Applied Soft Computing 35 (2015) 280–290

Contents lists available at ScienceDirect

Applied Soft Computing

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

Artificial neural networks for infectious diarrhea prediction using meteorological factors in Shanghai (China)

Yongming Wang[a,*], Jian Li[b], Junzhong Gu[a], Zili Zhou[c], Zhijin Wang[a]

[a] Department of Computer Science & Technology, East China Normal University, Shanghai 200241, China
[b] Department of Acute Infectious Disease, Shanghai Municipal Center for Disease Control & Prevention, Shanghai 200336, China
[c] School of Physics and Engineering, Qufu Normal University, Qufu, Shandong 273165, China

ARTICLE INFO

Article history:
Received 21 October 2013
Received in revised form
23 December 2014
Accepted 17 May 2015
Available online 3 July 2015

Keywords:
Artificial neural networks
Forecasting model
Infectious diarrhea
Meteorological factors
Sensitivity analysis

ABSTRACT

Infectious diarrhea is an important public health problem around the world. Meteorological factors have been strongly linked to the incidence of infectious diarrhea. Therefore, accurately forecast the number of infectious diarrhea under the effect of meteorological factors is critical to control efforts. In recent decades, development of artificial neural network (ANN) models, as predictors for infectious diseases, have created a great change in infectious disease predictions. In this paper, a three layered feed-forward back-propagation ANN (BPNN) model trained by Levenberg–Marquardt algorithm was developed to predict the weekly number of infectious diarrhea by using meteorological factors as input variable. The meteorological factors were chosen based on the strongly relativity with infectious diarrhea. Also, as a comparison study, the support vector regression (SVR), random forests regression (RFR) and multivariate linear regression (MLR) also were applied as prediction models using the same dataset in addition to BPNN model. The 5-fold cross validation technique was used to avoid the problem of overfitting in models training period. Further, since one of the drawbacks of ANN models is the interpretation of the final model in terms of the relative importance of input variables, a sensitivity analysis is performed to determine the parametric influence on the model outputs. The simulation results obtained from the BPNN confirms the feasibility of this model in terms of applicability and shows better agreement with the actual data, compared to those from the SVR, RFR and MLR models. The BPNN model, described in this paper, is an efficient quantitative tool to evaluate and predict the infectious diarrhea using meteorological factors.

1. Introduction

Infectious diarrhea which can be caused by a variety of bacterial, viral and parasitic organisms remains a major public health problem around the world [1]. As a kind of common and important infectious disease, infectious diarrhea has a serious threat to human health and leads to one billion disease episodes and 1.8 million deaths each year. Infectious diarrhea in young children is a killer illness, especially in developing countries [2,3]. In China, infectious diarrhea is a notified infectious disease which is the biggest developing country. In Shanghai of China, the incidence of infectious diarrhea has significant seasonality throughout the year and is particularly high in the summer and autumn of recent years. Moreover the incidence of infectious diarrhea has been the highest among the class A and B intestinal infectious diseases. Infectious diarrhea has become the focus of the prevention and control of infectious diseases over the years. Health forecasting is a novel area of forecasting, and a valuable tool for predicting future health events or situations such as demands for health services and healthcare needs. It facilitates preventive medicine and health care intervention strategies, by pre-informing health service providers to take appropriate mitigating actions to minimize risks and manage demand [4]. Hence, a robust short-term (week-ahead) prediction model for infectious diarrhea incidence is necessary for decision-making in policy and public health.

Infectious diseases have a closely relation with meteorological factors [5] and can affect infectious diseases in a linear or nonlinear fashion [6]. Over the past couple of decades, there has been a large scientific and public debate on climate change and its direct as well as indirect effects on human health [7]. The effects of meteorological factors, such as temperature, rainfall and relative humidity, on diarrhea diseases incidence have got much more concerning recently. As far as we are concerned with the prediction of diarrhea diseases in literature, many forecasting models

* Corresponding author. Tel.: +86 13482813260.
E-mail address: ymwang819@gmail.com (Y. Wang).

http://dx.doi.org/10.1016/j.asoc.2015.05.047
1568-4946/

based on statistical methods for diarrhea diseases forecasting have been devoted. Alexander et al. [8] evaluated monthly reports of diarrheal disease among patients presenting to Botswana health facilities and compared this to climatic variables (rainfall, temperature, and vapor pressure). Kolstad et al. [9] combined a range of linear regression coefficients to compute projections of future climate change (temperature)-induced increases in diarrhea. Zhao et al. [10] established multiple regression model rolling forecast of daily incidence of infectious diarrhea. Chou et al. [11] applied a climate variation-guided Poisson regression model to predict the dynamics of diarrhea-associated morbidity. The results indicated that the maximum temperature and extreme rainfall days were strongly related to diarrhea-associated morbidity. In Hashizume et al. [12], weekly rainfall, temperature and number of hospital visits for non-cholera diarrhea were analyzed by a Poisson regression model. McCormick et al. [13] studied of temporal and spatial patterns of diarrheal disease and construct a spatial panel regression model suing contemporary acute diarrhea disease and climatic data (temperatures and precipitation) and found there is a strong association between daily mean temperature and precipitation and the incidence of hospitalization. Singh et al. [14] examined diarrhea notifications in Fiji in relation to estimates of temperature and rainfall, using Poisson regression analysis of monthly data for 1978–1998. Result indicated that there were positive associations between diarrhea reports and temperature and between diarrhea reports and extremes of rainfall. Lloyd et al. [15] undertook a global cross-sectional study of diarrhea incidence in children under 5, and assessed the association with climate variables (temperature and rainfall) by linear regression method.

With regard to the fact that number of meteorological factor that effect infectious diarrhea are too much and the inter-relation among them is also very complicated, prediction models based on traditional statistics methods may not be fully suitable for such type of problems. In recent years, artificial intelligence techniques such as artificial neural network (ANN) and support vector machine (SVM) have been employed for developing predictive models in complex prediction problems. Nowadays, ANNs are considered to be one of the intelligent tools to understand the complex problems. As a powerful computational method, ANNs have been widely used in the medical and health field, such as medical diagnosis and disease prediction [16–23]. Moreover, ANN-based hybrid prediction model (hybrid neural networks) have been widely used in different fields and obtained the very good prediction result [24–27]. Compared with other forecast methods, ANNs are advantageous in terms of high data error tolerance, easy adaptability to online measurements.

The feed forward back propagation neural network (BPNN) as a typical ANNs is essentially a mapping function from input to output vector(s) without knowing the correlation between the data. BPNN can implement any complex nonlinear mapping function proved by mathematical theories, and approximate an arbitrary nonlinear function with satisfactory accuracy [28]. However BPNN are characterized by very poor convergence. Several improvements for BPNN were developed. The Levenberg–Marquardt back propagation neural network is a powerful optimization technique that was introduced to the neural net research because it provided methods to accelerate the training and convergence of the algorithm [29].

Based on a literature review, so far as I know to the best knowledge of the authors, there is no works has been carried out to utilize the ANNs method in predicting diarrhea disease, to say nothing of infectious diarrhea. In this paper, an attempt had been made to establish a new BPNN model to predict the weekly number of infectious diarrhea (*WNID*) in Shanghai of China with a set of meteorological factors as input variable. Also, as a comparison SVR, RFR and MLR models were developed for the same purpose in addition to BPNN model. Finally, since one of the drawbacks of ANNs is the interpretation of the final model in terms of the importance of variables, the calculation is performed using sensitivity analysis.

The rest of this paper is organized as follow. Study area and dataset that is used in this study are briefly described in Section 2. The prediction methods and performance evaluation criteria which are used in this paper are introduced in Section 3. These methods are ANN SVR, RFR and MLR. ANN and SVR, RFR, MLR models are developed in Sections 4 and 5, respectively. In order to investigate the performance of the established model, the prediction results of the ANN model are reported in comparison with the SVR, RFR and MLR models as discussed in Section 6. Section 7 illustrated the sensitivity analysis, and conclusions of this paper are concluded in Section 8.

Table 1
Summary of population in Shanghai [30].

Year	Population (million)	Population growth rate (%)
2004	13.52	
2005	13.60	0.582
2006	13.68	0.575
2007	13.79	0.788
2008	13.91	0.883

2. Study area and diarrhea-meteorological data

In this section, we present the study area in Section 2.1. We also describe the dependent and independent variables used in this paper in Section 2.2.

2.1. Study area

Shanghai is located in the eastern part of China which is the largest developing country in the world, and the city has a mild subtropical climate with four distinct seasons and abundant rainfalls. It is the most populous city in China comprising urban/suburban districts and counties, with a total area of 6340.5 km^2 and had a population of 13.9 million by the end of 2008 (Table 1) [30]. Generally, in epidemiological studies, the incidence is an important index that has become the focus for many researchers. The incidence is calculated by dividing the disease number with the total population. Since the study population was relatively stationary during the time period from 2005 to 2008 with the annual growth rate below 1% (Table 1), the trend of incidence during that time period could be similarly prescribed by the trend of disease cases number. Hence we used the number of infectious diarrhea as the response variable in our models. Description of the input (meteorological factors) and output (the weekly number of infectious diarrhea) parameters for constructing the BPNN model have been given in Table 2.

2.2. Dependent and independent variables

The *WNID* cases data from 2005.1.3 to 2009.1.4 were collected from National Disease Supervision Information Management System, which is a modifiable diseases reporting system of real-time, online, based-on-case information. The infectious diarrhea cases were all clinical or laboratory-confirmed cases of infectious diarrhea and reported by hospital diagnostic. In choosing meteorological factors, an important consideration is of course whether a factor is likely to have a significant influence on the infectious diarrhea. The Spearman Rank Correlation Analysis (ARCA) was carried to determine the effects of nine meteorological factors on infectious diarrhea. According to the ARCA results (see Table 3), T_{max}, T_{min}, T_{avg} and AP_{avg} are statistically significant correlation ($p < 0.01$) and RH_{min}, RH_{avg}, SD, WS_{avg} and R are statistically correlation ($p < 0.05$) with infectious diarrhea. This implies that it is entirely feasible to

Table 2
Description of input and output parameters for constructing the prediction models.

Input (independent variables)	Parameter (Unit)	Symbol
Meteorological factors	Weekly average maximum temperature (□)	T_{max}
	Weekly average minimum temperature (□)	T_{min}
	Weekly average temperature (□)	T_{avg}
	Weekly average minimum relative humidity (%)	RH_{min}
	Weekly average relative humidity (%)	RH_{avg}
	Weekly average atmospheric pressure (hPa)	AP_{avg}
	Weekly average sunshine duration (h)	SD
	Weekly average wind speed (m/h)	WS_{avg}
	Weekly average rainfall (mm)	R_{avg}
Output (dependent variables)	Parameter (Unit)	Symbol
Infectious diarrhea	Weekly number of infectious diarrhea (case)	*WNID*

use the meteorology factors to forecast the infectious diarrhea in Shanghai area. The weekly meteorological factors data in Shanghai from 2005.1.3 to 2009.1.4 were obtained from the Shanghai Meteorological Bureau of City Environmental Meteorological Center.

A total of 209 weeks data pairs were collected to develop the BPNN model. The basic statistical characteristics such as minimum, maximum, average, standard deviation of the input and output parameters are illustrated in Table 4. The temporal variation behavior of *WNID* cases and weekly average meteorological factors from 2005.1.3 to 2009.1.4 are illustrated in Figs. 1 and 2, respectively. In this study, the data sets were chosen and segregated in time order. In other words, the data sets of the earlier period were used for training and the data sets of the latest time period were used for testing. From the collected data sets, the data sets ahead 80% as calibration sets were selected for the development of the BPNN models, the remaining data sets as testing sets were used for the verifying the generalization capability of the established BPNN models.

3. The methodology and performance metrics

In this study, three of the most common regression forecasting models were chosen for modeling: the feed-forward back-propagation ANN (BPNN) model, support vector regression (SVR) and the random forests regression (RFR). Moreover,

Table 3
The Spearman rank correlation analysis between input and output variables.

Input variables	*WIDN*	
	p-value	*r*-value
T_{max}	<0.01**	0.7328
T_{min}	<0.01**	0.7569
T_{avg}	<0.01**	0.7563
RH_{min}	<0.05*	0.1671
RH_{avg}	<0.05*	0.1761
AP_{avg}	<0.01**	−0.5705
SD	<0.05*	0.1614
WS_{avg}	<0.05*	−0.1833
R	<0.05*	0.1238

* Statistically correlation.
** Statistically significant correlation.

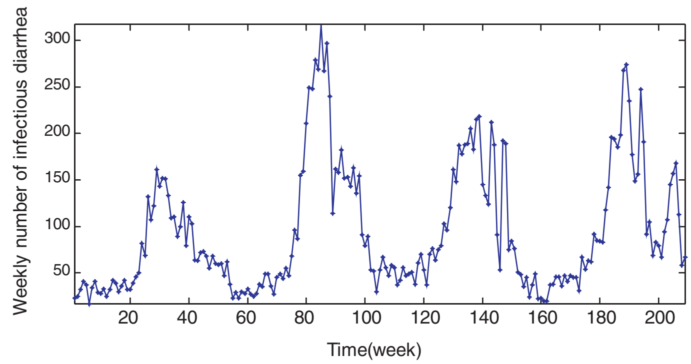

Fig. 1. Temporal variation of *WNID* in Shanghai from 2005.1.3 to 2009.1.4.

multiple linear regression (MLR) was also used to predict the next week *WNID* using meteorological factors.

3.1. Multiple linear regression analysis

As one of the most popular techniques for predictive modeling, multiple regression analysis (MRA) is widely used in engineering sciences to model and analyze the relationship between several independent or predictor variables and a dependent or criterion variable [31]. A multiple linear regression (MLR) prediction model based on the general MRA equation is given below:

$$y = \alpha + \sum_{i=1}^{n} b_i x_i + \varepsilon \quad (1)$$

where y is the dependent variable (predicted result), b_i is regression coefficients represent the independent contributions of each independent variable, α is a constant value, x_i are the values of the independent variables and ε represents the random error term.

3.2. Artificial neural network

Artificial neural networks (ANNs), which use the computer network system to simulate the biological neural network, are a class of flexible non-linear, self-adapting black box intelligent information process system. Detailed discussions on the ANNs can be found in [32–34], but a brief introduction about BPNN is provided here. The BPNN via Levenberg–Marquardt learning algorithm is the most widely used algorithm and has been used with great success to model numerous engineering applications [37,38]. BPNN is composed of an input layer, an output layer and one or more hidden layers that allow the network to learn relationships between input and output variables [35,36]. The following equation gives the mathematical expression of the output of the BPNN:

$$y = g\left(\theta + \sum_{j=1}^{m} v_j \left[\sum_{i=1}^{n} f(w_{ij} x_i + \beta_j)\right]\right) \quad (2)$$

In the Eq. (2), y is the prediction value of dependent variable; x_i is the input value of ith independent variable; w_{ij} is the weight of connection between the ith input neuron and jth hidden neuron; β_j is the bias value of the jth hidden neuron; v_j is the weight of connection between the jth hidden neuron and output neuron; θ is the bias value of output neuron; $g(\cdot)$ and $f(\cdot)$are the activation functions of output and hidden neurons, respectively.

In order for a BPNN to accomplish a required prediction task, BPNN should be trained with training dataset regarding the problem. During training, the weights and biases of the network are iteratively adjusted to minimize the global error between the network outputs and the actual outputs.

Table 4
Basic statistical characteristics of input and output variables.

Variables	Range		Mean	Median	Standard deviation
	Min	Max			
T_{max}	2.3857	36.4000	21.3072	22.7286	8.9847
T_{min}	−2.5000	27.8429	14.5494	15.1714	8.9827
T_{avg}	0.6500	31.8464	17.6415	18.6143	8.8901
RH_{min}	49	87.2857	69.8553	87.2857	7.3985
RH_{avg}	49	87.2857	69.8553	69.8571	7.3985
AP_{avg}	999.9750	1032.8571	1.0161e+03	1015.9143	8.6114
SD	0	10.9857	4.6139	4.6000	2.2250
WS_{avg}	1.3929	6.2179	3.1297	3.0607	0.6818
R	0	34.2571	3.1388	1.6571	4.4943
$WNID$	17	317	94.6268	68	67.5613

3.3. Support vector regression (SVR)

The SVR is a non-linear kernel-based regression method which tries to find the best regression hyper plane with smallest structural risk in a so-called high dimensional feature space [40]. Assume that the training dataset be $\{(x_i, y_i)\}_{i=1}^{n}$, where $x_i \in \Re^d$ is input data, $y_i \in \Re$ is the expected output value of the ith data point in the dataset, d is the dimensionality of samples and n is the number of samples.

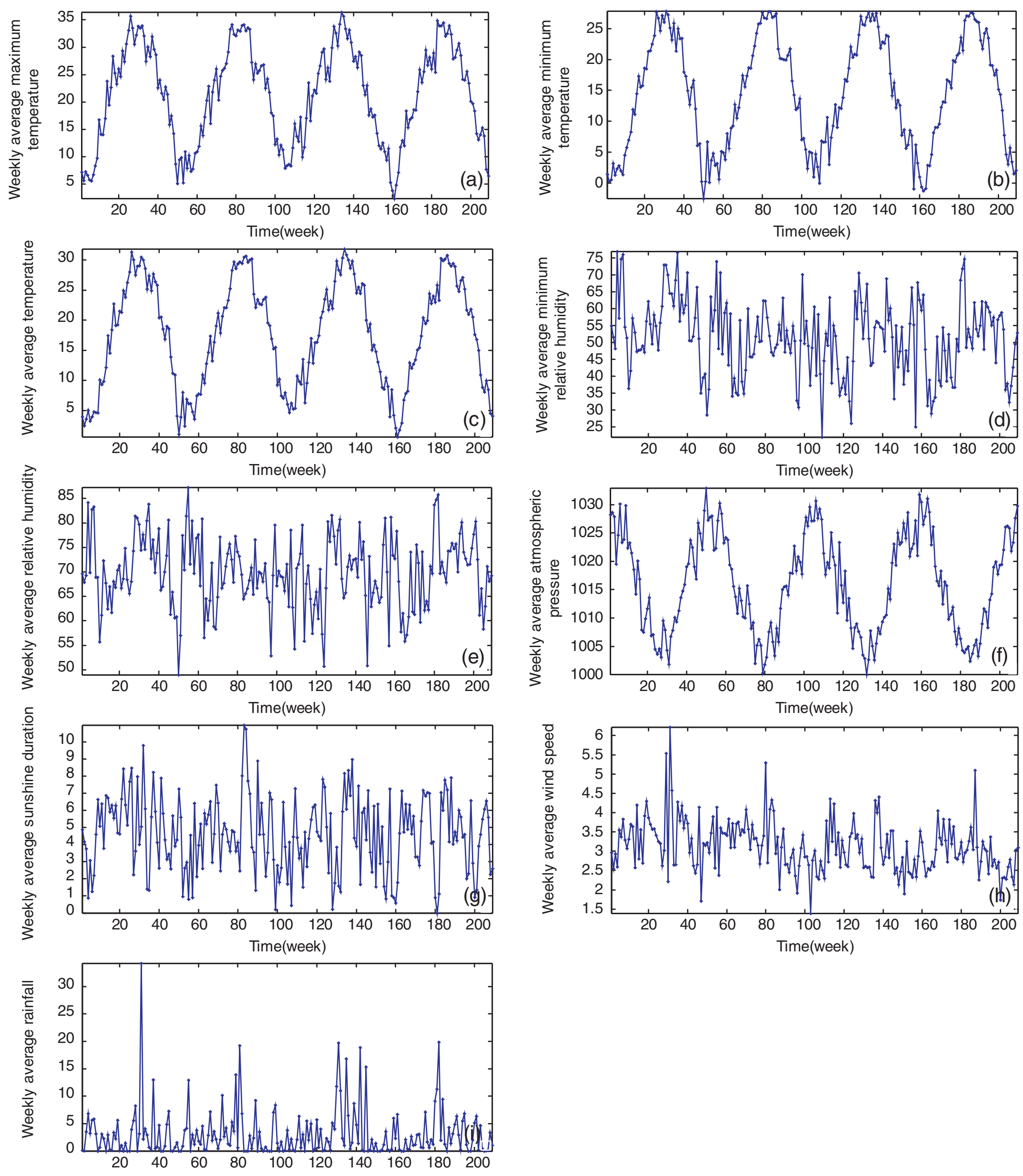

Fig. 2. Temporal variation of meteorological factors in Shanghai from 2005.1.3 to 2009.1.4.

Then, the nonlinear relationship between the input and the output can be formulated by a regression function as follows:

$$y = f(x) = w^T\varphi(x) + b \tag{3}$$

where $\varphi(\cdot)$ denotes kernel function which use a nonlinear mapping function to map the input space to the high-dimensional feature space. w is a vector of weight coefficients and b is a bias constant. When empirical risk ξ_i and structure risk ξ_i^* are considered together, w and b are estimated by minimizing the following optimization problem:

$$\begin{aligned} &\text{minimize} \quad \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \\ &\text{subjected to} \quad \begin{cases} y_i - \left(\left\langle w, \varphi(x_i)\right\rangle + b\right) \le \varepsilon + \xi_i \\ \left(\left\langle w, \varphi(x_i)\right\rangle + b\right) - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{cases} \end{aligned} \tag{4}$$

where C is a modifying coefficient representing the trade-off between empirical risk and structure risk, ε is band area width. The parameters values of C and ε are user-determined for SVR modeling.

3.4. Random forests regression (RFR)

Random forests (RF), proposed by Breiman [41], are an improved classification and regression tree (CART) method. RF regression (RFR) method consists of a collection of unpruned regression trees using different bootstrap samples of the training data. In each bootstrap sample, a random sample with replacement and with the same length, some of the data is repeated, and the left out samples are called out-of-bag (OOB). To construct a tree, at each node, a small random subset of size m of p explanatory variables ($m \le p$) is chosen, and the best split of the feature space is done using only this subset instead of using all variables. The value of m is held constant during the forest growing. Each tree is grown to the largest extent possible. Then the final output is taken as the average of the individual tree outputs. Since the OOB data are not used in the construction of each tree, therefore, it is regarded as a test set to get a running unbiased estimate of the prediction error and variable importance as trees are added to the forest [59]. The OOB error estimate is an unbiased estimate of the generalization error, and this is almost identical to that obtained by N-fold cross-validation [58]. In practice, the number of trees (n) and the size of the variable subset (m) should be optimized to reach the ideal forest by minimizing the OOB error.

3.5. Performance evaluation criteria (PEC)

To date, a variety of PECs that can be used for evaluation and inter-comparison of different models have been proposed and employed in the literature to evaluate the forecasting accuracy, but no single performance measure has been recognized as the universal standard. As a result, we need to assess the performance based on multiple metrics, and it is interesting to see if different metrics will give the same performance ranking for the models to be tested [42]. Among of PECs, the mean absolute error (MAE), the root mean square error (RMSE), the mean absolute percentage error (MAPE), the correlation coefficient (R) and the coefficient of determination (R^2) are the most widely used PECs and will be used in this study. These criteria can be expressed mathematically and calculated by Eqs. (5)–(9).

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \tag{5}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{6}$$

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|y_i - \hat{y}_i\right|}{\left|y_i\right|} \times 100\% \tag{7}$$

$$R = \sqrt{1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}} \tag{8}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}y_i^2} \tag{9}$$

In the above relations, y_t is the actual value, $\hat{y}_t$ is the predicted value, $\bar{y}$ is the average of the actual value, respectively; n is the total number of dataset. The models with the smallest MAE, RMSE and MAPE and the largest R and R^2 are considered to be the best models.

4. Development of BPNN model for predicting *WNID*

The BPNN modeling consists of two steps: the first step is to train the network using training data sets which are used to train the network and adjust the connection weights and biases; the second step is to test the network with testing data sets, which are not used in training step and are used to estimate the generalization error of trained networks. All BPNN models were implemented in Matlab software (Neural Network Toolbox of Matlab v.2013a, The Math Works Inc.).

To avoid the problem of overfitting, the 5-fold cross validation [43], considered to be the most effective method to ensure overfitting does not occur [44], was employed in this study. The implementation of the 5-fold cross validation was carried out by further splitting the calibration dataset into five equal-sized subsets. Four sets (80%) as training set was used for updating the network weights and biases, the remaining set (20%) as validation set was utilized to check the performance of the models and to determine when to stop training to prevent overfitting. The structure of data sets using 5-fold cross validation is shown in Fig. 3. In this study, optimal model architecture and parameters are chosen such that they minimized the root mean square error (RMSE) value between observed values and model outputs on the validation dataset [29,39,49,52,63].

4.1. Data pre-processing and post-processing

It was noticed that the range of input–output parameters are different. In order to improve the efficiency and generalization of neural network models, data pre-processing is necessary for efficient machine learning [46]. Normalization procedure is utilized so that larger input values do not overwhelm smaller inputs, helping reduce network error and speeding up the learning process [47]. There are several different normalization range including 0–1 [45,49], 0.1–0.9 [26], 0.05–1 [29], and 0.05–0.95 [50] in different literature. It is useful to normalize the input–output parameters to the range [0.05, 0.95] to avoid the problem of output signal saturation that can sometimes be encountered in ANN applications [50]. In this study, a comparison experiment was carried out to find the optimal normalization range. The results of comparison are showed in Fig. 6. Analyzing Fig. 6 indicates that the lowest RMSE-value for validation data is obtained using 0.05–0.95. So, in this study, all the input and output parameters were normalized to lie between 0.05 and 0.95, by using Eq. (10). Where x_{norm} is the normalized input/output value, x is the actual input/output value, $x_{\min}$ is the

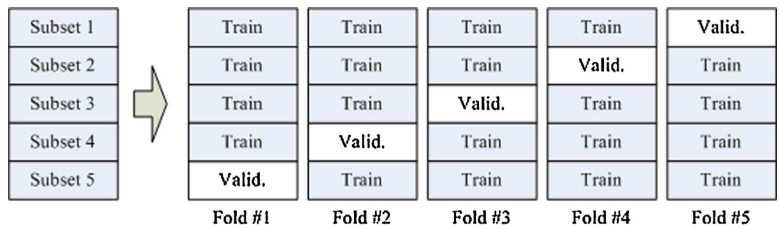

Fig. 3. Schematic illustration of the data partitioning for 5-fold cross validation.

maximum input/output value, x_{max} is the minimum input/output value.

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \times 0.9 + 0.05 \quad (10)$$

It should be noted that, as a result of normalizing the target values, the output of the ANN will correspond to the normalized range. Thus, to interpret the results obtained from the network, a reverse of this procedure is performed, transforming the ANN output data back to the original scale by using Eq. (11), then these data can be applied to the estimate of the model performance.

$$x = \frac{(x_{norm} - 0.05) \times (x_{max} - x_{min})}{0.9} + x_{min} \quad (11)$$

4.2. Determination of optimum network architectures and parameters

In the development of BPNN models, determination of an appropriate architecture (the number of input nodes, the number of hidden layers and hidden nodes and the number of output nodes) and parameters (the activation functions for hidden/output nodes and learning rate) for a particular problem are an important issue as the network architectures and parameters directly affects its computational complexity and its generalization capability [51]. There is no unified approach for determination of an optimal BPNN architecture and parameters [35].

For a given problem, the number of neurons of the input and output layer is determined according the modeling problem being tackled, and therefore in this study the input layer has nine neurons corresponding to the nine input meteorological factors variables and one neuron in the output layer corresponding to the output *WIDN* variable.

For the number of hidden layer, Hecht-Nielsen [36] indicated that a single-hidden layer BPNN with a sufficient number of hidden neurons is sufficient to approximate the corresponding desired outputs arbitrarily close. Therefore, one hidden layer was preferred in this study. To determine the number of hidden nodes, several heuristics approaches (summarized in Table 5) have been used in this study. As can be seen from Table 5, the number of neurons varies from 3 to 19. So, the technique used to determine the number of neurons in the hidden layer was to start with 3 nodes then increasing in steps by adding one neuron each time until to 19. During the training process, the RMSE error on the validation set was monitored. When the RMSE was amplified, the training was stopped and the minimum of the validation error was taken as an indicator for the performance [64]. The average prediction errors (RMSE) for validation data at different hidden neurons level are shown in Fig. 4. As demonstrated in Fig. 4, the network with 9–4–1 was found to be the optimum model architecture for the *WNID* prediction. For each number of hidden neurons, network was trained for 20 times to overcome the randomness of choosing initial weights and biases of neurons.

The activation function (AF) determines the relationship between inputs and outputs of a network and affects the performance of a BPNN model [38]. The linear function (*P*), sigmoid function (*L*) and hyperbolic tangent functions (*T*) are the most common AFs used in the literature [38]. In our study, various different combinations of three AFs with different normalization range were tried out to choose an optimal combination in hidden and output layer. The results in Fig. 5 revealed that the optimum combination is *L*–*P* with normalization range 0.05–0.95. So, sigmoid AF expressed as $f(x) = (1 + \exp(-x))^{-1}$ was used in the hidden layer and linear AF was used in the output layer, respectively.

The BPNN performance depends very much on the learning rate which influences the rate of weight and bias adjustment, and hence influences the rate of convergence [39,54–56]. There is no automatic way to select learning rate, and if incorrect value is specified, the convergence may be exceedingly slow, or it may not converge at all [54]. This study determined the optimum value of learning rate by varying their values from 0.01 to 1.0, the value 0.3 which gives the lowest validation RMSE was selected. Fig. 6 shows the effect of learning rate on the validation RMSE.

Table 5
The network architecture and parameters used in optimum network architecture.

Heuristic	Number of neuron	References
$<2 \times N_i + 1$	<19	Hecht-Nielsen [58]
$2N_i$	18	Kannellopoulas and Wilkinson [59]
$(N_i + N_O)/2$	5	Ripley [60]
$2N_i/3$	6	Wang [61]
$(N_i + N_O)^{1/2}$	3	Masters [62]

5. Development of other models for predicting *WNID*

In order to compare and evaluate the performance of developed BPNN model, three other forecasting models SVR, RFR and MLR also were developed for *WNID* prediction. To obtain meaningful comparisons, SVR, RFR and MLR models were developed with the same

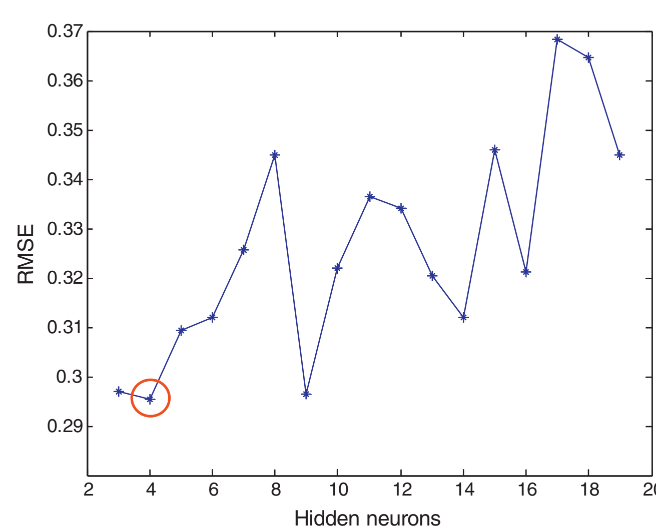

Fig. 4. Effect of the number of hidden neurons on validation RMSE.

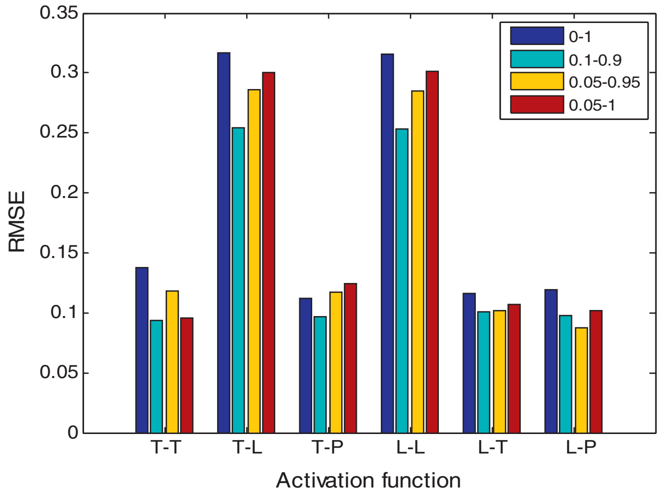

Fig. 5. Effect of activation functions and normalization ranges on validation RMSE.

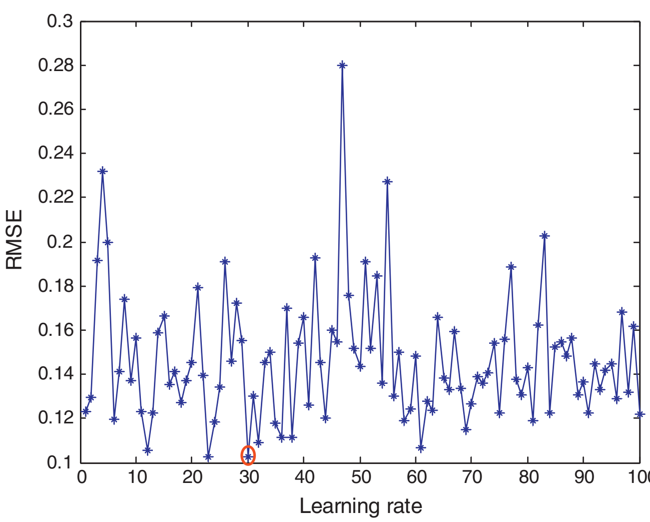

Fig. 6. Effect of the learning rates on validation RMSE.

data sets used for the BPNN model and the prediction results were also tested with the same datasets used to test the BPNN model, therefore, making the results comparable.

The software Matlab version R2012b was used to calculate the coefficients of MLR model. The MLR model for prediction of the *WNID* was derived as presented by the following mathematical equation:

$$\begin{aligned} WNID &= -1972.7903 - 10.9619T_{\max} + 20.8158T_{\min} - 2.6208T_{\text{avg}} \\ &-1.6506RH_{\min} + 0.2993RH_{\text{avg}} + 2.0902AP_{\text{avg}} + 5.7734SD \\ &-15.7205WS_{\text{avg}} + 1.6048R \end{aligned} \tag{12}$$

The detailed results of analysis of variance (ANOVA) for MLR are presented in Table 6. As calculated *F* is greater than the critical *F*, the null hypothesis is rejected and concluded that there is significant difference between regressions (parameters). And the calculated value of the *F*-ratio does not exceed the standard tabulated value of *F*-ratio for a desired level of confidence, say 95%. Therefore, it was concluded that the regression analyses were valid.

For building the SVR forecasting model, popular LibSVM [31][1] software package was adapted in this study. Radial basis function expressed as $K(x, y) = \exp(-\|x - y\|^2/\sigma^2)$ was used as a kernel function because of its advantages and simple implementation only one tuning parameter. Inappropriate parameter (the penalty parameter C and kernel parameter σ) settings may lead to poor regression performance of SVR models. The 5-fold cross validation grid search is selected to determine the best parameters for SVR and calculating the RMSE of the model. The parameter values which give the lowest validation RMSE were selected. The details can be seen from [57], herein we only give the optimal parameters setting: Best $C = 2$ and Best $\sigma = 0.5$ with the RMSE = 0.1286.

When development of RFR models, the number of trees (n) and the size of variable subset (m) are two key parameters that should be optimized to reach the ideal prediction performance. The values of the parameters n and m were optimized simultaneously by using the grid-search method ranging from 10 to 1000 (with step size 10) and from 1 to 9 (with step size 1), respectively. The parameter values which give the lowest RMSE of the OOB data were selected as an indicator for the performance. In this study, n and m values equal to 80 and 6, respectively, were found as optimal and were chosen for subsequent model. The RFR modeling was implemented by using an R package[2] in this study.

6. Prediction results and discussion

In this section, the results obtained from the developed BPNN and the SVR, RFR and MLR models for one-week-ahead *WNID* prediction are presented and compared. With the above selected optimal structure and parameters in Section 4, the BPNN model was trained for 20 times and the results are presented in Table 7 alongside minimum, maximum and mean MAE, RMSE, MAPE, R^2 and R values of the 20 results coming from the 20 trials for training, validation and testing sets. The average values have been used as results when comparing the final performance with SVR, RFR and MLR models. The model outputs were transformed back to the original range for all four models, then five key PECs were calculated using Eqs. (5)–(9) and compared between actual and predicted values for testing period.

The prediction performance statistical values (MAE, RMSE, MAPE, R^2 and R) of the BPNN model and SVR, RFR and MLR models for training, validation, and testing results are summarized in Table 8. According to the obtained results, it is clear that the MAE, RMSE and MAPE are lower for nonlinear models (BPNN, SVM and RFR). The correlation coefficients and coefficient of determination (R and R^2) given by BPNN, SVM and RFR models are also higher. This demonstrates that the performances of nonlinear models are better than that of linear MLR model. The MLR method was chosen only as a representative statistical model in this study. In many other studies [37,38], MLR method was also selected as a benchmark model to compare. The reason of better performances of the nonlinear models over MLR model may be attributed to the complex nonlinear relationship between infectious diseases and meteorological factors. This implies the nonlinearity of the investigated phenomenon.

The comparison of performance of the nonlinear models demonstrates that the BPNN model yields most satisfactory results when all PECs are considered simultaneously. This is evident from a lower MAE, RMSE and MAPE (MAE = 29.136, RMSE = 40.367 and MAPE = 0.2556) and a higher R and R^2 values (R = 0.7949 and R^2 = 0.9077). The reason for this may be related to the fact that ANNs are universal approximator which can approximate a large

[1] More details on this implementation can be found in http://www.csie.ntu.edu.tw/cjlin/libsvm.

[2] http://cran.r-project.org/web/packages/randomForest.

Table 6
Analysis of variance (ANOVA) results for MLR model.

Responses	Source of variance	Degree of freedom	Sum of squares	Mean squares	*F*-statistic	*p*-value
WIDN	Regression	9	486,536.824	54,059.647	34.214	0.000*
	Residual error	155	244,909.722	1580.063		
	Total	164	731,446.545			
	R = 0.816, R^2 = 0.665					

* Significant at the 95% confidence level.

Table 7
The results of BPNN model for testing, validation and testing periods.

PEC	Training			Validation			Testing		
	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean
MAE	18.339	19.270	19.167	23.649	27.013	24.402	27.757	30.645	29.136
RMSE	26.454	27.025	26.856	34.083	37.028	36.113	40.154	40.971	40.367
MAPE	0.2529	0.2878	0.2827	0.3352	0.3864	0.3561	0.2493	0.2745	0.2556
R	0.9084	0.9138	0.9096	0.8572	0.8805	0.8647	0.7879	0.7973	0.7949
R^2	0.9364	0.9391	0.9372	0.9143	0.9274	0.9184	0.9050	0.9102	0.9077

class of functions with a high degree of accuracy. It can be also concluded from Table 8 that the SVR and RFR models have almost equal accuracy and the RFR showed slightly better performance than the SVR model.

Furthermore, Fig. 7 shows time history based comparisons in between predicted and actual values in case of the BPNN and SVR, RFR and MLR models for the test period. It can be obviously seen from Fig. 7 that the BPNN-yielded *WNID* predicted values are in closer agreement with the corresponding actual values than those for SVR, RFR and MLR models, especially peak points. The BPNN model predicted the peak values fairly satisfactorily than SVR and RFR models in forecasting infectious diseases. The underestimation of the peak values and overestimation of the low values are much more for the SVR, RFR and MLR than the BPNN model. It can be observed from these figures that the MLR model seems to be insufficient for the forecasting *WNID*, especially the peak values.

Fig. 8 is scatter plots giving the comparison between the observed values and the predicted values using BPNN, SVR, RFR and MLR during the test period. The linear least square fit scatter plot, its equation, and the R^2 values were shown in Fig. 8. The figures depict that the estimates given by the BPNN model are denser in neighborhood of the straight line and have very less scattered estimates compared with that of SVR, RFR and MLR. In summary, it can be concluded that BPNN has provided more accurate results for prediction of *WNID* compared to SVR, RFR and MLR models.

Table 8
The results of BPNN, SVR, RFR and MLR models for testing, validation and testing periods.

Model	PEC	Training	Validation	Testing
BPNN	MAE	19.167	24.402	29.136
	RMSE	26.856	36.113	40.367
	MAPE	0.2827	0.3561	0.2556
	R	0.9096	0.8647	0.7949
	R^2	0.9372	0.9184	0.9077
SVR	MAE	19.392	26.684	38.461
	RMSE	30.884	35.727	49.909
	MAPE	0.2423	0.3745	0.3494
	R	0.8785	0.8678	0.6612
	R^2	0.9169	0.9202	0.8590
RFR	MAE	26.579	27.290	35.030
	RMSE	36.996	41.809	48.144
	MAPE	0.3714	0.3433	0.3048
	R	0.8156	0.8135	0.6899
	R^2	0.8789	0.8907	0.8687
MLR	MAE	28.892	33.405	40.128
	RMSE	38.291	42.370	54.280
	MAPE	0.4451	0.4451	0.4066
	R	0.8057	0.8080	0.5782
	R^2	0.8723	0.8877	0.8332

7. Sensitivity analyses

Since one of the drawbacks of ANNs is the interpretation physical meaning of the final model in terms of the relative importance of input variables [59], sensitivity analysis [60] can be carried out to quantify the relative significance of each individual input parameter on the objective (output) parameter in the modeling and to examine the contribution of an input variable to the outputs. Identifying sensitive inputs can be helpful for the model designers to carefully treat with such influential parameters in the designing models process [61]. Sensitivity analysis yields the information about the input parameters which must be measured most accurately and the effect of any small increment and decrement of that parameter on the overall design objectives.

The Cosine Amplitude Method (CAM) [62] is one of the sensitivity analysis methods that used to determine the effect of each individual input on the output and has been widely used in the relevant literature [60,63]. So, in this paper, CAM was employed to identify the most sensitive meteorological factors affecting *WIDN.*

In this method, the degree of sensitivity of each meteorological factor is assigned by establishing the strength of the relationship (r_{ij}) between the *WNID* and meteorological factors under consideration. The larger the value of CAM, the greater is the effect on the *WNID*. To use this method, all of the data pairs were expressed in common X-space. The data pairs used to construct a data array X defined as: $X = \{X_1, X_2, X_3, \ldots, X_m\}$. Each of the elements, X_i, in the data array X is a vector of lengths, that is: $X_i = \{x_{i1}, x_{i2}, x_{i3}, \ldots, x_{im}\}$. Thus, each of the data pairs can be thought of as a point in m-dimensional space, where each point requires m-coordinates for a full description. Each element of a relation, r_{ij} results in a pairwise comparison of two data pairs. Therefore, the strength of the relation between the data pairs, x_i and x_j, is given by using Eq. (13):

$$r_{ij} = \frac{\sum_{k=1}^{m} x_{ik}x_{jk}}{\sqrt{\sum_{k=1}^{m} x_{ik}^2 \sum_{k=1}^{m} x_{jk}^2}} \qquad (13)$$

Here, the strengths of relations (r_{ij} values) between the infectious diarrhea (output) and input parameters using the CAM method are shown in Fig. 9. As can be seen, the most effective parameters on the *WNID* are the temperature-related variables

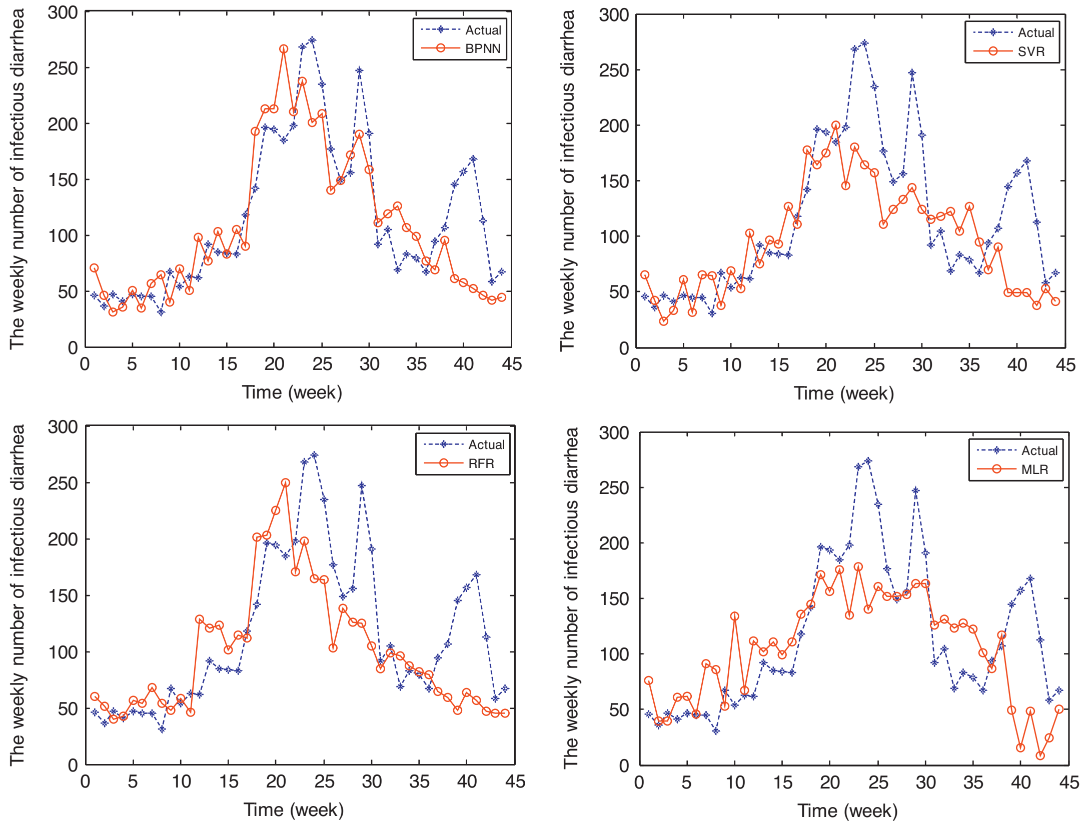

Fig. 7. The actual and predicted values by BPNN, SVR, RFR and MLR models in test period.

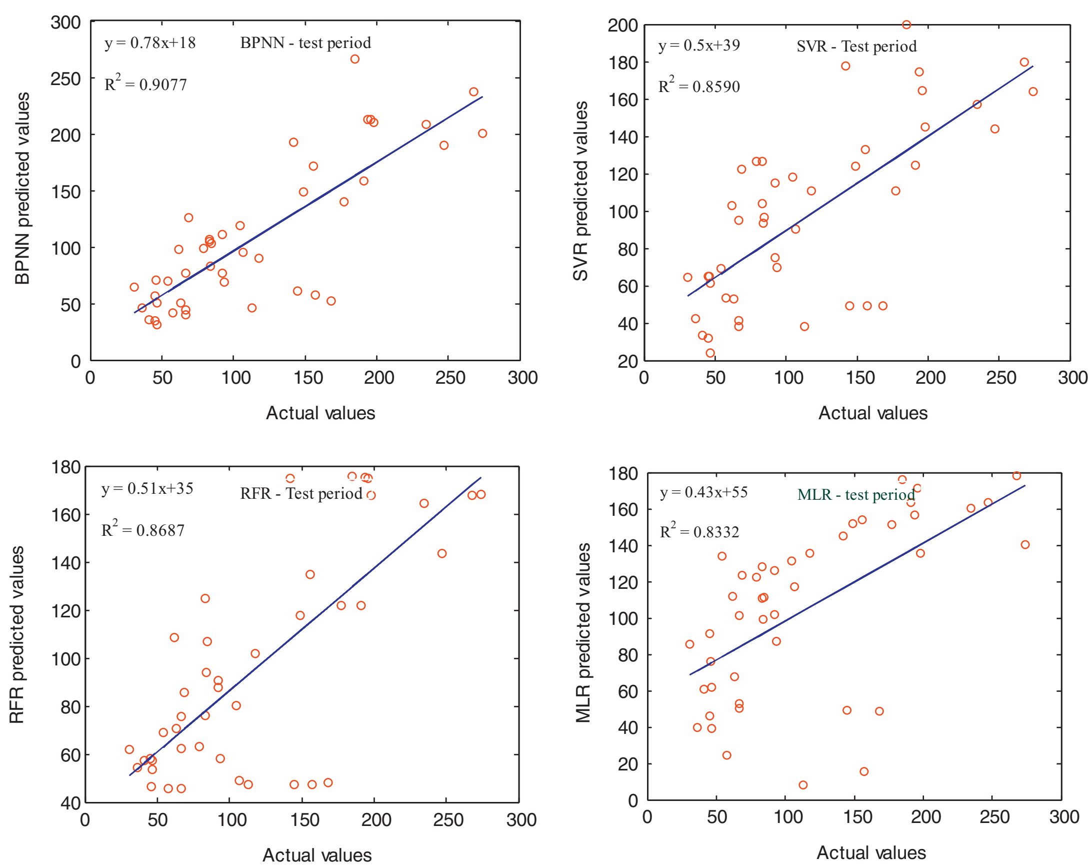

Fig. 8. The scatter plots of the BPNN, SVR, RFR and MLR models for testing period.

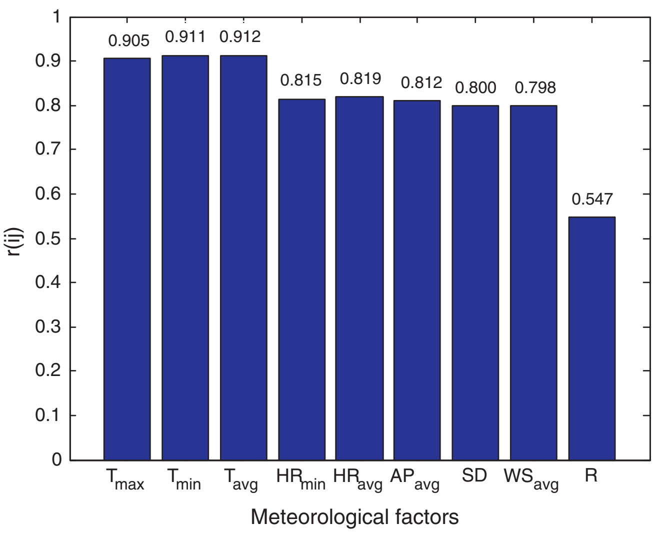

Fig. 9. The effect of meteorological factors on *WNID*.

Table 9
The Chi-square goodness-of-fit test *p*-values for temperature-related variables.

Variable	T_{max}	T_{min}	T_{avg}
T_{max}	–	0.2453	0.2452
T_{min}	0.2453	–	0.2412
T_{ave}	0.2452	0.2412	–

(T_{max}, T_{min}, and T_{avg}), whereas the weekly average rainfall is the least effective parameter on the infectious diarrhea.

In order to further evaluate whether there is a significant difference among T_{max}, T_{min}, and T_{avg}, following hypothesis is proposed: H_0: there is no difference among T_{max}, T_{min}, and T_{avg}; H_1: there is a significant difference among T_{max}, T_{min}, and T_{avg}. A Chi-square goodness-of-fit test is applied on r_{ij} values to test the hypothesis. The results are shown in Table 9. As it is shown, since all p-values > 0.05, so H_1 is rejected at the 95% confidence level. So, there is no significant statically difference among the r_{ij} values of the T_{max}, T_{min}, and T_{avg}.

8. Conclusions

In this study, feed-forward back-propagation neural network (BPNN) model was developed to predict the weekly number of infectious diarrhea of the Shanghai in China. For achieving this goal, an investigation was carried out to prove the potential of using meteorological factors dataset for producing higher accuracy of prediction. Nine meteorological factors were considered. In order to obtain true and effective evaluation of the performance of BPNN model, the same data sets were also trained and tested by the SVR, RFR and MLR models. The results presented in this paper suggested that a BPNN model with architecture 9–4–1 has the best accurate prediction results in prediction of the weekly number of infectious diarrhea. The experiments results indicate that BPNN model shows a good prediction performance as compared to the SVR, RFR and traditional MLR model. Moreover, sensitivity analysis revealed that most effective meteorological factors on the infectious diarrhea are temperature-related variables (T_{max}, T_{min}, and T_{avg}), whereas weekly average rainfall is the least effective parameter on the infectious diarrhea in this study. Considering the above results, it can be concluded that the BPNN model has good capability in predicting infectious diarrhea. Therefore, this technique can be used to predict one-week-ahead infectious diarrhea using meteorological factors in Shanghai of China.

Acknowledgments

This research is supported by the Fund of the Shanghai Science and Technology Development Foundation (Grant No. 13430710100). The authors are grateful to the editor and anonymous reviewers for their suggestions in improving the quality of the paper.

References

[1] World Health Organization, 2013. http://www.who.int/mediacentre/factsheets/fs330/en/index.html (accessed 2013).
[2] C.P. Larson, L. Henning, S. Luby, A.S.G. Faruque, Infectious childhood diarrhea in developing countries, in: Modern Infectious Disease Epidemiology, Springer, NY, 2010, pp. 291–308.
[3] S. Koletzko, S. Osterrieder, Acute infectious diarrhea in children, Dtsch. Arztebl. Int. 106 (33) (2009) 539–547.
[4] I.N. Soyiri, D.D. Reidpath, An overview of health forecasting, Environ. Health Prev. Med. 18 (1) (2013) 1–9.
[5] A.J. McMichael, R. Woodruff, Climate change and infectious diseases, in: K.H. Mayer, H.F. Pizer (Eds.), The Social Ecology of Infectious Diseases, Academic Press, London, 2008, pp. 378–407.
[6] K.D. Lafferty, The ecology of climate change and infectious diseases, Ecology 90 (4) (2009) 888–900.
[7] M. Grasso, M. Manera, A. Chiabai, A. Markandya, The health effects of climate change: a survey of recent quantitative research, Int. J. Environ. Res. Public Health 9 (5) (2012) 1523–1547.
[8] K.A. Alexander, M. Carzolio, D. Goodin, E. Vance, Climate change is likely to worsen the public health threat of diarrheal disease in Botswana, Int. J. Environ. Res. Public Health 10 (4) (2013) 1202–1230.
[9] E.W. Kolstad, K.A. Johansson, Uncertainties associated with quantifying climate change impacts on human health: a case study for diarrhea, Environ. Health Perspect. 119 (3) (2011) 299–305.
[10] N. Zhao, M. XH, G. Lu, Y.Y. Wei, P.Y. Sun, D.H. Zhang, Research on the application of medical–meteorological forecast model of infectious diarrhea disease in Beijing, in: Bio-Inspired Computing: Theories and Applications (BIC-TA), 2010 IEEE Fifth International Conference, 2010, pp. 100–103.
[11] W.C. Chou, J.L. Wu, Y.C. Wang, H. Huang, F.C. Sung, C.Y. Chuang, Modeling the impact of climate variability on diarrhea-associated diseases in Taiwan (1996–2007), Sci. Total Environ. 409 (1) (2010) 43–51.
[12] M. Hashizume, B. Armstrong, S. Hajat, Y. Wagatsuma, A.S. Faruque, T. Hayashi, D.A. Sack, Association between climate variability and hospital visits for non-cholera diarrhoea in Bangladesh: effects and vulnerable groups, Int. J. Epidemiol. 36 (5) (2007) 1030–1037.
[13] .B.J.J. McCormick, W.J. Alonso, M.A. Miller, An exploration of spatial patterns of seasonal diarrheal morbidity in Thailand, Epidemiol. Infect. 140 (7) (2012) 1236–1243.
[14] R.B. Singh, S. Hales, N. de Wet, R. Raj, M. Hearnden, P. Weinstein, The influence of climate variation and change on diarrheal disease in the Pacific Islands, Environ. Health Perspect. 109 (2) (2001) 155–159.
[15] S.J. Lloyd, R.S. Kovats, B.G. Armstrong, Global diarrhea morbidity, weather and climate, Clim. Res. 34 (2) (2007) 119–127.
[16] W.G. Baxt, Application of artificial neural networks to clinical medicine, Lancet 346 (8983) (1995) 1135–1138.
[17] J.F. Xu, X.N. Zhou, Application of artificial neural networks in infectious diseases, Chin. J. Parasitol. Parasit. Dis. 29 (1) (2011) 49–54.
[18] Q.K. Al-Shayea, Artificial neural networks in medical diagnosis, Int. J. Comput. Sci. 8 (2) (2011) 150–154.
[19] H.A. Abbass, An evolutionary artificial neural networks approach for breast cancer diagnosis, Artif. Intell. Med. 25 (3) (2002) 265–281.
[20] S. Kara, F. Dirgenali, A system to diagnose atherosclerosis via wavelet transforms, principal component analysis and artificial neural networks, Expert Syst. Appl. 32 (2) (2007) 632–640.
[21] S. Kia, S. Setayeshi, M. Shamsaei, M. Kia, Computer-aided diagnosis (CAD) of the skin disease based on an intelligent classification of sonogram using neural network, Neural Comput. Appl. 22 (6) (2012) 1049–1062.
[22] Y.X. Ma, S.G. Wang, The application of artificial neural network in the forecasting on incidence of a disease, in: Biomedical Engineering and Informatics (BMEI), 3rd International Conference on IEEE, 2010, pp. 1269–1272.
[23] J.M. Flores-Fernández, E.J. Herrera-López, F. Sánchez-Llamas, Development of an optimized multi-biomarker panel for the detection of lung cancer based on principal component analysis and artificial neural network modeling, Expert Syst. Appl. 39 (12) (2012) 10851–10856.
[24] L.P. Wang, S. Gupta, Neural networks and wavelet de-noising for stock trading and prediction, in: Witold Pedrycz, Shyi Ming Chen (Eds.), Time Series Analysis, Modeling and Applications, Springer, 2013, pp. 229–247.
[25] L.P. Wang, K.K. Teo, Z.P. Lin, Predicting time series with wavelet packet neural networks, in: 2001 IEEE International Joint Conference on Neural Networks (IJCNN 2001), 2001, pp. 1593–1597.
[26] Yu Fang, Kamaladdin Fataliyev, Lipo Wang, Xiuju Fu, Yaoli Wang, Improving the genetic-algorithm-optimized wavelet neural network approach to stock market prediction, in: 2014 International Joint Conference on Neural Networks, 2014.

[27] M. Zhu, L.P. Wang, Intelligent trading using support vector regression and multilayer perceptrons optimized with genetic algorithms, in: The 2010 International Joint Conference on Neural Networks (IJCNN), 2010, pp. 1–5.
[28] L. Zhang, K. Wu, Y. Zhong, P. Li, A new sub-pixel mapping algorithm based on a BP neural network with an observation model, Neurocomputing 71 (2008) 2046–2054.
[29] S. Asadi, J. Shahrabi, P. Abbaszadeh, S. Tabanmehr, A new hybrid artificial neural networks for rainfall-runoff process modeling, Neurocomputing 121 (2013) 470–480.
[30] W.P. Ma, X.D. Sun, Y.Y. Song, F.F. Tao, W. Feng, Y. He, N.Q. Zhao, Z.A. Yuan, Applied mixed generalized additive model to assess the effect of temperature on the incidence of bacillary dysentery and its forecast, PLOS ONE 8 (4) (2013) e62122.
[31] E. Doğan, A.P. Akgüngör, Forecasting highway casualties under the effect of railway development policy in Turkey using artificial neural networks, Neural Comput. Appl. 22 (5) (2011) 869–877.
[32] V. Kecman, Learning and Soft Computing, Support Vector machines, Neural Networks and Fuzzy Logic Models, The MIT Press, Cambridge, MA, 2001.
[33] L.P. Wang, X.J. Fu, Data Mining with Computational Intelligence, Springer, Berlin, 2005.
[34] L.P. Wang (Ed.), Support Vector Machines: Theory and Application, Springer, Berlin, 2005.
[35] M.A. Shahin, M.B. Jaksa, H.R. Maier, State of the art of artificial neural networks in geotechnical engineering, Electron. J. Geotech. Eng. 8 (2008) 1–26.
[36] R. Hecht-Nielsen, Kolmogorov's mapping neural network existence theorem, in: Proceedings of the First IEEE International Conference on Neural Networks, San Diego, CA, 1987, pp. 11–14.
[37] J.J. Moré, The Levenberg–Marquardt algorithm: implementation and theory, in: Numerical Analysis, Springer, Berlin, Heidelberg, 1978, pp. 105–116.
[38] Guoqiang Zhang, B. Eddy Patuwo, Michael Y. Hu, Forecasting with artificial neural networks: the state of the art, Int. J. Forecast. 14 (1998) 35–62.
[39] S.M.M. Vaghefi, S.M.M. Vaghefi, Prediction of phosphorus content of electroless nickel–phosphorous coatings using artificial neural network modeling, Neural Comput. Appl. 20 (2011) 1055–1060.
[40] C.Y. Yeh, C.W. Huang, S.J. Lee, A multiple-kernel support vector regression approach for stock market price forecasting, Expert Syst. Appl. 38 (2011) 2177–2186.
[41] L. Breiman, Mach. Learn 45 (2001) 5.
[42] G. Li, J. Shi, On comparing three artificial neural networks for wind speed forecasting, Appl. Energy 87 (7) (2010) 2313–2320.
[43] B. Wah, M. Qian, Time-series predictions using constrained formulations for neural-network training and cross validation, in: Proceedings of the International Conference on Intelligent Information Processing, 16th IFIP World Computer Congress, 2000, pp. 220–226.
[44] M. Smith, Neural Networks for Modeling, Van Nostrand Reinhold, NY, 1993.
[45] P.H.S.W. Kulatilake, Qiong Wu, T. Hudaverdi, C. Kuzu, Mean particle size prediction in rock blast fragmentation using neural networks, Eng. Geol. 114 (2010) 298–311.
[46] V.M. Janakiraman, XuanLong Nguyen, Dennis Assanis, Nonlinear identification of a gasoline HCCI engine using neural networks coupled with principal component analysis, Appl. Soft Comput. 13 (2013) 2375–2389.
[47] P. Hosseinzadeh Talaee, Multilayer perceptron with different training algorithms for streamflow forecasting, Neural Comput. Appl. 24 (3–4) (2014) 695–703.
[49] D.C. Moosmayer, Alain Yee-Loong Chong, J.L. Martin, B. Schuppar, A neural network approach to predicting price negotiation outcomes in business-to-business contexts, Expert Syst. Appl. 40 (2013) 3028–3035.
[50] M. Smith, Neural Networks for Statistical Modeling, Wiley, NY, 1993.
[51] S.K. Jain, P.C. Nayak, K.P. Sudheer, Models for estimating evapotranspiration using artificial neural networks, and their physical interpretation, Hydrol. Process. 22 (13) (2008) 2225–2234.
[52] S. Mandal, P.V. Sivaprasad, S. Venugopal, K.P.N. Murthy, Artificial neural network modeling to evaluate and predict the deformation behavior of stainless steel type AISI 304L during hot torsion, Appl. Soft Comput. 9 (2009) 237–244.
[54] Ahmad Taher Azar, Fast neural network learning algorithms for medical applications, Neural Comput. Appl. 23 (3–4) (2013) 1019–1034.
[55] R.A. Jacobs, Increased rates of convergence through learning rate adaptation, Neural Netw. 1 (1988) 295–307.
[56] S. Pandey, D.A. Hindoliya, R. Mod, Artificial neural networks for predicting indoor temperature using roof passive cooling techniques in buildings in different climatic conditions, Appl. Soft Comput. 12 (3) (2012) 1214–1226.
[57] Ruan Junhu, Xuping Wang, Yan Shi, Developing fast predictors for large-scale time series using fuzzy granular support vector machines, Appl. Soft Comput. 13 (9) (2013) 3981–4000.
[58] T. Hastie, R. Tibshirani, J. Friedmanm, The Elements of Statistical Learning: Data Mining, Inferences and Prediction, Springer, NY, 2009;
N. Goudarzi, D. Shahsavani, F. Emadi-Gandaghi, et al., Application of random forests method to predict the retention indices of some polycyclic aromatic hydrocarbons, J. Chromatogr. A 1333 (2014) 25–31.
[59] Markus Brameier, Wolfgang Banzhaf, A comparison of linear genetic programming and neural networks in medical data mining, IEEE Trans. Evolut. Comput. 5 (1) (2001) 17–26.
[60] B. Acherjee, S. Mondal, B. Tudu, D. Misra, Application of artificial neural network for predicting weld quality in laser transmission welding of thermoplastics, Appl. Soft Comput. 11 (2) (2011) 2548–2555.
[61] M. Monjezi, Z. Ahmadi, A.Y. Varjani, M. Khandelwal, Backbreak prediction in the Chadormalu iron mine using artificial neural network, Neural Comput. Appl. 23 (3–4) (2013) 1101–1107.
[62] Y.H. Jong, C.I. Lee, Influence of geological conditions on the powder factor for tunnel blasting, Int. J. Rock Mech. Min. Sci. 41 (3) (2004) 533–538.
[63] A. Salsani, J. Daneshian, S. Shariati, A. Yazdani-Chamzini, M. Taheri, Predicting roadheader performance by using artificial neural network, Neural Comput. Appl. 24 (2014) 1823–1831.
[64] Mohammad Zounemat-kermani, Ozgur Kisi, Taher Rajaee, Performance of radial basis and LM-feed forward artificial neural networks for predicting daily watershed runoff, Appl. Soft Comput. 13 (2013) 4633–4644.