Building and Environment 78 (2014) 111–117

Contents lists available at ScienceDirect

Building and Environment

journal homepage: www.elsevier.com/locate/buildenv

Prediction of particulate matter at street level using artificial neural networks coupling with chaotic particle swarm optimization algorithm

Hong-di He [a], Wei-Zhen Lu [b,*], Yu Xue [c]

[a] Logistics Research Center, Shanghai Maritime University, Shanghai 200135, China
[b] Department of Civil and Architectural Engineering, City University of Hong Kong, Hong Kong
[c] Institute of Physics Science and Engineering, Guangxi University, Nanning 53004, China

ARTICLE INFO

Article history:
Received 13 February 2014
Received in revised form
29 March 2014
Accepted 16 April 2014
Available online 30 April 2014

Keywords:
Particulate matter
Artificial neural networks
Chaotic particle swarm optimization
PM_{10}

ABSTRACT

The time series of particulate matter at urban intersection consists of complex linear and nonlinear patterns and are difficult to forecast. Artificial neural networks (ANNs) have been applied to air quality forecasting in urban areas, but they have limited accuracy owing to their potential convergence to a local minimum and over-fitting. Chaotic particle swarm optimization (CPSO) algorithm is chaos-based searching algorithms and can recognize nonlinear patterns. Hence, a novel hybrid model combining ANN and CPSO algorithm is proposed to improve forecast accuracy. The proposed model, together with the ANN model with the traditional algorithms (Levenberg–Marquardt and PSO), is examined with the measured data in spring and winter respectively. The proposed model is found to provide the best results among them, implying that the hybrid model can be an effective tool to improve the particulate matter forecasting accuracy. Additionally, the proposed model is found to perform better for fine particles than for coarse particles. The model is also verified to predict better in winter than in spring. The outputs of these findings demonstrate the potential of the proposed model to be applied to forecast the trends of air pollution in similar meso-to mega-cities.

1. Introduction

Recently, there appeared many evidences that particulate matter (PM) concentration at urban intersection is mainly originate from traffic volume and is strongly governed by certain meteorological conditions. Obviously, the relationships among these variables are all nonlinear, and as such it is difficult to apply linear models to examine them. For a better understand of them, many approaches, such as computational fluid dynamics model (CFD) [1,2] and statistical model [3], have been proposed to investigate it. Among them, the artificial neural network (ANN) technique, and in particular, the multilayer perception (MLP) model, has been recently developed and verified as a cost-effective approach for analyzing nonlinear environmental problems [4,5].

MLP works by employing a training algorithm to find an optimal or at least a near-optimal weight matrix in the model's weight space, minimizing an error function that normally takes the form of a difference measure between observations and predictions. The optimal weight matrix stores this knowledge, which can then be used to interpret nonlinear relationships. When a new input condition is presented to the MLP model, it generalizes to give a prediction by using the stored knowledge learned from the training data. The predictive performance of the MLP technique is thus mainly determined by the training algorithm used [6,7].

The existing training algorithms for MLP can be generally divided into two categories [8]. The first comprises the local optimization algorithms, which employ an individual-based searching technique. The back-propagation (BP) algorithm and its variants such as gradient descent, conjugated gradient descent and Levenberg–Marquardt (LM) belong to this category. These algorithms have acquired a reputation for potential convergence to a local minimum and over-fitting [9], which means that finding the optimal weight matrix by using these algorithms may not be guaranteed. The second category comprises the stochastic global optimization algorithms, which employ a population-based searching technique. This technique randomly initializes a group of weight matrixes in the weight space, and then parallel performs a searching task by updating a group of weight matrixes iteratively.

* Corresponding author. Tel.: +852 27844316; fax: +852 27887612.
E-mail address: bcwzlu@cityu.edu.hk (W.-Z. Lu).

http://dx.doi.org/10.1016/j.buildenv.2014.04.011
0360-1323/

Algorithms in this category, including genetic algorithms (GA) and particle swarm optimization (PSO), are more efficient and effective in terms of avoiding the local minimum and attempting to find the global minimum, or at least the near global minimum, than the local optimization algorithms [10–12].

Recently, stochastic global optimization algorithms have attracted increasing attention, and in particular PSO due to its straightforward logic, simple realization, and underlying intelligence. Lu et al. [13] employed PSO to train an MLP model to predict the levels of several air pollutants, and suggested that this global training algorithm outperformed the BP algorithm in terms of MLP predictive performance. Chau [14] applied it to train an MLP model for stage prediction in the Shing Mun River, and concluded that the PSO technique was an effective alternative training algorithm for MLP. Wang and Lu [15] proposed a hybrid Monte Carlo (HMC) method to sample the weight matrix from the posterior probability distribution of the MLP optimal matrix, and found that HMC initialization was able to assign the population of weight matrixes in a more promising region of the weight space than could be achieved by random initialization before the PSO started to evolve the MLP model. However, although the traditional PSO algorithm performed well in these studies, it still inevitably suffers from the problem of becoming trapped in local optima because it greatly depends on its parameters [16,17].

To avoid this problem, chaos and chaos-based searching algorithms have been explored in this paper. Chaos optimization algorithms describe the complex behavior of a nonlinear deterministic system and can carry out overall searches at higher speeds than stochastic ergodic searches that depend on probabilities. This paper investigates an improved CPSO algorithm as well as LM and PSO algorithms to train an MLP model to predict PM concentrations with the indentified influencing factors. The aim of this work was: (i) to evaluate the relative influence of precursor concentrations and meteorological variables on PM variation at intersection; and (ii) to predict PM concentrations, through a new methodology based on chaos optimization algorithm and PSO algorithm.

2. Data acquisition

The measurement of PM levels was performed for one and half hours at each time of day over a week each in spring and winter at a selected traffic intersection in Mong Kok along Nathan Road, a major commercial and transport line in Kowloon Peninsula, which is surrounded by high-rise buildings with high traffic volume. The measurement equipment is located 20 cm above the ground on the roadside of the selected location (Fig. 1). The measurement of PM was carried out by the Fluke 983 Particle Counter, which counts the particle number every second by using the laser diffraction technique for size differentiation from submicron to millimeter. In order to compare our measurement with mass concentration, the transformation is performed from number concentration to mass concentration according to the formula between particle diameter and mass [18]. The fine particles (PM_1) and the coarse particles (PM_{10}) are calculated respectively. A digital camera was also dismounted for recording the traffic counts and the traffic signals while a Q-track model 7565 with an air velocity probe 962 was used to detect the wind speed, pitot velocity, pressure, temperature, relative humidity (RH), dew point, wet bulb, and baro pressure.

According to the initial statistical analysis, it is verified that the dispersion of PM at urban traffic intersections is strongly affected by many factors, including traffic conditions and meteorological conditions listed in Table 1. Hence, in order to predict the PM concentrations during each green-light period (PM_{10}_G or PM_1_G) with high precision, eleven variables are selected as input variables for the MLP model, i.e., the PM concentrations in the red-light period (PM_{10}_R or PM_1_R), the green-light period, the red-light period, the diesel vehicle count, the petrol vehicle count, the ratio of diesel vehicles to petrol vehicles (Ratio_DP), the average wind speed, the maximum wind speed, the minimum wind speed, temperature, and relative humidity. The PM_{10}_R and PM_1_R are chosen to represent the background concentrations, while the green-light period, red-light period, diesel vehicle count, petrol vehicle count, and the ratio between these two are selected to describe the traffic characteristics. Furthermore, the wind speed, temperature, and relative humidity are selected to represent the meteorological conditions. It should be noted that, as the traffic volume is calculated for each green-light period, the PM variables selected here are all the average values for the corresponding green-light period.

Fig. 1. Measurement site and equipment setting.

3. Methodology

MLP modeling involves the nonlinear transformation of input data to approximate output data. The model can learn from experimental data examples, and exhibits some ability to generalize beyond the training data. The most common MLP network is the feed-forward MLP model with three layers, that is, the input layer, hidden layer, and output layer. As explained before, eleven variables that represent the background concentrations, traffic conditions, and meteorological conditions were chosen as input neurons, and the variable of PM concentration in the green-light period was chosen as the output (Fig. 2). These settings were inherently determined by the problem, and the number of neurons in the input and output layers remained the same in the whole simulation. The only variation was the numbers of neurons in the hidden layer and the corresponding weights and bias. Thus, the algorithms were mainly used to optimize these parameters in the MLP model to achieve a more accurate prediction.

3.1. Levenberg–Marquardt (LM) algorithm

The LM algorithm is a gradient-based algorithm that includes the first and second derivative of the error term. It adjusts the weights in the direction of steepest descent in which the performance function is decreasing most rapidly. When employed to train MLP models, the advantage of the LM algorithm over the traditional BP algorithm is that it provides a faster (second-order) convergence rate and retains relative stability [19]. During the training, the weight matrix is updated using the following equation: