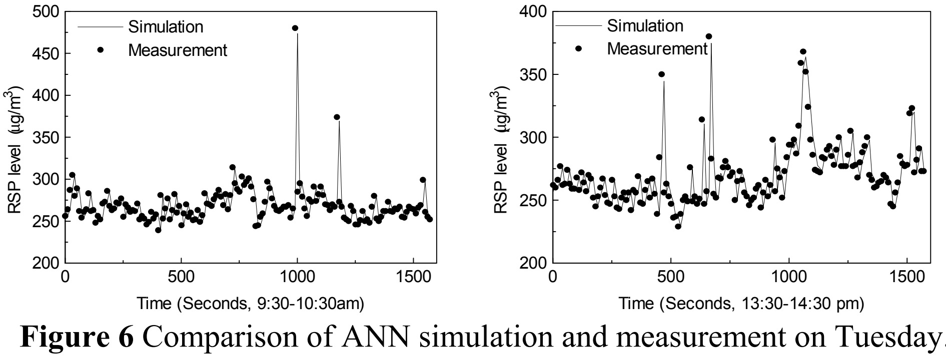


Figure 6 Comparison of ANN simulation and measurement on Tuesday.

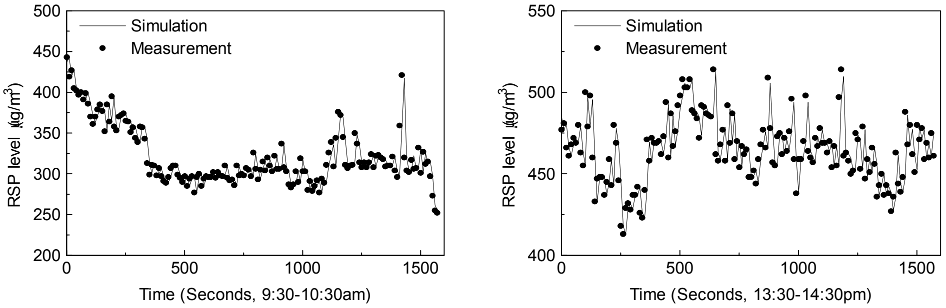


Figure 7 Comparison of ANN simulation and measurement on Wednesday.

CONCLUSIONS

In this study, an ANN model has been established to predict the RSP level in public TTIs. In the ANN model, an additive complex factor is produced by the stochastic function and used as input vector instead of conventional parameters. It is found that the climatic condition such as humidity has great impact on the RSP levels in TTI, while temperature is less important. Although the air movement affects the characteristics of RSP, it is feasible to ignore it because of a small change in magnitude of air velocity. Concerning the effect of taxi flow-rate on RSP level, it is necessary to get more field data in different TTIs in future. It can be concluded that the established ANN approach for forecasting pollutant levels inside TTI is feasible, effective and validated by applying to a practical case of simulating the RSP levels in a typical TTI in Hong Kong.

ACKNOWLEDGEMENT

The authors acknowledge the financial support of the Strategic Research Grants #7001463(BC) and 7001371(BC), City University of Hong Kong, HKSAR.

REFERENCES

Chan, L.Y. and Kwok, W.S. (2001). Roadside suspended particulate at heavily trafficked urban sites of Hong Kong—seasonal variation and dependence on meteorological conditions. *Atmospheric Environment* **35**, 3177–3182.

Chan, L.Y., Chan, C.Y. and Qin, Y (1999). The effect of commuting microenvironment on commuter exposures to vehicular emission in Hong Kong. *Atmospheric Environment* **33**, 1777–1787.

Environmental Protection Department Report (2000). Air Quality in Hong Kong.

Karayiannis, N.B. and Venetsanopoulos, A.N. (1993). *Artificial Neural Networks: Learning Algorithms, Performance Evaluation, and Applications*. Boston/Dordrecht/London: Kluwer Academic Publishers.

Neubauer, K.R., Johnston, M.V. and Wexler, A.S. (1998). Humidity effects on the mass spectra of single aerosol particles. *Atmospheric Environment* **32** (14/15), 2521–2529.

Perez, P. and Reyes, J. (2001). Prediction of particulate air pollution using neural techniques. *Neural Computing and Application* **10**, 165–171.

Qin, Y., Chan, C.K. and Chan, L.Y. (1997). Characteristics of chemical compositions of atmospheric aerosols in Hong Kong: spatial and seasonal distributions. *The Science of the Total Environment* **206**, 25–37.