PM2.5 Estimation Based on Image Analysis

Xiaoli Li[1,2,3*], Shan Zhang[2] and Kang Wang[2]
[1]Beijing Key Laboratory of Computational Intelligence and Intelligent System; Engineering
Research Center of Digital Community, Ministry of Education, Beijing, 100124, China
[2]Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China
[3] Beijing Future Network Science and Technology Innovation Center, Beijing 100124, China
[e-mail: lixiaolibjut@bjut.edu.cn]
*Corresponding author: Xiaoli Li

Abstract

For the severe haze situation in the Beijing-Tianjin-Hebei region, conventional fine particulate matter (PM2.5) concentration prediction methods based on pollutant data face problems such as incomplete data, which may lead to poor prediction performance. Therefore, this paper proposes a method of predicting the PM2.5 concentration based on image analysis technology that combines image data, which can reflect the original weather conditions, with currently popular machine learning methods. First, based on local parameter estimation, autoregressive (AR) model analysis and local estimation of the increase in image blur, we extract features from the weather images using an approach inspired by free energy and a no-reference robust metric model. Next, we compare the coefficient energy and contrast difference of each pixel in the AR model and then use the percentages to calculate the image sharpness to derive the overall mass fraction. Furthermore, the results are compared. The relationship between residual value and PM2.5 concentration is fitted by generalized Gauss distribution (GGD) model. Finally, nonlinear mapping is performed via the wavelet neural network (WNN) method to obtain the PM2.5 concentration. Experimental results obtained on real data show that the proposed method offers an improved prediction accuracy and lower root mean square error (RMSE).

1. Introduction

$\mathbf{T}$he seasonal deterioration in air quality demands serious attention. Rapid economic development, population expansion, accelerated urbanization, and increased energy consumption are all affecting the air quality in China. In particular, the air pollution problem caused by fine particulate matter (PM2.5) particles has become increasingly prominent with regard to urban air quality and has gradually become the focus of attention [1,2,3]. In 2012, a new ambient air quality standard (GB 3095-2012) was released in China. The level of fine PM2.5 particles, with an aerodynamic diameter of 2.5 μm or less, is an important new monitoring indicator [4]. Therefore, PM2.5 particles as a major component of air pollution, have received widespread attention.

As main air pollutant, the pollution of PM2.5 remains in the air for a long time, because of its relatively small size, it will result in a decrease in visibility, which will, in turn, affect health, daily life, traffic and the urban landscape as well as crop yields and ecosystems [5]. To effectively deal with the problem of PM2.5 pollution, it is necessary not only to develop an effective model for analyzing and predicting PM2.5 concentration, but also to better grasp the temporal and spatial distribution of PM2.5 particles in the atmosphere. It is of great significance to provide timely and accurate air quality information to environmental monitoring departments and management personnel. At present, the commonly used PM2.5 prediction methods include process-based PM2.5 prediction methods [6] and data-based PM2.5 estimation methods [7,8]. In a process-based PM2.5 prediction method, the processes of pollutant generation, transportation, transformation and sedimentation are simulated. However, the model resolution, meteorological boundary conditions, inventory of emission sources and other parameters are difficult to determine and compute. The actual atmospheric physicochemical processes that cause air pollution are approximated and simplified, affecting the accuracy of the predictions. Relatively speaking, PM2.5 estimation methods based on data offer stronger versatility. Such a method does not require an in-depth analysis of the physicochemical mechanisms in the atmosphere. Based on monitoring data, a mapping relationship between the relevant variables and the PM2.5 concentration can be established, from which the PM2.5 concentration can be predicted. Data-based PM2.5 estimation methods can be divided into linear regression(LR)-based nonlinear modeling methods [9,10] and neural network (NN)-based nonlinear modeling methods [11]. However, LR modeling is a linear method. It has difficulty accurately describing nonlinear relationships, and it is not suitable for modeling the nonlinear processes occurring in the atmosphere. The NN approach has advantages that make up for the shortcomings of LR and other classical algorithms, just like the NN in literature [12], [13]. In practical applications, it has been fully demonstrated that NN models can be effectively used to solve nonlinear modeling problems involving complex environmental systems [14,15,16]. The numerical stability and accuracy of such a model are much better than those of traditional models. However, the quality of the acquired monitoring data greatly affects the accuracy of the forecasting results.

In view of the monitoring problem of PM2.5, researches generally choose to use sensors monitoring data. Considering the shortcomings of the sensor, such as high price, high maintenance cost and limited coverage, etc, the study chooses a convenient method to avoid it. The concentration of PM2.5 can be estimated using pictures acquired at any time and locations, so that air quality index can be easily and quickly obtained. After building the proposed prediction model, we can only use the daily weather photos to estimate the PM2.5 concentration, rather than monitoring data from air quality sensors. In conclusion, it can be seen that the proposed method is not only convenient, but also can avoid the errors caused by the accuracy of monitoring equipment.

At present, with the popularity of the Internet, smart phones are indispensable tools in our daily lives. Given a feasible means of image acquisition, images can be analyzed to quantify and predict the PM2.5 concentration in the environment. In view of this, PM2.5 estimation based on extensive image analysis has become a simple, inexpensive and fast method of monitoring PM2.5 levels. This approach avoids the problem of using data obtained through separate monitoring, which can result in poor prediction accuracy. However, little research of this type has been done. Liu et al. proposed a method of extracting six features from an image [17]. The dark channel priority (DCP) is first used to estimate the image transmission, and then, the image contrast and entropy are measured as additional features. Furthermore, the effects of sky color and solar location on PM2.5 concentration estimation are taken into account. Support vector regression (SVR) [18] is also applied to predict PM2.5 concentrations by combining all features. However, the model relies heavily on manually premarked reference regions at different depths, which considerably limits its scope of application.

This paper proceeds as follows. Two important models used in this study are introduced in Section 2. Section 3 explains our algorithm and predictive model in detail. The experimental process and a discussion of the results are presented in Section 4. Finally, Section 5 offers a summary of the entire study.

2. Related Works

2.1 A Linear Autoregressive (AR) Model

The premise of the so-called free energy principle is assumed to be similar to the Bayesian brain hypothesis regarding how the cognitive process is controlled by a generative model in the brain [19, 20]. The generative model obtained through this method can be regarded as a probabilistic model and includes two components: a prior probability and a posterior probability. However, it mainly allows the brain to identify useful knowledge based on a given input signal and discards the remaining uncertain information. Such an internally generated model is always formed when a scene is presented. In fact, however, there is always a difference between what the brain model predicts and the actual external scene. Thus, the brain-controlled visual analysis system can reverse the possibilities to determine the posterior probability. It is generally believed that this difference is closely related to the quality of visual perception [21, 22]. Therefore, it is a suitable basis for measuring image

sharpness. Simplified-reference (RR), no-reference (NR), and full-reference (FR) image quality assessment (IQA) [23, 24, 25] can be performed through the use of free energy and on-site statistics using these methods. Appropriate analysis and integration to propose a universal model are desirable to achieve higher performance.

In particular, it is assumed that the model m that is generated internally depends on changes in visual perception and that the situation of the external scene is represented by modifying the sizes of the parameters τ. When an external signal c is presented, its entropy is calculated from the joint distribution of two components generated by $p(c, \tau \mid m)$:

$$-\log p(c/m) = -\log \int p(c, \tau / m) \mathrm{d}\tau \tag{1}$$

We introduce an additional term $q(\tau \mid c)$ in the denominator and numerator to derive the following from equation (1):

$$-\log p(c/m) = -\log \int q(\tau/c) \frac{p(c, \tau/m)}{q(\tau/c)} \mathrm{d}\tau \tag{2}$$

where $q(\tau \mid c)$ is an additional posterior distribution of the parameters in the model for a particular signal c. This distribution can be regarded as a true posterior approximation of the model parameters $p(\tau \mid c, m)$ generated by the brain. Once the adjusted parameter set τ or the image signal c is sensed, the best explanation for signal c in $q(\tau \mid c)$ is sought. The generative model of the brain can greatly reduce the difference between the generated approximate posterior $q(\tau \mid c)$ and the actual value.

For simplicity, we gradually reduce the dependency of the generative model m until the two are irrelevant.

By introducing Jensen's inequality into equation (2), we obtain:

$$-\log p(c) \le -\int q(\tau \mid c) \log \frac{p(c, \tau)}{q(\tau, c)} \mathrm{d}\tau \tag{3}$$

The right side of the above formula is an upper bound on the free energy, which is defined as follows:

$$f(\tau) = -\int q(\tau/c) \log \frac{p(c, \tau)}{q(\tau, c)} \mathrm{d}\tau \tag{4}$$

The principle of free energy can be straightforwardly applied based on the difference between the best interpretation of an input visual signal and the internally generated model. Thus, the free energy can be considered a natural representation of the visual perceptual quality of an image. As a result, the free energy estimate for an image can be expressed as:

$$f(s) = f(\hat{\tau}) \tag{5}$$

And:

$$\hat{\tau} = \arg \min_{\tau} f(\tau \mid m, c) \tag{6}$$

However, models with higher levels of expression better approximate the functioning of the brain, and the complexity of the calculation process is also higher. According to previous experience and theory, the more parameters a selected model contains, the higher its cost [26]. Therefore, it is difficult to estimate suitable models from observations.

An AR model can be used to simulate various natural scenes by adjusting its parameters in a simple and effective way [27, 28]. Its parameters are represented as immutable object conversions. Linear autoregressive (AR) models are used to approximate m, and we choose to generate our model as a linear AR model because such a model is simple and easy to construct and can be made to represent various natural scenes by modifying its parameters [26]. For an input image signal c, the AR model is defined as:

$$c_n = \gamma^t(c_n)\kappa + c_n \tag{7}$$

where c_n is the pixel of interest. $\gamma^t(c_n)$ is a vector of t nearest neighbors of c_n. $\kappa = (\kappa_1, \kappa_2, ..., \kappa_t)$ is the coefficient vector in the AR model. The character t indicates transposition. The integer ε_n represents an error term. A matrix representing a linear system is obtained based on the following premise:

$$\hat{\kappa} = \arg\min_{\kappa} \|c - \gamma\kappa\|^2 \tag{8}$$

where $c = (c_1, c_2, ..., c_t)^T$ and $\gamma(i,:) = \gamma^t(c_i)$. The least-squares method is used to solve the linear system. Next, we estimate the solution using the following formula:

$$\hat{c}_n = \gamma^t(c_n)\hat{\kappa} \tag{9}$$

According to a previous analysis in the literature [21], predictive coding has a great influence on the minimization of free energy. Therefore, the key point is that it can accurately express the entropy of the prediction deviation between the fixed-order AR model c and $\hat{c}$, as described by efficient coding theory [29] and Infomax theory [30]. In fact, it is assumed that the internal generative model is an AR model, and the free energy minimization process described in the previously cited article is performed using the AR model of the input signal c based on the least bit [31]. Therefore, when an input image signal is received from the outside world, its free energy is expressed as:

$$f(c) = -\sum_i p_i(c_\Delta) \log p_i(c_\Delta) \tag{10}$$

where c_Δ is the error between the input value and the corresponding prediction and $p_i(c_\Delta)$ is the probability density of gray value i given c_Δ.

2.2 GGD model

It has been found that a decorrelating effect can be achieved by applying local nonlinear operations to log-contrast luminance to separate zero log contrast from local average displacements and to normalize the local variance of the log contrast. In addition, the resulting normalized brightness values of natural images are well characterized by a unit normal Gaussian, and this property has been used to mimic the contrast gain masking of human vision [32]. Therefore, we first calculate the normalized coefficients of average contrast for distorted images based on the methods used in [33, 34].

Next, it is assumed that the allocation of the aforementioned coefficients has statistical features that are affected by distortion. For example, as Ruderman [30] found, the coefficients of a given natural image tend to coincide with a Gaussian appearance, whereas Gaussian blurring gives these coefficients a more Laplacian appearance. In addition, a generalized Gaussian distribution (GGD) can be used to effectively capture the broader

statistics of distorted images. Therefore, we use the relevant definitions provided in [35] to estimate a GGD with zero mean:

$$f(x; \lambda, \mu^2) = \frac{\lambda}{2\beta \mathrm{H}(\frac{1}{\lambda})} \exp\left(-\left(\frac{|x|}{\beta}\right)^\lambda\right) \tag{11}$$

where:

$$\beta = \lambda \sqrt{\left(1 + \frac{1}{\lambda}\right) \Big/ \left(1 + \frac{3}{\lambda}\right)} \tag{12}$$

Here, the gamma function $H(\cdot)$ is given by:

$$\mathrm{H}(x) = \int_0^\infty t^{x-1} e^{-t} \mathrm{d}t \qquad x > 0 \tag{13}$$

In equation (11), for each input x, the parameter λ controls the trend of the distribution, and the variance is represented by μ^2. In this study, a distribution with a mean value of 0 is selected due to the symmetry of the Mean Subtracted Contrast Normalized (MSCN) coefficients. We use this parameter model to accommodate the empirical MSCN distributions in both distorted and undistorted images. For each given image, we estimate a set of parameters (λ, μ^2) from the GGD fit to the MSCN coefficients - the original ratio and the reduced resolution obtained after low-pass filtering - and we then subsample these factors to form a set of features.

3. Predictive model

3.1 Wavelet theory

Wavelet neural network (WNN) [36] was proposed by Zhang Qinghua et al. to combine the merits of wavelet analysis and artificial neural network (ANN). Such a network can avoid falling into local optima and has the characteristics of simultaneous local time and frequency analysis [37, 38]. The WNN approach has been widely used in many fields [39, 40, 41].

At present, the importance of wavelet analysis theory and its extensive applications are leading to widespread concern within the scientific and technological communities. The emergence of wavelet analysis is considered to be a milestone in Fourier analysis. Wavelet analysis has enabled many breakthroughs in approximation, differential equations, identification, computer vision, image processing, and nonlinear science. Wavelet analysis was developed to overcome the deficiencies of Fourier transformation. A serious problem in Fourier transformation is that time information is discarded. Thus, the time characteristics of a Fourier-transformed signal cannot be judged. A wavelet is a waveform with a certain length and a mean value of 0. The wavelet transform function consists of a mother wavelet function. After size scaling and translation, wavelet analysis turns a signal into a number of wavelet functions.

The wavelet transform [42] can shift the basic wavelet function $\phi(t)$ into different ranges θ, and then, the signal $\omega(t)$ can be analyzed at different scales x by means of the

inner product. The wavelet transform can shift the basic wavelet function into different ranges.

$$f_\omega(x,\theta) = \frac{1}{\sqrt{x}} \int_{-\infty}^{+\infty} \omega(t)\phi(\frac{t-\theta}{x})\mathrm{d}t \quad x>0 \tag{14}$$

where n is the formula, varying θ is equivalent to moving a lens parallel to the target, and varying x is equivalent to moving the lens closer to or farther away from the target.

It can be seen from this equation that by transforming the wavelet basis function, wavelet analysis can be used to analyze different parts of the signal characteristics, and in both cases, it offers signal direction selectivity. Therefore, this technique has been extensively applied both as a mathematical foundation as and an analysis technique.

3.2 Network structure

The three basic principles of WNN are as follows:
(1) Topology
The WNN topology is based on the back propagation neural network (BPNN) topology, and wavelet basis functions (WBF) are applied to extract the functions of the hidden-layer nodes of the WNN in this study via the backward propagation of error during signal propagation. **Fig. 1** shows the topology of the wavelet neural network.
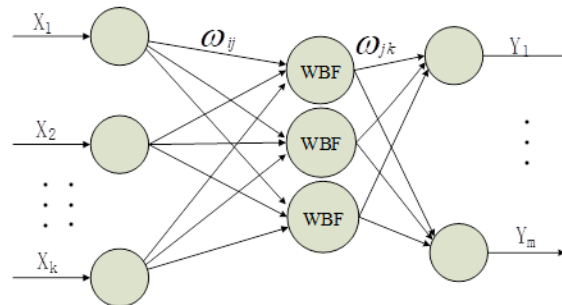

Fig. 1. The topology of the wavelet neural network

In this figure, the input parameters are $X_1, X_2,..., X_K$; $Y_1, Y_2,..., Y_m$ are the predictive outputs. The ω_{ij} and ω_{jk} are the weights of the WNN. When given $x_i (i = 1, 2,..., k)$ as the input signal sequence, the output of the hidden layer is:

$$h(j) = h_j((\sum_{i=1}^{k} \omega_{ij} x_i - b_j)/a_j) \quad j=1,2,...,l \tag{15}$$

In equation (15), $h(j)$ is the output value of the jth node in the hidden layer, ω_{ij} is the weight of a connection from the input layer to the hidden layer, a_j and b_j are WBF expansion factors, and h_j is a wavelet function.

In this paper, we select Morlet wavelet functions, for which the mathematical expression is as follows:

$$y = \cos(1.75x)e^{-x^2/2} \tag{16}$$

The output of the WNN is calculated as:

$$y(k) = \sum_{i=1}^{l} \omega_{ik} h(i) \qquad k = 1,2,...,m \qquad (17)$$

where the ω_{ik} and $h(i)$ are the connection weights and outputs, respectively, of the hidden layer, l and m represent the numbers of hidden and output-layer nodes, respectively.

(2) Network weight correction

Similar to the case of a BPNN, the weights of the WNN and the WBF parameters in this paper are established according to the gradient correction method. Specifically, the prediction error of the network is calculated as follows:

$$e = \sum_{k=1}^{m} yn(k) - y(k) \qquad (18)$$

where $yn(k)$ and $y(k)$ represent the expected values and the prediction outputs of the WNN, respectively.

The WNN connection weights and the WBF coefficients are both corrected according to the error e:

$$\omega_{n,k}^{i+1} = \omega_{n,k}^{i} + \Delta\omega_{n,k}^{i+1} \qquad (19)$$

$$a_k^{i+1} = a_k^i + \Delta a_k^{i+1} \qquad (20)$$

$$b_k^{i+1} = b_k^i + \Delta b_k^{i+1} \qquad (21)$$

among them:

$$\Delta\omega_{n,k}^{(i+1)} = -\eta \frac{\partial e}{\partial \omega_{n,k}^i} \qquad (22)$$

$$\Delta a_k^{i+1} = -\eta \frac{\partial e}{\partial a_k^{(i)}} \qquad (23)$$

$$\Delta b_k^{i+1} = -\eta \frac{\partial e}{\partial b_k^{(i)}} \qquad (24)$$

where η is the learning rate.

(3) Algorithm process

The learning process of the WNN algorithm is described as follows:

1) Network initialization. Randomly initialize the parameters of the wavelet function: the expansion factors a_k, the shift factors b_k and the weights ω_{ij} and ω_{jk}. In addition, set the learning rate of the network.

2) Sample grouping. Separate the samples into training and test samples in accordance with a ratio of 4:1.

3) Predictive output. Present the training samples as input to the NN to compute the values predicted by the network and the error e between the actual and desired values.

4) Weight adjustment. On the basis of the error e, revise the wavelet function parameters and the NN weights such that the actual predicted values will be close to the expected values.

5) Check whether the goal has been accomplished, if not, return to step 3).

In summary, the framework of the proposed method is presented in **Fig. 2**.
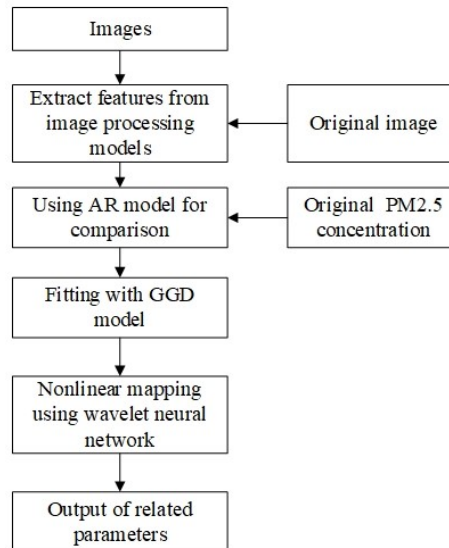
**Fig. 2.** The flowchart of the proposed method

4. Discussion and analysis of results

4.1 Data Sources

Nowadays, many smartphones are equipped with high-quality imaging and powerful computing power, which can detect and quantify PM2.5 in the air by analyzing photos of outdoor scenes. Therefore, we choose to use images to detect and quantify PM pollution. In this study, Image data set plays a fundamental role in image-based prediction of PM2.5 concentration. By capturing pictures of different scenes, such as schools, parks, buildings, roads, lakes and so on, the used image data set involves 750 pictures. The image resolution ranges from 1082*1458 to 4032*3024, and the image size ranges from 0.1-10M. All photos were collected by different partners with different resolutions. The devices used are apple 6splus, apple 8plus, Nikon D750, Fuji XT20, SONY ILCE7M2, etc. For each photograph in the database, we find the PM2.5 concentration of the corresponding photograph taking time from the PM2.5 historical data of the monitoring station. **Fig. 3** presents some examples in the database. Two basic rules should be followed to ensure the quality of the samples. The first principle is to take photos on a clear day without wind or breeze, and the shooting location should be within 1 km of the air quality monitoring point radius. By limiting the wind and site, the PM2.5 value corresponding to the acquired image can be guaranteed to be close to the real PM2.5 value reported by the monitoring point. The second rule is that the captured image must contain the sky, which accounts for about 1 / 3, 1 / 2 of the top of the area image, and avoid facing the sun.

This article uses an experimental dataset consisting of 750 photographs of schools, parks, highways, and buildings acquired in different seasons and different hardware during the past three years. First, in the experiment, all photos are divided into two groups, one for

training and the other for testing, in accordance with ratios of 9:1. In the experiment, the neural network is a typical three-layer structure. The WNN used in the method has a frame structure of 10-10-1. The total number of iterations was chosen to be 400. The wavelet basis function is Morlet mother wavelet basis function. The details are shown in Section 3.
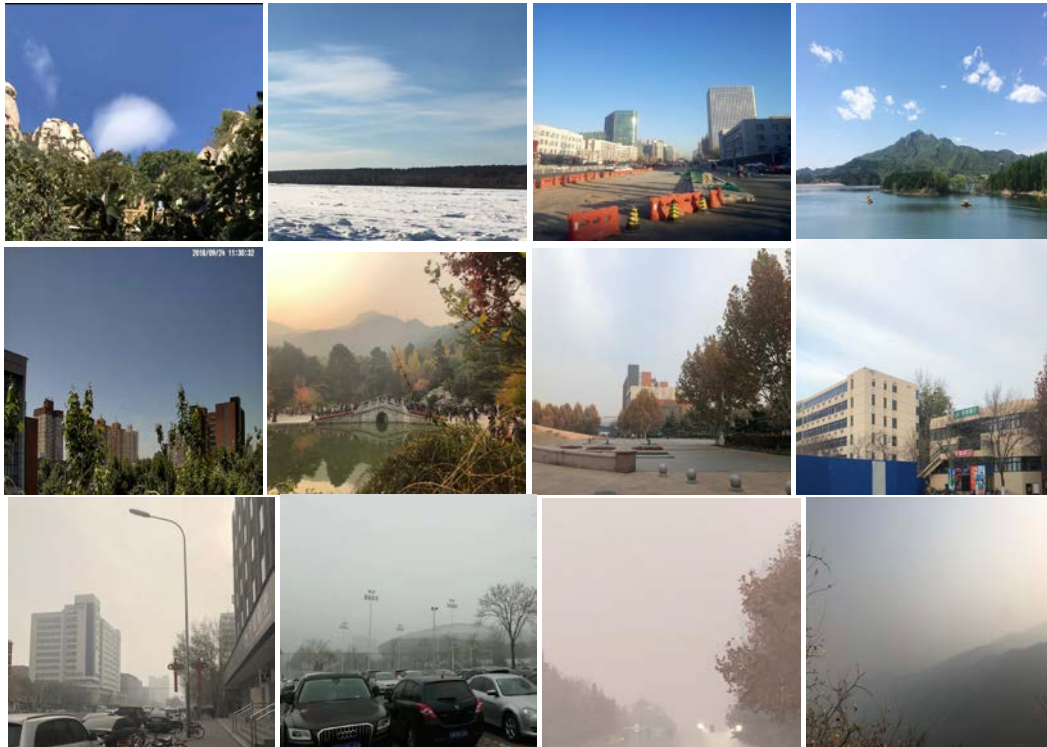
Fig. 3. Weather Images at Different Concentrations of PM2.5

This section mainly explains the superiority of our proposed model in terms of its predictive performance. For this purpose, the following discussion addresses three aspects of our analysis: an introduction to the methods considered for comparison, a numerical comparison, and a visual comparison.

4.2 The methods considered for comparison

It is found that different levels of haze pollution will greatly affect the contrast and visual quality of images. Therefore, the evaluation model measurement method of three kinds of image quality information extraction is applied to the whole image database, and the training model integrates the characteristics of color and contrast to predict the visual quality of the image. Several popular related prediction models are considered in this paper, which can be divided into three categories based on their application scenarios.

The first category consists of three newly designed models based on natural statistics, namely, Integrated Local Natural Image Quality Evaluator (IL-NIQE) [43] and

Accelerated Screen Image Quality Evaluator (ASIQE) [44]. In the cited studies, the authors used high-quality images to create several NS models. When multiple changes were made to the images, the quality scores decreased.

The second category of models includes a pair of highly advanced models, namely, No Reference Image Quality Metric for Contrast distortion (NIQMC) [45] and Blind Image Quality Measure of Enhanced images (BIQME) [46]. These two studies focused specifically on the quality assessment of images with varying contrast to guide automatic image enhancement [47]. These methods are related to the image-based prediction of PM2.5 concentrations because PM2.5 particles have an enormous effect on the images recorded by cameras.

The last category corresponds to a new Picture-based Predictor of PM2.5 Concentration (PPPC) [48] that generates real-time estimates of PM2.5 concentrations using images captured by cell phones or cameras. This model possesses great advantages in terms of prediction accuracy and implementation efficiency. Another one is Photo-Based PM2.5 Concentration Estimation (PPCE) [49]. In the study, two types of features (including gradient similarity and distribution shape of pixel values in saturation images) were extracted from captured photographs to estimate PM2.5 concentration. The proposed method requires only easily available pictures, which greatly reduces the complexity and difficulty of the experiment. At the same time, compared with the above model, it also improves the accuracy and performance.

Table 1. Performance comparison between the proposed method and other models

Model	PLCC	KRCC	RMSE
IL-NIQE	32.37%	17.26%	83.12
ASIQE	12.13%	9.996%	87.21
NIQMC	41.56%	29.60%	79.92
BIQME	54.28%	37.19%	73.78
PPPC	80.8%	60.79%	51.88
PPCE	78.23%	58.09%	57.69
Proposed	81.78%	61.05%	60.49

4.3 A numerical comparison

For the sake of testing and verifying the efficacy of the presented WNN model, we consider that the performance of an image quality analysis (IQA) method is usually evaluated from the perspective of predictive ability, that is, predictive accuracy and predictive monotonicity. The calculation of the correlation performance usually requires a parameter regression process to eliminate the nonlinearity of the prediction score. In general, there are three kinds of logistic functions that can be used for nonlinear mapping. We use a four-parameter logistic function to nonlinearly map the prediction result x to the subjective score:

$$f(x) = \frac{\lambda_1 - \lambda_2}{1 + \exp(-\frac{x - \lambda_3}{\lambda_4})} + \lambda_2 \tag{25}$$

where x represents the input score and $F(x)$ is the mapped score. The free parameters $\lambda_j (j=1,2,3,4)$ are determined during curve fitting. Next, to illustrate the performance of this indicator, we calculate the following three measures, which are typically used for this purpose:

(1) The Kendall rank correlation coefficient (KRCC) is used to measure the order consistency between two inputs:

$$KRCC = \frac{Q_c - Q_d}{0.5Q(Q-1)} \tag{26}$$

where Q_c and Q_d represent the numbers of consistent and inconsistent items in the test set, respectively, and Q represents the total number of images in the test dataset.

(2) The Pearson linear correlation coefficient (PLCC) describes the correlation between the input and output.

The larger the absolute value is of this measure is, the more accurate the prediction. It is defined as:

$$PLCC = \frac{\sum_i (a_i - \overline{a}) \cdot (b_i - \overline{b})}{\sqrt{\sum_i (a_i - \overline{a})^2 \cdot \sum_i (b_i - \overline{b})^2}} \tag{27}$$

where a_i and $\overline{a}$ are the subjective rating of the ith image and the overall mean subjective rating of all images, respectively, and similarly, b_i and $\overline{b}$ are the objective score of the ith image after nonlinear regression and the corresponding average value.

(3) The root mean square error (RMSE), which is focused on predictive consistency, is defined as follows:

$$RMSE = \sqrt{\frac{1}{Q} \sum_i \left(a_i - b_i \right)^2} \tag{28}$$

The KRCC and PLCC represent the correlation between the subjective opinion score and the result of the objective mass fraction prediction. The closer these values are to 1, the higher the prediction accuracy of the model. Based on the above three evaluation criteria, the prediction quality of a model is better if it has a higher PLCC and KRCC while also having a smaller RMSE.

Based on the above test dataset and evaluation criteria, we measured the relative performance of each test model, and the resulting numerical comparisons are shown in **Table 1**. Obviously, according to the PLCC, KRCC and RMSE values, the proposed model achieves the best performance. Compared with the other methods, our technique achieves clear performance gains. Specifically, the energy and contrast differences for the AR model coefficients of each pixel are measured individually. Then, the global quality score is derived by computing the image sharpness using percentage pooling. Moreover, the proposed model is a unique model with PLCC and KRCC values of more than 80%. Compared with the PPPC method, which shows the best performance among the other models, our prediction method achieves relative performance gains of 51% and 118% in terms of the PLCC and KRCC, respectively. In addition, the relative performance gains

between the proposed model and the BIQME algorithm are 1% in terms of the PLCC and 33% in terms of the KRCC.

4.4 A visual comparison

The scatter plot is an intuitive way of illustration. The scatter plot is generated by using the PM2.5 values estimated by the model and the actual values, as shown in **Fig. 4**. The abscissa is the number of test samples, and the ordinate is the concentration of PM2.5. Through the intuitive comparison, we can obviously find that the sample points of the prediction effect of the method used have better convergence and linearity than other tested prediction models. This shows that our predictor can produce predictions that are more consistent with the truth value.
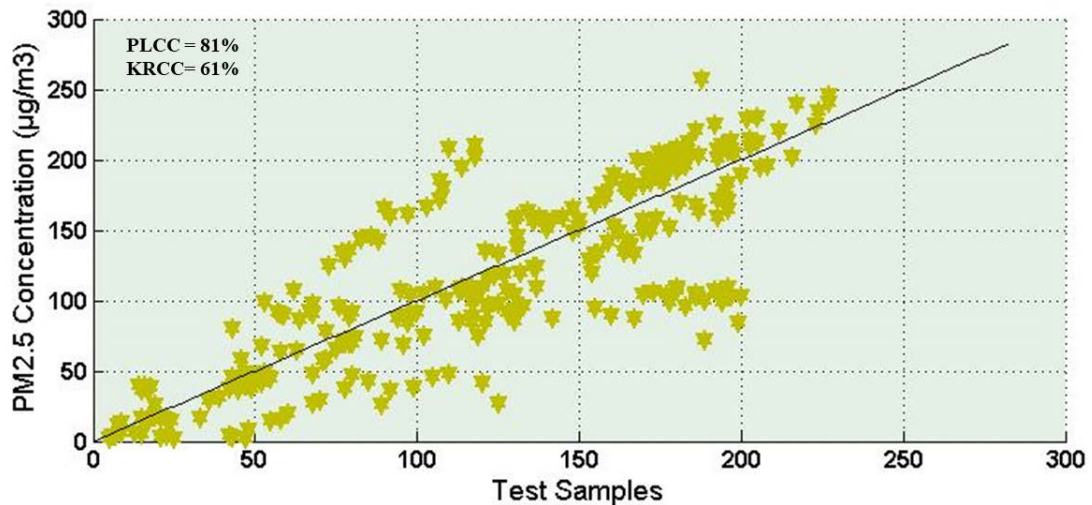

Fig. 4. The scatter plot of the proposed method

5. Conclusions

In this paper, a PM2.5 concentration monitoring method based on image analysis technology is proposed with the aim of improving the overall prediction accuracy and reducing the cost of forecasting. First, on the basis of an analysis of the parameters of an autoregressive (AR) image model and the principle of free energy and inspired by the no-reference free energy based quality metric (NFEQM) model, features are extracted from images captured by mobile phones. The level of image blurring is estimated from the similarity of the local estimated AR parameters. The basic assumption of similarity introduces a significant performance gain compared to the previous NFEQM metric. Specifically, the energy and contrast differences for the AR model coefficients of each pixel are measured individually. Then, the global quality score is derived by computing the image sharpness using percentage pooling. The original image is compared with the

predicted results. The residuals obtained with the two methods are related to the PM2.5 concentration, and the parameters are fitted with a GGD model. Finally, the wavelet neural network (WNN) method is used to perform nonlinear mapping to obtain the predicted PM2.5 concentration. The final experimental results prove that the presented WNN prediction method has higher prediction accuracy and a lower RMSE than those of other state-of-the-art methods. However, to date, we have designed only a simple and experience-based approach, and other advanced technologies in the field of machine learning will be further studied in the future.

References

[1] K. Gu, J.-F. Qiao, and W.S. Lin, "Recurrent air quality predictor based on meteorology- and pollution-related factors," *IEEE Trans. Ind. Informat*, Jan.2018. Article (CrossRef Link)

[2] S. C. Park, M. K. Park, and M. G. Kang, "Estimating ground-level PM2.5 in China using satellite remote sensing." *Journal of Environment Science & Technology*, vol. 48, no. 13, pp. 7436-7444, 2014. Article (CrossRef Link)

[3] Mohammad Arhami, Nima Kamali, Rajabi, and Mahdi Mohammad," Predicting hourly air pollutant levels using artificial neural networks coupled with uncertainty analysis by Monte Carlo simulations." *Journal of Environmental Science and Pollution Research*, vol. 20, pp. 4777-4789, Jul. 2013. Article (CrossRef Link)

[4] Ministry of environmental protection, *Ambient Air Quality Standard GB3095-2012*, Environmental Science Press, Beijing, China, 2012.

[5] J. O. Anderson, J. G. Thundiyil, and A. Stolbach, "Clearing the air: A review of the effects of particulate matter air pollution on human health," *Journal of Medical Toxicology*, vol. 8, no. 2, pp. 166, 2012. Article (CrossRef Link)

[6] Heidi M. Waldrip, Rotz C. Alan, and Sasha D. Hafner, "Process-based modeling of ammonia emission from beef cattle feedyards with the integrated farm systems model," *Journal of Environment Quality*, vol. 43, no. 4, pp. 1159-1168, 2014. Article (CrossRef Link)

[7] Perez and Patricio, "Combined model for PM10 forecasting in a large city," *Journal of Atmospheric Environment*, vol. 60, pp. 271-276, 2012. Article (CrossRef Link)

[8] Pablo E. Saide, Gregory R. Carmichael, and Scott N. Spak, "Forecasting urban PM10 and PM2.5 pollution episodes in very stable nocturnal conditions and complex terrain using WRF-Chem CO tracer model," *Journal of Atmospheric Environment*, vol. 45, no. 16, pp. 2769-2780, 2011. Article (CrossRef Link)

[9] Gavin Pereira, Hyung Joo Lee, and Michelle Bell, "Development of a model for particulate matter pollution in Australia with implications for other satellite based models" *Environmental Research*, vol. 159, pp. 9-15, 2017. Article (CrossRef Link)

[10] A. Vlachogianni, P. Kassomenos, and Ari Karppinen, "Evaluation of a multiple regression model for the forecasting of the concentrations of NO_x and PM10 in Athens and Helsinki," *Science of the Total Environment*, vol. 409, no. 8, pp. 1559-1571, 2011. Article (CrossRef Link)

[11] Maher Elbayoumi, Nor Azam Ramli, and Noor Faizah Fitri Md Yusof, "Multivariate methods for indoor PM10 and PM2.5 modeling in naturally ventilated schools buildings," *Atmospheric Environment*, vol. 94, pp. 11-21, 2014. Article (CrossRef Link)

[12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. of IEEE Conference on computer vision and pattern recognition*, pp. 770-778, 2016. [Article (CrossRef Link)](#)

[13] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," in *Proc. of IEEE conference on computer vision and pattern recognition*, pp. 2261-2269, 2017. [Article (CrossRef Link)](#)

[14] Paisan Kittisupakorn, Piyanuch Thitiyasook, and M. A. Hussain, "Neural network based model predictive control for a steel pickling process," *Journal of Process Control*, vol. 19, no. 4, pp. 579-590, 2009. [Article (CrossRef Link)](#)

[15] JB Ordieres, EP Vergara, and RS Capuz, "Neural network prediction model for fine particulate matter (PM2.5) on the US-Mexico border in Texas and Chihuahua," *Environmental Modelling & Software*, vol. 20, no. 5, pp. 47-59, 2005. [Article (CrossRef Link)](#)

[16] M.A. Elangasinghe, N. Singhal, and K.N. Dirks, "Complex time series analysis of PM10 and PM2.5 for a coastal site using artificial neural network modeling and k-means clustering," *Atmospheric Environment*, vol. 94, pp. 106-116, 2014. [Article (CrossRef Link)](#)

[17] C.B. Liu, Tsow Francis, and Y. Zou, "Particle pollution estimation based on image analysis," *PloS one*, vol. 11, no. 2, 2016. [Article (CrossRef Link)](#)

[18] Chih-Chung Chang and Chih-Jen Lin, "Libsvm: a library for support vector machines," *ACM Trans. Intel. Syst. Technol*, vol. 2, no. 3, 2011. [Article (CrossRef Link)](#)

[19] K. Friston, "The free-energy principle: A unified brain theory," *Nature Reviews Neuroscience*, vol. 11, pp. 127-138, 2010. [Article (CrossRef Link)](#)

[20] D.C. Knill and A. Pouget, "The Bayesian brain: The role of uncertainty in neural coding and computation," *Trends Neurosci*, vol. 27, no. 2, pp. 712-719, 2004. [Article (CrossRef Link)](#)

[21] G. Zhai, X. Wu, X. Yang, W. Lin, and W. Zhang "A psychovisual quality metric in free-energy principle," *IEEE Trans. Image Process*, vol. 21, no. 1, pp. 41-52, 2012. [Article (CrossRef Link)](#)

[22] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang, "Visual saliency detection with free energy theory," *IEEE Signal Processing Letters*, vol. 22, no. 10, pp.1552-1555, 2015. [Article (CrossRef Link)](#)

[23] K. Gu, Vinit Jakhetiya, J.-F. Qiao, X. Li, W. Lin, and Daniel Thalmann, "Model-based referenceless quality metric of 3D synthesized images using local image description," *IEEE Transactions on Image Processing*, vol. 27, no. 1, pp. 394-405, 2018. [Article (CrossRef Link)](#)

[24] Y. Liu, G. Zhao, K. Gu, X. Liu, and D. Zhao, "Reduced-reference image quality assessment in free-energy principle and sparse representation," *IEEE Transactions on Multimedia*, vol. 20, no. 2, pp. 379-391, 2018. [Article (CrossRef Link)](#)

[25] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Hybrid no-reference quality metric for singly and multiply distorted images," *IEEE Transactions on Broadcasting*, vol. 60, no. 3, pp. 555-567, 2014. [Article (CrossRef Link)](#)

[26] X. Wu, G. Zhai, X. Yang, and W. Zhang, "Adaptive sequential prediction of multidimensional signals with applications to lossless image coding," *IEEE Trans. Image Process*, vol. 20, no. 1, pp. 36-42, 2011. [Article (CrossRef Link)](#)

[27] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Using free energy principle for blind image quality assessment," *IEEE Trans. Multimedia*, vol. 17, no. 1, pp. 50-63, 2015. [Article (CrossRef Link)](#)

[28] K. Gu, G. Zhai,W. Lin, X. Yang, and W. Zhang, "No-reference image sharpness assessment in autoregressive parameter space," *IEEE Trans. Image Process*, vol. 24, no. 10, pp. 3218-3231, 2015. [Article (CrossRef Link)](#)

[29] H. Barlow, "Principles Underlying the Transformation of Sensory Messages," *Sensory Communication, MIT Press: Cambridge, MA, USA*, pp. 217–234, 1961. Article (CrossRef Link)

[30] R. Linsker, "Perceptual neural organization: Some approaches based on network models and information theory," *Annu. Rev. Neurosci*, vol. 13, pp. 257-281, 1990. Article (CrossRef Link)

[31] H. Attias, "A variational bayesian framework for graphical models," in *Proc. of Adv. Neural Inf. Process. Syst*, vol. 12, pp. 209-215, 2000.

[32] M. Carandini, D.J. Heeger, and J.A. Movshon, "Linearity and normalization in simple cells of the macaque primary visual cortex," *J. Neurosci*, vol. 17, no.21, pp. 8621-8644, 1997. Article (CrossRef Link)

[33] A. Mittal, A.K. Moorthy, and A.C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Process*, vol. 21, no. 12, pp. 4695- 4708, 2012. Article (CrossRef Link)

[34] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a completely blind image quality analyzer," *IEEE Signal Process*, vol. 22, no. 3, pp. 209-212, 2013. Article (CrossRef Link)

[35] K. Sharifi and A. Leon-Garcia, "Estimation of shape parameter for generalized Gaussian distributions in subband decompositions of video," *IEEE Trans. Circuits Syst. Video Technology*, vol. 5, no. 1, pp. 52-56, 1995. Article (CrossRef Link)

[36] Q. Zhang and Albert Benveniste, "Wavelet networks," *IEEE Transactions on Neural Networks*, vol. 3, no. 6, 1992. Article (CrossRef Link)

[37] W. Andrew and P. Donald, "Wavelet methods for times series analysis," *Cambridge University Press*, 2000.

[38] Q. Zhang, "Using wavelet network in nonparametric estimation," *IEEE Trans. Neural Networks*, vol. 8, no. 2, pp. 227-237, 1997. Article (CrossRef Link)

[39] Roberto Besteiro, Tamara Arang, and Ortega J. Antonio "Prediction of carbon dioxide concentration in weaned piglet buildings by wavelet neural network models," *Computers and electronics in agriculture*, vol. 143, pp. 201-207, 2017. Article (CrossRef Link)

[40] El-Diastya Mohammed, "Hybrid harmonic analysis and wavelet network model for sea water level prediction," *Applied Ocean Research*, vol. 70, pp.14-21, 2018. Article (CrossRef Link)

[41] Daniel Dunea, Alin Pohoata, and Stefania Iordache, "Using wavelet feed forward neural networks to improve air pollution forecasting in urban environments," *Environmental monitoring and assessment*, vol. 187, no. 7, 2015. Article (CrossRef Link)

[42] C. Xiao and F. S.Wang, "43 cases analysis of neural network using matlab," *Beihang University Press*, 2013.

[43] L. Zhang, L. Zhang, and A.C. Bovik, "A feature-enriched completely blind image quality evaluator," *IEEE Trans. Image Process*, vol. 24, no. 8, pp. 2579-2591, 2015. Article (CrossRef Link)

[44] K. Gu, J. Zhou, J.-F. Qiao, G. Zhai, W. Lin, and A. C. Bovik, "No reference quality assessment of screen content pictures," *IEEE Trans. Image Process*, vol. 26, no. 8, pp. 4005-4018, 2017. Article (CrossRef Link)

[45] K. Gu, W. Lin, G. Zhai, X. Yang, W. Zhang, and C.W. Chen, "No-reference quality metric of contrast-distorted images based on information maximization," *IEEE Trans. Cybern*, vol. 47, no. 12, pp. 4559-4565, 2017. Article (CrossRef Link)

[46] K. Gu, D. Tao, J.-F. Qiao, and W. Lin, "Learning a no-reference quality assessment model of enhanced images with big data," *IEEE Trans. Neural Netw. Learning Syst*, 2017. Article (CrossRef Link)

[47] K. Gu, G. Zhai, W. Lin, and M. Liu, "The analysis of image contrast: From quality assessment to automatic enhancement," *IEEE Transactions on Cybernetics*, vol. 46, no. 1, pp. 284-297, 2016. Article (CrossRef Link)

[48] K. Gu, J.-F. Qiao, and X. Li, "Highly efficient picture-based prediction of PM2.5 concentration," *IEEE Trans. Ind. Elearon*, vol. 66, no. 4, pp. 3176-3184, 2019. Article (CrossRef Link)

[49] G. Yue, K. Gu, and J.-F. Qiao, "Effective and Efficient Photo-Based PM2.5 Concentration Estimation," *IEEE Transactions on instrumentation and measurement*, 2019. Article (CrossRef Link)

Xiaoli Li received the B.E. and M.E. degrees from Dalian University of Technology in 1994 and 1997, Ph.D. degree from Northeastern University, China in 2000. From 2000 to 2003, he was a Research Fellow of Tsinghua University, China, and University Libre de Bruxelles, Belgium. Now he is a professor of Beijing University of Technology. His research interests include intelligent control, multiple model control, adaptive control, robust control.

Shan Zhang received the B.S. degree in Computer Department from Hebei University of Architecture, Hebei, China, in 2016. In 2019, she received her master's degree from Beijing University of Technology, Beijing, China. Her research interests include neural networks, modeling and control in complex system systems.

Wang Kang post-doctor at Faculty of Information Technology, Beijing University of Technology. From University of Science and Technology Beijing, he received the bachelor degree in 2012, and Ph. D. degree in control science and engineering in 2018. His research interest covers optimal control and intelligent control.