Atmospheric Environment 104 (2015) 264–272

Contents lists available at ScienceDirect

Atmospheric Environment

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

Fine-scale estimation of carbon monoxide and fine particulate matter concentrations in proximity to a road intersection by using wavelet neural network with genetic algorithm

Zhanyong Wang [a], Feng Lu [b], Hong-di He [c], Qing-Chang Lu [a], Dongsheng Wang [a], Zhong-Ren Peng [a, d, *]

[a] Center for ITS and UAV Applications Research, State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
[b] School of Geographic Science, Nantong University, Nantong 226007, China
[c] Logistics Research Center, Shanghai Maritime University, Shanghai 200135, China
[d] Department of Urban and Regional Planning, University of Florida, PO Box 115706, Gainesville, FL 32611-5706, USA

HIGHLIGHTS

- A hybrid model (GA-WNN) is proposed to predict CO and $PM_{2.5}$ near an intersection.
- GA-WNN model is better in forecasting 5-min series of pollutants than BPNN model.
- GA-WNN model shows good predictions in both peak and off-peak traffic time periods.
- The GA-WNN performance is affected by distance from the road and other local conditions.

ARTICLE INFO

Article history:
Received 5 July 2014
Received in revised form
21 December 2014
Accepted 22 December 2014
Available online 24 December 2014

Keywords:
Carbon monoxide
Fine particulate matter
Fine-scale estimation
Wavelet neural network
Genetic algorithm
Road intersection

ABSTRACT

At road intersections, vehicles frequently stop with idling engines during the red-light period and speed up rapidly in the green-light period, which generates higher velocity fluctuation and thus higher emission rates. Additionally, the frequent changes of wind direction further add the highly variable dispersion of pollutants at the street scale. It is, therefore, very difficult to estimate the distribution of pollutant concentrations using conventional deterministic causal models.

For this reason, a hybrid model combining wavelet neural network and genetic algorithm (GA–WNN) is proposed for predicting 5-min series of carbon monoxide (CO) and fine particulate matter ($PM_{2.5}$) concentrations in proximity to an intersection. The proposed model is examined based on the measured data under two situations. As the measured pollutant concentrations are found to be dependent on the distance to the intersection, the model is evaluated in three locations respectively, i.e. 110 m, 330 m and 500 m. Due to the different variation of pollutant concentrations on varied time, the model is also evaluated in peak and off-peak traffic time periods separately. Additionally, the proposed model, together with the back-propagation neural network (BPNN), is examined with the measured data in these situations. The proposed model is found to perform better in predictability and precision for both CO and $PM_{2.5}$ than BPNN does, implying that the hybrid model can be an effective tool to improve the accuracy of estimating pollutants' distribution pattern at intersections. The outputs of these findings demonstrate the potential of the proposed model to be applicable to forecast the distribution pattern of air pollution in real-time in proximity to road intersection.

1. Introduction

Road intersections have been identified as air pollution hotspots in urban road networks because of the increasing emissions resulted from the complex traffic flows (e.g., free, interrupted,

* Corresponding author. School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Shanghai 200240, China; Department of Urban and Regional Planning, University of Florida, PO Box 115706, Gainesville, FL 32611-5706, USA.
E-mail addresses: zrpeng@sjtu.edu.cn, zpeng@dcp.ufl.edu (Z.-R. Peng).

http://dx.doi.org/10.1016/j.atmosenv.2014.12.058
1352-2310/

congested) and vehicular state (e.g., idling, acceleration, deceleration, cruise) (He et al., 2009; Soulhac et al., 2009; Gokhale, 2011). Recent research further shows that the air quality impact of major roads is significant within a range of up to 250–500 m distance (Zhou and Levy, 2007; HEI, 2010). Such a long range from the intersection tends to cover a dense population, and thus the estimation of air pollution level is of great concern and it is crucial to select an appropriate method to do precise forecast.

As reviewed by Gokhale and Khare (2004), the majority of deterministic models fail to estimate the 'extreme' or short-time variations of pollutant concentrations, and statistical distribution models cannot predict the entire concentration range. Subsequently, these models are extended but cannot be widely used, since they require many input variables such as detailed information of reaction mechanisms, chemical kinetics, transport and other parameters which are usually not well known yet (Inal, 2010). Although traditional statistical models such as regression and linear stochastic models are easy to use, they depend critically on time-series data (Gokhale and Khare, 2004). Typically, data with discontinuous noise often come across in environmental studies and thus influence the prediction accuracy of these models. Moreover, they cannot well model the complex nonlinear environment behavior (Singh et al., 2012). Fortunately, artificial neural network (ANN) that can deal with these limitations have been applied for estimating the ground-level air pollution with good outcomes (Kukkonen et al., 2003; Grivas and Chaloulakou, 2006; Singh et al., 2012; Elangasinghe et al., 2014).

The back-propagation neural network (BPNN), which is one of the mature and most popular ANN models (Rumelhart and McClelland, 1986), has been frequently used in modeling the nonlinear air pollution process recently (Cai et al., 2009; Zhang and Peng, 2014). However, the conventional BPNN has low learning speed and difficulty in choosing the proper size of network, and is easy to fall into local minima (Solhmirzaei et al., 2012). Therefore, the wavelet neural network (WNN) is introduced by Zhang and Benveniste (1992) that it combines the time-frequency characteristic of wavelet transformation with the self-learning of conventional neural network. Using a wavelet function as the activation function of hidden neurons to reveal properties of the measured signal in localized regions of the joint time-frequency space, the neural network structure will achieve powerful approximation and prediction effects, and the local optima problem of gradient-based algorithm is also avoided (Lu et al., 2009). As reported, the pollutants dispersions are complex at intersections owing to the instantaneously variable traffic-related emission intensity and local meteorology, especially at fine-time scales (e.g., minutes) (Tiwary et al., 2011). Hence, WNN provides strong advantages in dealing with nonlinear mapping and on-line estimates. So far, the application of WNN in air quality evaluation is less reported among the literature reviewed. Furthermore, the network training determines prediction stability, but many previous studies concentrate on the best prediction and neglect the training results (Feng et al., 2011). Genetic algorithm (GA) serves as a global optimization method on the solution of complex problems (Lu et al., 2009). Based on GA, the weights and bias of neural network can be optimized to increase the stability of forecasting air pollutants and to fulfill the aim of online estimates, which has demonstrated good results in ozone concentration forecast (Feng et al., 2011).

In this paper, an iterative method combining the strength of back-propagation (BP) in weight learning and GA's capability of searching the satisfying solution is proposed for optimizing WNN, named GA–WNN. The proposed model is developed to estimate the air pollutant concentrations of 5-min scale in proximity to a road intersection. We take into account two pollutants characterized with different nature, i.e. carbon monoxide (CO) and fine particulate matter ($PM_{2.5}$), to test the models responding to varying pollutants. To date, the estimation of air pollutants around intersections is less addressed at fine-time scales (e.g., 5 min) but crucial since decision makers can adopt real-time traffic control measures on intersections that are responsible for high levels of air pollution in surroundings to relieve their effects (Zito et al., 2008; Galatioto and Zito, 2009).

2. Field experiment and data collection

2.1. Experimental area

A suburban site in Shanghai, China was selected for data collection (Fig. 1a), which involves a busy signalized intersection crossing by the 4-lane Dongchuan Rd. and 2-lane Cangyuan Rd. (Fig. 1b). Local traffic constitutes the dominant source of air pollutants nearby. Pollution monitoring was conducted simultaneously at roadside (black solid dot in Fig. 1b) and setbacks (circles with numbers (No.1–3) in Fig. 1b). Roadside site lies in sidewalk along the campus of Shanghai Jiao Tong University (SJTU), and setbacks sit on SJTU campus with 110 m, 330 m and 500 m from the intersection, respectively. The campus is considered for setback layout because it is open from the intersection to setbacks without high-rise buildings blocking the dispersion of pollutants, which is ideal to estimate the immediate effect of local traffic emissions on neighborhoods. Three setbacks characterize different spatial locations from the intersection, and can help to understand air pollution conditions of distance gradient effect.

2.2. Data collection

The field campaign lasted for four sunny days in spring 2013, and daily measurement was split into morning (7:00–9:00), midday (11:00–14:00) and afternoon (16:00–18:00). Morning and afternoon are taken as peak traffic periods while midday is identified as off-peak period of traffic.

The actual pollution level varies remarkably at fine-time scale around intersections (Zito et al., 2008), and thus the measurement of mass concentrations of CO (ppm) and $PM_{2.5}$ (ug/m^3) was run at minute scale. Minute-by-minute concentrations of $PM_{2.5}$ were collected using TSI Sidepak AM510, a measurer with light-scattering techniques. Five-second instantaneous concentrations of CO were recorded by Langan T15n electrochemical sensor, and then 20 continuous observations were averaged into one 1-min sample. The devices were both factory calibrated and estimated at outdoor stations in Shanghai prior to this field campaign. In this study, two sets of portable monitors were used to detect both pollutants at roadside and setbacks, respectively, which were set up 1.7 m above the ground and close to breathing zone of adult pedestrians. Finally, two days' data was collected at No.1 setback, and data collection continued for one day at No.2 and No.3 setbacks. It is noted that roadside measurements are used to estimate the pollutant decay with setback from the intersection rather than to evaluate the models because of a limited observations collected there.

In street-scale traffic environments, air pollution depends heavily on traffic, meteorology and spatial locations (Buonanno et al., 2011). In this study, the traffic condition was recorded by two cameras (circles in lower right of Fig. 1b), and traffic variables were manually counted on 5-min interval which is a rather stable interval to reduce sampling randomness. According to vehicle structure and fuel types, traffic volumes (vehicles/5 min) were subdivided by four categories, (1) PCMUV, passenger-cars and medium-utility vehicles (e.g., taxi, jeep); (2) LDV, light-duty vehicles (e.g., coach, bus); (3) MDV, medium-duty vehicles (e.g., truck);