Atmospheric Pollution Research 7 (2016) 557–566

HOSTED BY

ELSEVIER

Contents lists available at ScienceDirect

Atmospheric Pollution Research

journal homepage: http://www.journals.elsevier.com/locate/apr

Original article

Air pollutants concentrations forecasting using back propagation neural network based on wavelet decomposition with meteorological conditions

Yun Bai [a], Yong Li [b], Xiaoxue Wang [c], Jingjing Xie [d,*], Chuan Li [b,e]

[a] College of Architecture, Anhui Science and Technology University, Anhui 233100, China
[b] School of Environmental and Biological Engineering, Chongqing Technology and Business University, Chongqing 400067, China
[c] Nanan District Environmental Monitoring Station of Chongqing, Chongqing 400060, China
[d] College of Resource and Environment, Anhui Science and Technology University, Anhui 233100, China
[e] State Research Center of System Health Maintenance, Chongqing Technology and Business University, Chongqing 400067, China

ARTICLE INFO

Article history:
Received 23 August 2015
Received in revised form
15 January 2016
Accepted 16 January 2016
Available online 13 February 2016

Keywords:
Air pollutants concentrations
Forecasting
Back propagation neural network
Meteorological data
Stationary wavelet transform

ABSTRACT

Air quality forecasting is an effective way to protect public health by providing an early warning against harmful air pollutants. In this paper, a model W-BPNN using wavelet technique and back propagation neural network (BPNN) is developed and tested to forecast daily air pollutants (PM_{10}, SO_2, and NO_2) concentrations. Firstly, stationary wavelet transform (SWT) is applied to decompose historical time series of daily air pollutants concentrations into different scales, of which the information represents wavelet coefficients of air pollutant concentration. Secondly, the wavelet coefficients are used to train a BPNN model at each scale. The input data for forecasting contain the wavelet coefficients of the air pollutants concentrations 1-day in advance, and local meteorological data. The suitable groups of the input variables are determined by correlation analysis method. At last, the estimated coefficients of the BPNN outputs for all of the scales are employed to reconstruct the forecasting result through the inverse SWT. The proposed approach is tested using data during 1/1/2011 to 26/12/2011 in Nan'an District of Chongqing, China. The results show that the W-BPNN model has better forecasting performance for the three air pollutants than mono-BPNN model in terms of the statistics indexes (mean absolute percentage error, root mean square error and correlation coefficient criteria) and the forecasting accuracy of the number of relevant days of individual air quality index.

1. Introduction

Air pollution is a significant risk factor affecting human's health containing respiratory system and cardiovascular system, and environmental conditions. Recently, various environmental protection agencies of different countries have established the national ambient air quality standards to protect public health and the environment. The public is informed of Air Quality Index (AQI) calculated from air pollutants concentrations forecasted and associated health risks through government announcements (Zhang et al., 2012). Therefore, an accurate and reliable model for forecasting air pollutants concentrations is important since it can provide advanced air pollution information at an early stage such that guiding the works of air pollution control and public health protection.

Generally, air pollution is caused by two factors: pollutant emissions (Gantt et al., 2010; Urbanski et al., 2011; Gao et al., 2014) and meteorological conditions (Wang et al., 2013; L. Wang et al., 2014; Wu et al., 2014,b; Russo et al., 2015). The pollutant emissions are the sources of pollution, and the meteorological conditions are the controlling factors for air pollutants' transfer and diffusion in atmosphere environment. He et al. (2013) found that the meteorological conditions play an essential role in the daily fluctuation of air pollutants concentrations. Recently, much research in air pollution forecasting has been devoted to the formulation and development of models with the meteorological data-for example, statistics model (Ozel and Cakmakyapan, 2015), autoregressive integrated moving average

* Corresponding author. No. 9, Donghua Road, Fengyang, Anhui Province, 233100, China.
E-mail address: xiejingjing0914@foxmail.com (J. Xie).

Peer review under responsibility of Turkish National Committee for Air Pollution Research and Control (TUNCAP).

http://dx.doi.org/10.1016/j.apr.2016.01.004

(Samia et al., 2012; Jian et al., 2012), artificial neural network (ANN) (Chaudhuri and Acharya, 2012; Elangasinghe et al., 2014; Feng et al., 2015; Pauzi and Abdullah, 2015), community multi-scale air quality model (CMAQ) (Chen et al., 2014; Wu et al., 2014a; Djalalova et al., 2015), weather research and forecasting model with chemistry (WRF-Chem) (Saide et al., 2011; Chuang et al., 2011), fuzzy inference system (Domańska and Woktylak, 2012), grey model (Pai et al., 2013a), and other hybrid methods (Chen et al., 2013; Corporation, 2013; Russo and Soares, 2014; Yahya et al., 2014). These methods have achieved good performances for air pollution forecasting result from their functions giving possibilities for discovering the new dependencies between data gathered in sets.

Among these models, ANN, which has the capabilities of nonlinear mapping, self-adaption, and robustness, has been proved its superiority and widely used in forecasting fields. Recently, various structures of the ANN have been developed for improving the forecasting performances of air pollutants concentrations. Feng et al. (2011) applied back propagation neural network (BPNN) to forecast ozone concentration. Wu et al. (2011) considered dust storms when improving Elman network in predicting PM_{10} in Wuhan, China. Paschalidou et al. (2011) used multilayer perceptron (MLP) and radial basis function (RBF) techniques to forecast hourly PM_{10} concentrations in Cyprus. Pai et al. (2013b) adopted neural network and fuzzy learning approach (ANFIS) to forecast oxidant concentration in 24-h. Antanasijević et al. (2013) employed general regression neural network (GRNN) to forecast PM_{10} concentration. The comparison with the PCR (combination of principal component analysis and multilinear regression) model has shown that the GRNN has significantly better performance than the PCR model. Ababneh et al. (2014) designed a using three-layer feed forward neural network (FFNN) and recurrent Elman network to forecast PM_{10} concentrations 1 day advance in Yilan County, Taiwan.

Although these studies have been already done to improve the ANN and obtain better forecasting results, it is still required to improve the forecasting accuracy. It is a universal truth that the structures of the ANN only depend upon the input of the historical data (meteorology and pollutants) without directly taking into account the underlying physical or chemical processes and thus entail much less input and parameter data. However, the excellent forecasting performances of the ANN model generally depend on data representation, because different representations can entangle and hide more or less the different explanatory factors of variation behind the data (Li et al., 2015). Therefore, feature extraction, learning, and analysis of historical data is the key to ensure forecasting accuracy.

Considering the representations and the mutual relations of the historical data (meteorology and pollutants) mentioned above, in this paper, the authors address a model using the BPNN based on wavelet decomposition with meteorological conditions. The nonlinear mapping capability of the BPNN and the multi-resolution characteristics of the wavelet transformation are integrated to improve the forecasting accuracy: (1) to enhance the characterization of the daily air pollutants concentrations, stationary wavelet transform (SWT) is used to decompose the single mixed features into multiple simple features; (2) to identify the relationships of the input variables (the pollutants and the meteorological conditions), correlation analysis method is applied; and (3) to simulate the changes of the pollutants concentrations, a BPNN model is employed to generate the wavelet coefficient of the next-day air pollutants concentrations in each SWT scale, and then the inverse SWT is subsequently employed to reconstruct the outputs of all the scales into the final results.

The rest of the paper is organized as follows. The modeling methodologies are described in Section 2. The air pollutants concentrations and meteorological data are presented in Section 3. The forecasting performance criteria are also introduced in this section. Section 4 illustrates the results and discussion. Conclusions are given in Section 5.

2. Methodologies

In this section, the systematic methodology of a developed BPNN by wavelet decomposition (W-BPNN) approach is described in detail. In the first subsection, the basic BPNN model is introduced for building the regression model. The SWT is introduced in the second subsection for the multi-scale analysis of the regression model. The final subsection provides modeling steps of the W-BPNN approach.

2.1. Basic BPNN model

The BPNN, multilayer feed forward network, is one of the common used neural networks. In the BPNN architecture, the artificial neurons are organized as layers and the information strictly flows forward, and the errors of the network are propagated backwards. The architecture of this network is consisted of input layer, one or more hidden layers and output layers. Each layer is composed of a number of neurons. A general structure of the three layers BPNN is shown in Fig. 1.

The mathematic expressions of the outputs of hidden layer and output layer are as follows (Trigo and Palutikof, 1999),

$$h_j = f_{hidden}\left(\sum\nolimits_{i=1}^{m} w_{ij}x_i\right), \text{ and } Y_k = f_{output}\left(\sum\nolimits_{j=1}^{n} w_{jk}h_j\right), \qquad (1)$$

where $\mathbf{X} = (x_1, x_2,\ldots, x_i)$ $(i = 1,2,\ldots, m)$ represents the inputs, $\mathbf{h} = (h_1, h_2,\ldots, h_j)$ $(j = 1,2,\ldots, n)$ represents the outputs of the hidden layer, $\mathbf{Y} = (Y_1, Y_2,\ldots, Y_k)$ $(k = 1,2,\ldots, p)$ represents the outputs of the network, $\mathbf{w}$ represents the weight matrix between two layers, and $f_{hidden}(.)$ and $f_{output}(.)$ are transfer functions of hidden layer and the output layer, respectively. In this paper, nonlinear transfer function as sigmoid is used in the hidden layer and pure linear is used in the output layer as a linear transfer function.

Back propagation, meaning “error backward propagation”, requires a known, desired output for each input value in order to calculate the loss function gradient. The gradient is fed to the optimization method which in turn uses it to update the weights, in an attempt to minimize the loss function (Lalis et al., 2014). Generally, least mean square error is applied to compare the networks outputs with the actual outputs (Rumelhart et al., 1986).

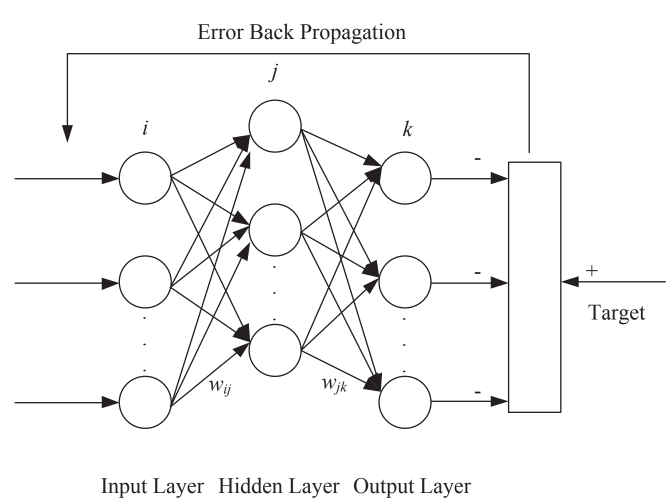

Fig. 1. Structure diagrams of BPNN.