Knowledge-Based Systems 88 (2015) 12–23

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

Diarrhoea outpatient visits prediction based on time series decomposition and multi-local predictor fusion

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

[a] Department of Computer Science and Technology, East China Normal University, Shanghai 200241, PR China
[b] College of Physics and Engineering, Qufu Normal University, Qufu 273165, PR China

ARTICLE INFO

Article history:
Received 24 December 2014
Received in revised form 2 August 2015
Accepted 4 August 2015
Available online 25 August 2015

Keywords:
Diarrhoea outpatient visits
Prediction model
Time series decomposition
Ensemble empirical mode decomposition
Generalized regression neural networks
Multi-predictor fusion

ABSTRACT

Accurate and reliable prediction of diarrhoea outpatient visits is necessary for the health authorities to ensure the appropriate action for the control of the outbreak. In this study, a novel method based on time series decomposition and multi-local predictor fusion has been proposed to predict the diarrhoea outpatient visits. For time series decomposition, the Ensemble Empirical Mode Decomposition with Adaptive Noise (EEMDAN) is used to decompose diarrhoea outpatient visits time series into a finite set of Intrinsic Mode Function (IMF) components and a residue. The IMF components and residue are modeled and predicted respectively by means of Generalized Regression Neural Network (GRNN) as local predictor. Then the prediction results of all components are fusioned using another independent GRNN as fusion predictor to obtain final prediction results. This is the first study on using a EEMDAN and GRNN to constructing an prediction model for diarrhoea outpatient visits prediction problems. The pre-procession and post-processing techniques are used to take into account the seasonal and trend effects in the datasets for improving the prediction precision of proposed model. The performance of the proposed EEMDAN–GRNN model has been compared with Seasonal Auto-Regressive Moving Average (SARIMA), Single GRNN, Wavelet-GRNN and also with EEMD–GRNN by applying them to predict four real world diarrhoea outpatient visits. The results indicate that the proposed EEMDAN–GRNN model provides more accurate prediction results compared to the other traditional techniques. Thus EEMDAN–GRNN can be an alternate tool to facilitate the prediction of diarrhoea outpatient visits.

1. Introduction

Diarrhoea is the passage of three or more loose or liquid stools per day, or more frequently than is normal for the individual [1]. Diarrhoeal caused by a variety of bacterial, viral and parasitic organisms is usually a symptom of gastrointestinal infection. Infection is spread through contaminated food or drinking-water, or from person to person as a result of poor hygiene [1]. Diarrhoeal disease is a serious threat to the health and well-being of the citizens of the world, especially in the developing countries like China, African countries and India. Thousands of people, in particular for children less than five years old, suffer from this disease every year. In spite of many studies on the diarrhoea still there were nearly 1.7 billion cases of diarrhoeal disease every year [1]. Diarrhoeal disease is the second leading cause of death in children under five years old, and is responsible for killing around 760,000 children every year. Government authorities are incurring huge cost to control and eliminate the outbreaks of diarrhoea. Thus, accurately and timely predict diarrhoea outpatient visits in advance outbreaks is an important issue for various national governments and international organizations. 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 [2].

However, many problems have occurred in prevention and control of diarrhoeal disease. One of the main problems associated with prediction of diarrhoeal disease is the complexity and diversity of influence factors that affect the diarrhoeal incidences, such as malnutrition, meteorological, living surroundings, and living habits. Especially due to global warming, rapid climate changes are occurring which result in the increase of diarrhoeal disease incidence depending upon the specific micro-climate of that particular region [56–58]. The complexity and diversity of the influence factors make it is great challenge for the researchers to predict the diarrhoeal disease outbreaks in advance. In the absence of knowledge about probabilistic attack of these diseases and exogenous factors are often limited by the availability of data,

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

http://dx.doi.org/10.1016/j.knosys.2015.08.013
0950-7051/

government fails to provide adequate treatment facility on time. Thus, it is necessary to forecast the occurrence of these diseases in advance so that its devastating impact on the society can be reduced.

There have been wide attempts to capture the relationship between the available information using some straightforward linear regression models like Auto-Regressive Integrated Moving Average (ARIMA) [3–10]. Such traditional techniques do require minimum computational efforts to set up prediction models which are considered to be an advantage. However, with non-linear nature of diarrhoeal disease it becomes difficult to use these techniques. In recent years, many recent studies focus on the use of machine learning techniques, such as Artificial Neural Networks (ANNs), to build a prediction model for time series prediction problems. Unlike traditional statistical models, ANNs are data-driven and non-parametric models. They do not require strong model assumptions and can map any nonlinear function without a priori assumption about the properties of the data, even though the underlying relationships are unknown or hard to describe. Related works have shown that machine learning techniques outperform many traditional models [11–15].

Due to the nonlinear and non-stationary characteristic, accurate prediction diarrhoea outpatient visits by the building a single global predictor is often not possible. One of the best ways to solve these problems is using decomposition-and-ensemble principle. Firstly, a complex problem is decomposed into a set of sub-problems according to the inherent class relations among training data, then gives a local predictor to learn each of the sub-problems, finally, combination the multi-local predictors into a solution to the original problem [16–19]. In the area of time series forecasting (TSF), decomposition-and-ensemble principle has proven to be a method superior to single global predictor. Many recent studies in the different area have been shown to perform better than single models [20–28]. However, so far as I know, this is the first study on using a decomposition-and-ensemble principle to constructing a prediction model for diarrhoea outpatient visits prediction problems in health forecasting.

In order to decompose a complex time series prediction task into several relative simple subtasks, time series decomposition methodologies have been widely used in different studies. These techniques can divide the data into local characteristic time scale and extracting meaningful features embedded implicitly in data. The Empirical Mode Decomposition (EMD) [34] and its improved version named Ensemble Empirical Mode Decomposition (EEMD) [35], have been widely used as a promising alternative for nonlinear and non-stationary time series and successfully applied to different areas [22–24,29–33]. In this paper, a variation of the EEMD algorithm, called Ensemble Empirical Mode decomposition with Adaptive Noise (EEMDAN) [36] is used for the time series decomposition. The EEMDAN provides an exact reconstruction of the original signal and a better spectral separation of the modes, with a lower computational cost.

Among ANN models, multilayer perceptron (MLP) trained by the standard Back-Propagation (BP) learning algorithm is popular ANNs for predicting time series. Despite the advantages of MLP, they have several weaknesses (e.g., a large number of design parameters, long training time, and suffering from local minima) which make modeling more difficult [37]. In this paper, we attempt to develop local predictors using generalized regression neural networks (GRNN), a special type of ANNs. GRNN has only a single design parameter and is simple and fast in training. Our effort in this paper focuses on designing a modeling scheme to take full advantage of EEMDAN and GRNN properties for diarrhoea outpatient visits prediction. Therefore, a novel prediction algorithm call EEMDAN–GRNN is proposed. To increase the accuracy of the prediction, we perform data pre-processing techniques such as data transformation, detrending and deseasonalizing. In order to improve the robustness and error tolerance of proposed model, a trainable fusion method (another independent GRNN predictor) is used to fusion the prediction results of multi-local predictors. The strength of the proposed prediction method is tested on four real word monthly diarrhoea outpatient visits time series datasets from three different geographical location of China.

In summary, the primary innovation and contributions of our study can be outlined as follows:

(1) Based on a literature review, there is no works has been carried out to utilize the decomposition-and-ensemble principle method in predicting diarrhoea outpatient visits. In this study, following the decomposition-and-ensemble principle, a novel model based on time series decomposition and multi-local predictor fusion has been proposed to predict the diarrhoea outpatient visits.
(2) A variation of the EEMD algorithm, called EEMDAN is used for the diarrhoea outpatient visits time series decomposition. The GRNN is used as local predictor. So far, EEMDAN and GRNN have not been used in this direction. The proposed EEMDAN–GRNN algorithm adequately makes use of the advantages of the EEMDAN decomposition method and GRNN and integrates them well, which conduce to boosting the model prediction ability and enhancing prediction efficiency.
(3) Our proposed EEMDAN–GRNN algorithm uses a dynamic nonlinear weighted scheme to fusion the multi-local predictor into a single predictor. Each local GRNN predictor independently predicts the output. A fusion predictor is then trained to predict the final output from the outputs of local predictor. Consequently, fusion predictor can capture interactions among local predictors.

The remainder of this paper is organized as follow. The methodologies that are used in this study are briefly described in Section 2. The proposed EEMDAN–GRNN modeling framework is presented in detail in Section 3. Section 4 illustrates the experimental design and methodologies implementation in details. Following that, in Section 5, the experimental results obtained from four real diarrhoea outpatient visits datasets are presented and discussed. Finally, the study is concluded in Section 6.

2. Methodologies

Before starting to present the proposed method, it is necessary to describe the theory of the acquired methodologies in the proposed approach. In this section, the decomposition technique of EEMDAN and the theory of GRNN algorithm are briefly introduced.

2.1. EEMD with adaptive noise

The EMD [34] is an adaptive signal processing technique introduced to analyze non-linear and non-stationary time series. It consists in a local and fully data-driven separation of a time series in fast and slow oscillations. However, EMD experiences some problems, such as the presence of oscillations of very disparate amplitude in a mode, or the presence of very similar oscillations in different modes, named as mode mixing [36]. To overcome these problems, the EEMD method was proposed [35]. It performs the EMD over an ensemble of the signal plus Gaussian white noise. The addition of white Gaussian noise solves the mode mixing problem by populating the whole time–frequency space to take advantage of the dyadic filter bank behavior of the EMD [38]. However it creates some new ones. Indeed, the reconstructed signal includes

residual noise and different realizations of signal plus noise may produce different number of modes. In order to overcome these situations, Torres et al. [36] proposed a variation of the EEMD algorithm that provides an exact reconstruction of the original signal and a better spectral separation of the modes, with a lower computational cost. The improved EEMD with adaptive noise (EEMDAN) is depicted by the following algorithm [36]:

Step 1. Decompose by EMD realizations $x[n]+\varepsilon_0\omega^i[n]$ $(i=1,2,\ldots,I)$ to obtain their first modes and compute $\widehat{IMF}_1[n]=\frac{1}{I}\sum_{i=1}^{I}IMF_1^i[n]=\overleftrightarrow{IMF}_1[n]$;
Step 2. At the first stage ($k=1$) calculate the first residue: $r_1[n]=x[n]-\widehat{IMF}_1[n]$;
Step 3. Decompose realizations $r_1[n]+\varepsilon_1E_1(\omega^i[n])(i=1,2,\ldots,I)$, until their first EMD mode and define the second mode:

$$\widehat{IMF}_2[n]=\frac{1}{I}\sum_{i=1}^{I}E_1(r_1[n]+\varepsilon_1E_1(\omega^i[n]))$$

Step 4. For $k=2,3,\ldots,k$ calculate the kth residue: $r_k[n]=r_{(k-1)}[n]-\widehat{IMF}_k[n]$;
Step 5. Decompose realizations $r_k[n]+\varepsilon_kE_1(\omega^i[n])(i=1,2,\ldots,I)$, until their first EMD mode and define the $(k+1)$th mode as:

$$\widehat{IMF}_{(k+1)}[n]=\frac{1}{I}\sum_{i=1}^{I}E_1(r_k[n]+\varepsilon_kE_k(\omega^i[n]))$$

Step 6. Go to step 4 for next k.
Steps 4–6 are performed until the obtained residue is no longer feasible to be decomposed (the residue does not have at least two extrema). The final residue satisfies:

$$R[n]=x[n]-\sum_{k=1}^{K}\widehat{IMF}_k$$

With K is the total number of modes. Therefore, the given signal $x[n]$ can be expressed as:

$$x[n]=\sum_{k=1}^{K}\widehat{IMF}_k+R[n] \tag{1}$$

Eq. (1) makes the proposed decomposition complete and provides an exact reconstruction of the original data.

Where operator $E_j(\cdot)$ produces the jth mode obtained by EMD, the w^i is white noise with $N(0,1)$, $x[n]$ is the targeted data and the decomposition modes will be noted as $\widehat{IMF}_k$ for EEMDAN and $\overleftrightarrow{IMF}$ for EEMD.

2.2. Generalized regression neural network

Generalized Regression Neural Network (GRNN), developed by [39], is a kind of radial basis function (RBF) networks which is based on a standard statistical technique called kernel regression [26]. A typical GRNN is organized using four layers, namely the input layer, the pattern layer (radial basis layer), the summation layer, and the output layer. The hidden layer has radial basis neurons, while neurons in the output layer have a linear transfer function. A typical schematic diagram of the GRNN architecture is presented in Fig. 1. Given a sufficient number of neurons, GRNN can approximate a continuous function to an arbitrary accuracy [39].

As a standard regression technique, GRNN is used for the estimation of continuous variables. It is related to the radial basis function network and is based on a standard statistical technique called kernel regression. By definition, the regression of a dependent variable y on an independent $\boldsymbol{x}$, estimates the most probable value for y, given $\boldsymbol{x}$ and a training set. The regression method will produce the estimated value of y which minimizes the Mean Squared Error (MSE). The GRNN is a method for estimating the joint probability density function of $\boldsymbol{x}$ and y, given only a training dataset. Because the probability density function is derived from the data with no preconceptions about its form, the system is perfectly general.

If $f(\boldsymbol{x},y)$ represents the known joint continuous probability density function of a vector random variable $\boldsymbol{x}$, and a scalar random variable y, the conditional mean of y given $\boldsymbol{x}$, also called the regression of y on $\boldsymbol{x}$, is given by:

$$E[y|\boldsymbol{x}]=\frac{\int_{-\infty}^{\infty}yf(y,\boldsymbol{x})dy}{\int_{-\infty}^{\infty}f(y,\boldsymbol{x})dy} \tag{2}$$

where $\boldsymbol{x}=[\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_N]^T$ is input and $\boldsymbol{x}_i=(x_{i1},x_{i2},\ldots,x_{id})$ is a d-dimensional input vector; $y=[y_1,y_2,\ldots,y_N]^T$ is target output; $E[y|\boldsymbol{x}]$ is the expected value of the output y, given the input vector $\boldsymbol{x}$, $f(y,\boldsymbol{x})$ is the joint probability density function of $\boldsymbol{x}$ and y.

When the density $f(y,\boldsymbol{x})$, is not known, it must usually be estimated from a sample of observations of x and y. Given N input–output pairs $\{\boldsymbol{x}_i,y_i\}_{i=1}^{N}\in\Re^d\times\Re^1$ and as the training samples, assume the original design of GRNN, that is, the number of hidden neurons is equal to the number of training samples. For a desired estimate of system output vectors y, under the input vectors $\boldsymbol{x}$, is achieved by a regression calculation $\bar{y}_i=\hat{f}(\boldsymbol{x}_i)$, where

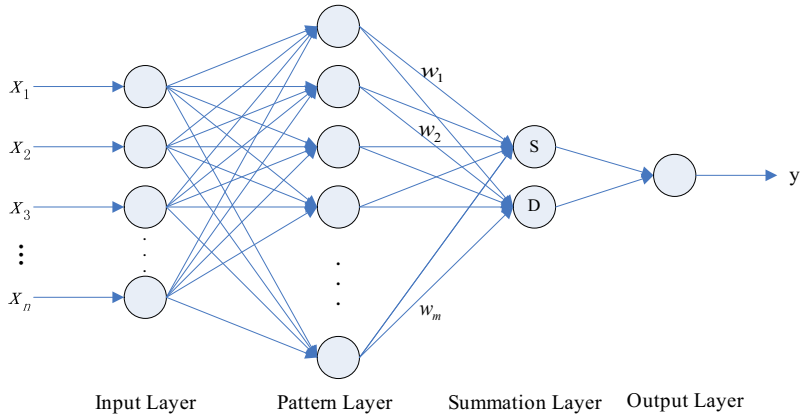

Fig. 1. Typical GRNN structure.

$f(\cdot): \Re^d \to \Re$ is the predictor function. When probability density function adopt Gaussian function, the procedure of the GRNN model can be represented as:

$$\hat{f}(\boldsymbol{x}_i) = \frac{\sum_{i=1}^{N} y_i \exp\left(-D_i^2/2\sigma^2\right)}{\sum_{i=1}^{N} \exp\left(-D_i^2/2\sigma^2\right)} \tag{3}$$

where D_i^2 is defined as $D_i^2 = (\boldsymbol{x} - \boldsymbol{x}_i)^T(\boldsymbol{x} - \boldsymbol{x}_i)$, σ denotes the smoothing parameter, $\boldsymbol{x}$ is the input variable of the network, $\boldsymbol{x}_i$ is a specific training vector of the neuron i in the pattern layer.

A good performance of GRNN method depends on smoothing/spread factor σ (in Eq. (3) above), which is very important in using GRNN for prediction and determines the generalization capability of the GRNN. The smoothing factor (SF) is only free parameter, apart from the input and output layer, involved in the designing of the network. The recommended value for smoothing factor is smaller than the typical distance between input vectors. Smoothing factor is smaller and the function approximation capability will be stronger, while smoothing factor becomes larger, the performance will tend to be smoother [26].

As discussed in Section one, MLP networks have several shortcomings that make design of MLP a difficult task. GRNN, on the other hand, have several advantages [37], including: (1) it has one design parameter (i.e., smoothing factor); (2) it is easy to train since it is a one-pass algorithm; (3) it can accurately approximate functions from sparse and noisy data; (4) it can converge to the conditional mean surface by increasing the number of data samples; and (5) ability to model from a relatively small data set, and ability to handle outliers. It is these unique advantages that make us to choose GRNN in our ANN modeling scheme as predictor. Some recent papers [40–43] employed GRNN to TSF and concluded that GRNN have a higher degree of prediction accuracy than MLP neural networks.

3. The proposed EEMDAN–GRNN algorithm

Considering the aforementioned points in Section one, in the current research, the powerful combination of positive aspects of EEMDAN and GRNN algorithm is presented to one-step-ahead diarrhoea outpatient visits time series prediction problem. The research scheme of the proposed EEMDAN–GRNN algorithm is presented in Fig. 2. As shown in Fig. 2, the proposed algorithm can be separated into five steps: (1) the data pre-processing; (2) the time series decomposition; (3) the local predictor construction; (4) the multi-local predictor fusion and (5) prediction results post-processing. The detailed illustration of each step in the proposed algorithm is described as follows:

Step 1: Data pre-processing.

Kotsiantis et al. [44], Azadeh et al. [45] and Kuvulmaz et al. [46] pointed out that input data pre-processing has significant effects for improving prediction performance of supervised learning models. Besides the conventional rescaling or normalization of data, pre-processing methods such as detrending and deseasonalization were used in this paper. A pre-processing method should contain the capability of transforming pre-processed data into its original scale (called post-processing). Since GRNN perform prediction based on the similarity of the input point to the historical data points in the input space, GRNN are inherently ineffective in modeling trend [37]. Thus effective detrending is more important for GRNN than other ANN models. The same conclusions were also drawn for highly seasonal time series that deseasonalization can significantly improve prediction accuracy [26]. For data normalization, there are different normalization algorithms, such as Min–Max normalization, Z-score normalization and sigmoid normalization.

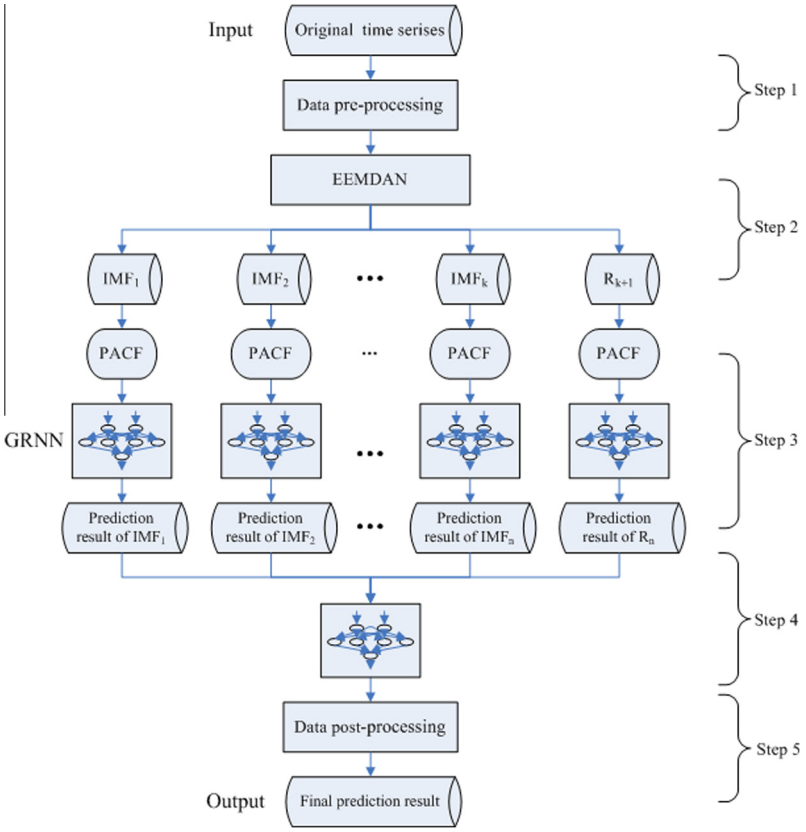

Fig. 2. The proposed EEMDAN–GRNN algorithm framework.

In this paper we use Min–Max normalization (Eq. (4)) which is a common approach in this field.

$$y^*_{new} = \left(\frac{y_{old} - y_{\min}}{y_{\max} - y_{\min}}\right)(y^*_{\max} - y^*_{\min}) + y^*_{\min} \tag{4}$$

where y^*_{new} is the normalized value, y_{old}, $y_{\max}$, $y_{\min}$ are the original, maximum and minimum values of the raw data, respectively and $y^*_{\max}$, $y^*_{\min}$, are the maximum and minimum of the normalized data, respectively. Given an original one-dimensional time series $Y_t = \{y_t\}_{t=1}^N \in \Re^1$, the output of this step is a one-dimensional time series $Y^*_t = \{y^*_t\}_{t=1}^N \in \Re^1$, where N represents the length of the series.

Step 2: Decompose time series using EEMDAN.

In this Step, the EEMAN is used to decomposes the pre-processed time series data into a finite and often a small number of intrinsic mode functions (IMFs) and plus a residue.[1] Each IMF component can represent the local characteristic time scale by itself. Then, these subseries can be predicted more accurately. The output of this step is k IMFs and a residue and k is the number of IMFs:

$$Y^*_t = \begin{bmatrix} IMF_{1,t} \\ IMF_{2,t} \\ IMF_{3,t} \\ \dots \\ IMF_{k,t} \\ R_{k+1,t} \end{bmatrix} = \begin{bmatrix} x_{1,1} & x_{1,2} & x_{1,3} & \dots & x_{1,N} \\ x_{2,1} & x_{2,2} & x_{2,3} & \dots & x_{2,N} \\ x_{3,1} & x_{3,2} & x_{3,3} & \dots & x_{3,N} \\ \dots & \dots & \dots & \dots & \dots \\ x_{k,1} & x_{k,2} & x_{k,3} & \dots & x_{k,N} \\ x_{k+1,1} & x_{k+1,2} & x_{k+1,3} & \dots & x_{k+1,N} \end{bmatrix} \tag{5}$$

Step 3: Local predictor construction.

After the components (IMFs and a residue) are adaptively extracted via EEMDAN, each component is modeled by an independent GRNN model which is used to generate local predictor to predict the component series respectively. So, for each IMF_i (ith IMF series), following three sub-steps were executed:

(1) Input selection for IMF_i

Several studies, for example [47–49], have indicated that selecting model input lags is probably the most critical task for a time series prediction model, since it contains important information embedded in the data. The statistical approach to examine auto- and partial-auto-correlation of the observed time series was recognized as a good and parsimonious method in the determination of model inputs [50,51]. So, in this study, the model inputs are mainly determined by the plot of partial-auto-correlation function (PACF).

(2) Construction of input/output pairs for IMF_i

In this sub-stage, a single-variable IMF_i time series is embedded from the original one-dimensional space into a p-dimensional reconstruction space, which indicates the system state at different time. Let $IMF_{it} = \{x_{i1}, x_{i2}, \dots, x_{iN}\} \in \Re(i = 1, 2, \dots, k)$ stand for ith IMF series, we can reconstructed the original IMF_i series into an input–output pair format:

$$\begin{Bmatrix} x_{1i} \\ x_{2i} \\ x_{3i} \\ \dots \\ x_{Ni} \end{Bmatrix}^T \rightarrow \underbrace{\begin{Bmatrix} x_{i,1} & x_{i,2} & x_{i,3} & \dots & x_{i,p} \\ x_{i,2} & x_{i,3} & x_{i,4} & \dots & x_{i,p+1} \\ x_{i,3} & x_{i,4} & x_{i,5} & \dots & x_{i,p+2} \\ \dots & \dots & \dots & \dots & \dots \\ x_{i,N-p+1} & x_{i,N-p} & x_{i,N-p-1} & \dots & x_{i,N-1} \end{Bmatrix}}_{input} \underbrace{\begin{Bmatrix} x_{i,p+1} \\ x_{i,p+2} \\ x_{i,p+3} \\ \dots \\ x_{i,N} \end{Bmatrix}}_{output} \tag{6}$$

where p is a positive integer often referred to as the embedded dimension of the model and $N - p + 1$ stands for the number of input/output pairs. The $N - p + 1$ new training dataset is used to construction ith local predictor for ith IMF.

(3) Train local predictors for IMF_i

The main problem for local predictor modeling for IMF_i is constructing such a function with satisfactory prediction accuracy, $f_{local}(x_{input}, x_{output}) : \Re^p \rightarrow \Re^1$, with the following form:

$$x_{out} = f_{local}(x_{input}, \phi) + \varepsilon_t \tag{7}$$

where local-predictors $f_{local}(\cdot)$ are trained and defined using Eq. (3) in different regions of the input space, ϕ is a parameter vector to be determined, and ε_t denotes noise usually regarded as Gaussian white noise independent of the previous observations.

Through the implementation of Step 3, $k + 1$ prediction results can be obtained and organized into a $k + 1$-dimensional vector $z = \{\bar{x}_{ti}\} \in \Re^{k+1}(i = 1, 2, \dots, k + 1, t = 1, 2, \dots, N)$.

Step 4: Muti-local predictor fusion.

To capture the correlation of local predictors more robust and flexibly, in this Step, trainable fusion method can be adopted in our work. Based on results of Step 3 and actual pre-processed data, a new training dataset $Z = \{\bar{x}_{ti}, y^*_t\}$ is formed. Based on Z, a fusion predictor is trained, in which GRNN can be adopted as well:

$$\bar{y}_t = f_{fusion}(f_{local\_j}(x_{t-1}, x_{t-2,} \dots x_{t-p_j}), \phi) + \varepsilon_t, \ (j = 1, 2, \dots, k + 1) \tag{8}$$

where $f_{fusion}(\cdot)$ is a fusion predictor; $\bar{y}_t$ is prediction results for pre-processed data and p_j denotes model inputs for jth local predictor. In trainable fusion model, the correlation of local predictors is depicted by function $f_{fusion}(\cdot)$, which is not based on any assumption such as linear restriction. As a result, the local predictor can be combined in a better-organized way according to data's intrinsic distribution.

Step 5: Post-processing data.

Post-processing (rolling back of the preprocessing step performed, such as deseasonalization, detrending and returned to original scale) of the estimated results is done in this step and the final prediction value of the raw diarrhoea outpatient visits data can be obtained.

4. Experimentation design

In this section, the details about the experimental design will be presented. In Section 4.1, the briefly describes of the datasets are given in our experimentations. Section 4.2 describe the experimental setting and implementations of proposed EEMDAN–GRNN algorithm in detail. Section 4.3 presents the selected counterparts for comparison and the implementations of counterpart models. Our experiments are conducted on four real word diarrhoea outpatient visits datasets. The hardware used is an AMD 2.20 GHz CPU with 4 GB memory. Programs were written in Matlab (R2013b, The Math works, Inc., Natick, MA, USA) and run using Windows 7.

4.1. Experimental datasets

To evaluate the performance of proposed prediction algorithm EEMDAN–GRNN, four real-world monthly numbers of diarrhoea outpatient visits datasets were collected from the Shanghai Municipal Center for Disease Control & Prevention (SCDC) and National Disease Supervision Information Management System (NDSIMS), respectively. A brief summary of the four data sets is given in Table 1. The temporal variation behavior of four datasets are illus-

[1] In this paper, the residual also be considered as an IMF.

Table 1
Diarrhoea outpatient visits time series dataset.

Dataset	Seasonal	Trend	Size	Periods	Location	Source
SH_Chindren	12	Yes	84	2006.1–2012.12	Shanghai	SCDC
SH_Adult	12	Yes	84	2006.1–2012.12	Shanghai	SCDC
BJ	12	Yes	108	2004.1–2012.12	Beijing	NDSIMS
GD	12	Yes	108	2004.1–2012.12	Guangdong	NDSIMS

trate in Fig. 3. It is clear that four diarrhoea outpatient visits time series exhibit upward increasing trend pattern together with seasonal pattern as shown in Fig. 3.

4.2. Experimental setting

To test the algorithms with different training and testing periods, all methods were assessed using 5-fold time ordered cross-validation [55] procedure for each dataset in this study. The implementation of the 5-fold time ordered cross-validation was carried out by splitting (using time order) the original dataset into five equal-sized subsets. Any four of the five subsets as training datasets are selected to perform training. The remaining part as testing dataset will be executed to test the generalization performance of model. The structure of data sets using 5-fold cross-validation is shown in Fig. 4.

As a result, each part will be trained and tested five times. The values reported are averages of the cross-validation estimates over these different testing datasets. The root mean squares error (RMSE) metric is calculated on the testing subsets in order to evaluate the generalization capabilities of the tested methods.

According to the proposed EEMDAN–GRNN approach, in Stage 1, the original datasets is first scaled into the range of [0,1] using Eq. (4). Once the final prediction results are obtained using scaled data, inverse transformation is carried out to obtain final prediction results. Since the diarrhoea outpatient visits time series exhibit strong seasonal component and trend pattern (see Fig. 3), we conduct deseasonalizing by means of the multiplicative seasonal decomposition using SPSS.19 statistical software. In addition, detrending is performed by the 13-term moving average and then dividing the estimated trend from the series.

In fact, there are three types of parameters to be set for our EEMDAN–GRNN algorithm (Input lag p for each IFM, smoothing

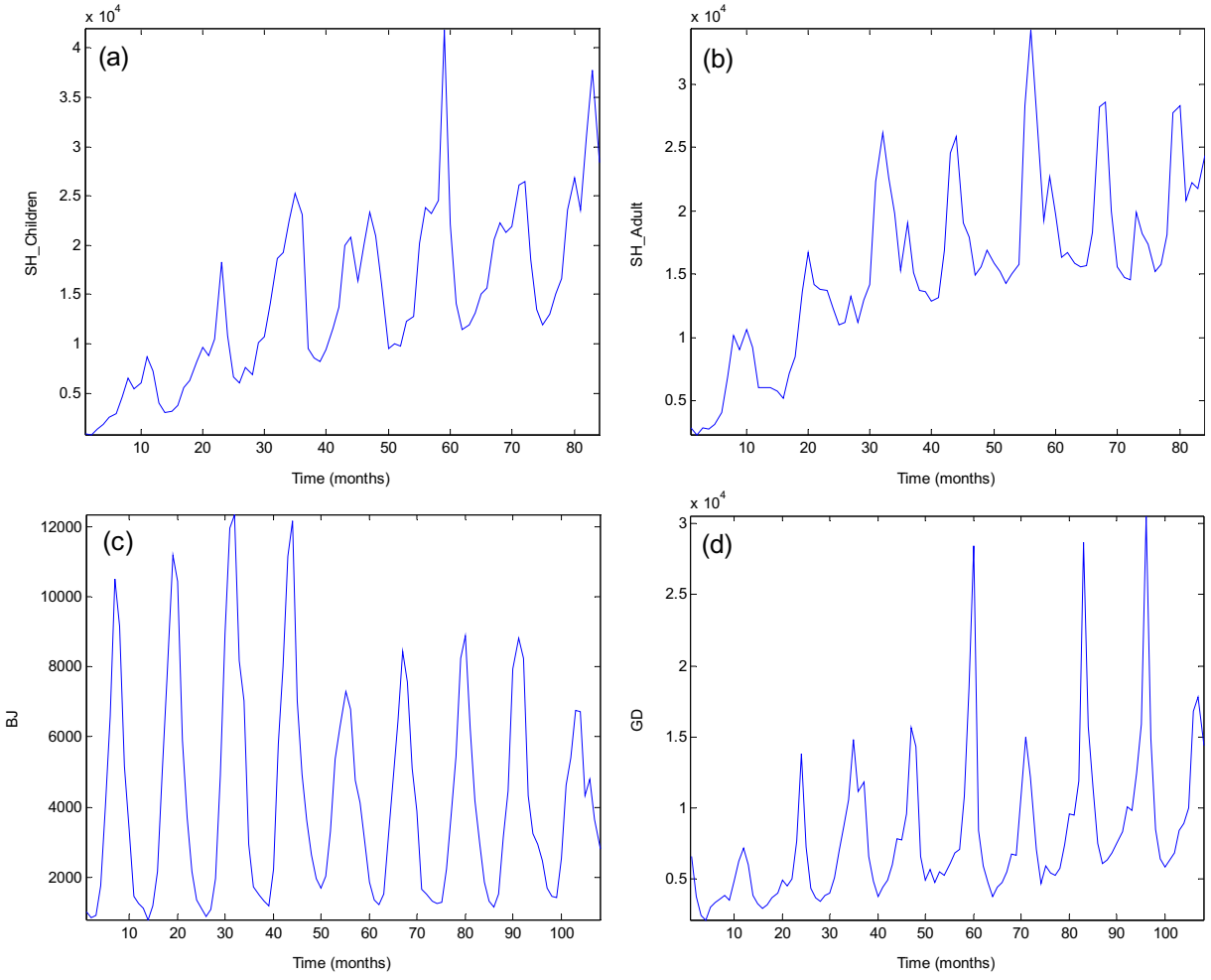

Fig. 3. Monthly diarrhoea outpatient visits. (a) SH_Children; (b) SH_Adult; (c) BJ; (d) GD.

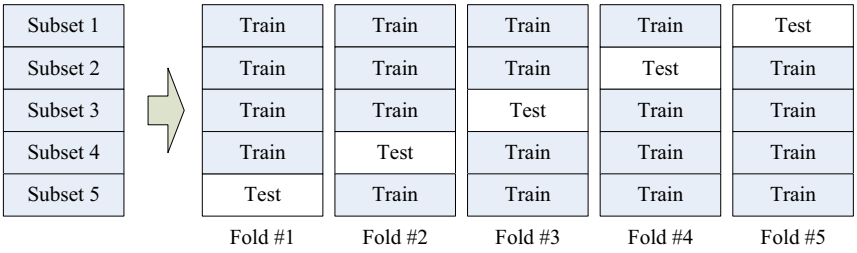

Fig. 4. Schematic illustration of the data partitioning for 5-fold time ordered cross-validation.

Table 2
Values of input lags and smoothing factors provide the average cross-validation errors in the different datasets in developing local GRNN predictors for each IMF.

IMF	SH_Children		SH_Adult		BJ		GD	
	Lag	SF	Lag	SF	Lag	SF	Lag	SF
IMF_1	2	0.04	2	0.02	2	0.01	2	0.05
IMF_2	4	0.02	4	0.01	4	0.02	2	0.01
IMF_3	2	0.01	3	0.01	3	0.01	3	0.01
IMF_4	3	0.01	2	0.01	3	0.01	3	0.05
IMF_5	2	0.01	3	0.01	4	0.01	3	0.01
IMF_6	1	0.01	2	0.01	5	0.01	4	0.03
IMF_7	–	–	1	0.01	1	0.01	1	0.01

factor σ for each local predictor and smoothing factor σ for fusion predictor). In this study, EEMDAN[2] is implemented using the program provided by Torres et al. [36] and the PACF was used to determine the numbers of input lag for the local GRNN predictors for each IMF. After determining the number of lagged inputs, the GRNN is optimized with respect to the smoothing factor. In this study, the value of smoothing factor is selected according to 5-fold cross-validation performance on the testing datasets. And in each fold experiment, an iterative optimization procedure is implemented in this study. The range of the smoothing factor parameter is selected large enough (between 0.01 and 1, increases at the step of 0.01) to ensure convergence for all-time series and the optimal smoothing factor for each IFM is then selected based on the corresponding minimum the root mean square error (RMSE) value between observed values and model outputs on the testing datasets.

In Stage 4, an independent GRNN model is used as trainable fusion strategy to obtain a more robust and accurate result. In this study, the inputs of the fusion GRNN predictor are the prediction results of IMFs, and the smoothing factor are selected through the implementation of iterative optimization procedures, described above. The final prediction value of the raw diarrhoea outpatient visits was obtained by post-processing in Stage 5.

4.3. The selected counterparts for comparison

In order to reflect the model superiority, it is necessary to build other models to compare with the proposed model. Seasonal ARIMA (SARIMA), Single GRNN, Wavelet-GRNN and EEMD–GRNN are selected as counterparts for the purpose of comparison. It should be noted the reason for selecting single GRNN is to justify the effectiveness of based on time series decomposition modeling framework, for the selection of SARIMA is due to the exhibited characteristics of strong seasonality of the diarrhoea outpatient visits data sets, and for the selection of Wavelet-GRNN is the similar modeling mechanism shared by EEMDAN-based and Wavelet-based modeling frameworks. The reason for selecting EEMD–GRNN is to justify the effectiveness of EEMDAN-based modeling framework. All methods were assessed using 10-fold cross-validation procedure.

In this study, the implementation of SARIMA model has three steps: model identification, parameter estimation, and diagnostic checking. We used the SPSS.19 to formulate the SARIMA model. The parameters in SARIMA models have been estimated using the ordinary least squares method and the most suitable model structure is identified using the Akaike Information Criterion (AIC) value. The model obtained from the 5-fold cross-validation are SARIMA$(2,1,1)\times(1,1,1)_{12}$ for children, SARIMA$(2,1,0)\times(1,1,1)_{12}$ for adult, SARIMA$(0,1,0)\times(0,1,1)_{12}$ for BJ and SARIMA $(1,1,1)\times(0,1,1)_{12}$ for GD. According to the above SARIMA models, the prediction results can be obtained for out-of-sample testing samples.

The Wavelet toolbox in Matlab is used to implement the discrete Wavelet transform. Generally speaking, when Wavelet transform is employed to construct a prediction model, the Wavelet basis functions and decomposition stages need to be determined first. In this study, Daubechies's Wavelets of order 4 is adopted through preliminary simulation in a trial-error fashion. To determine the number of decomposition levels, $L = \text{int}[\log(N)]$ is used [52], where L presents the decomposition level and N denotes the length of the data series.

In this paper, the EEMD[3] is also implemented using the program provided by [36]. The simulation of the Wavelet-GRNN and EEMD–GRNN are in general similar to the proposed EEMDAN–GRNN model and can be implemented following the above procedures. In order to save space and keep paper concise, detailed introduction of implementation procedures to these methods will not be repeated here.

4.4. Performance evaluation criteria

To numerical assessment the effectiveness of the different prediction models' accuracy, no single accuracy measure can capture all the distributional features of the errors when summarized across data series [22]. To identify the best model quantitatively, two criteria were used to evaluate and compare the models. These evaluation criteria include the Root Mean Square Error (RMSE) and

[2] Matlab code is available at: http://perso.ens-lyon.fr/patrick.flandrin/emd.html.

[3] Matlab code is available at: http://perso.ens-lyon.fr/patrick.flandrin/emd.html.

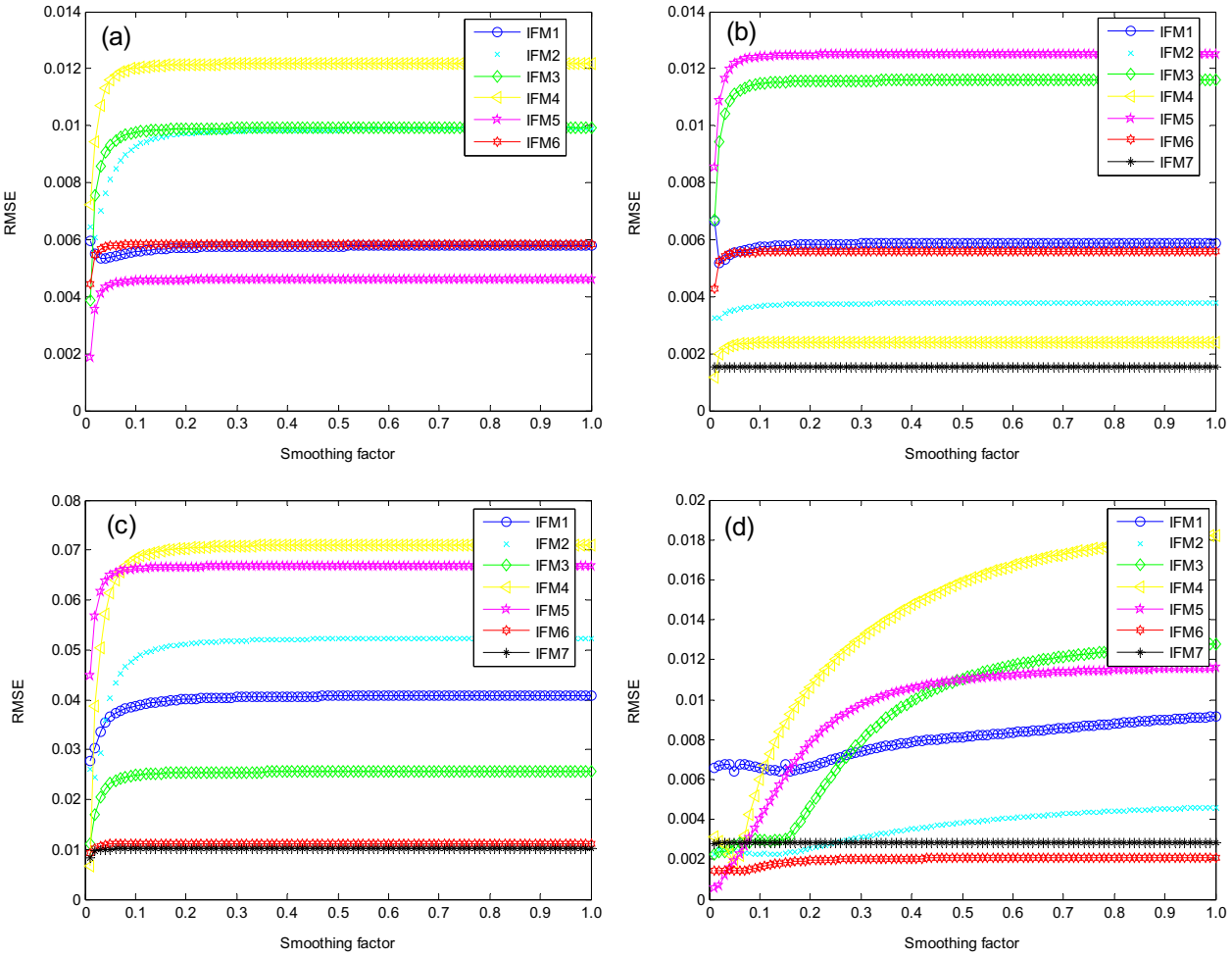

Fig. 5. Average RMSE of the local predictors on four datasets with different smoothing factor value. (a) SH_Children, (b) SH_Adult, (c) BJ, (d) GD.

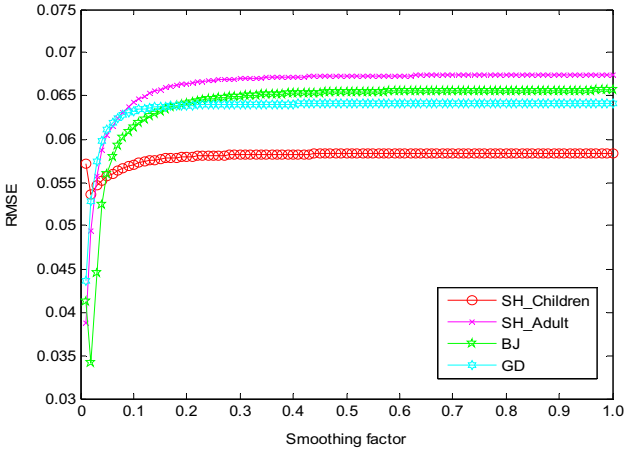

Fig. 6. Average RMSE of the fusion predictors on four datasets with different smoothing factor.

the Mean Absolute Percentage Error (MAPE) and are defined as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(y_i - \hat{y}_i)^2} \tag{9}$$

$$\mathrm{MAPE} = \frac{1}{N}\sum_{n=1}^{N}\frac{|y_i - \hat{y}_i|}{y_i} \times 100\% \tag{10}$$

Table 3
Values of smoothing factors provide the average cross-validation errors in the different datasets in developing fusion GRNN predictors.

Dataset	SF
SH_Children	0.02
SH_Adult	0.01
BJ	0.02
GD	0.01

where y_i is the actual observation value for a time period i, $\hat{y}_i$ is the prediction value for the same period and N is the number of observation in the hold-out sample. RMSE and MAPE were used to measure the correctness of a prediction in terms of levels and the deviation between the actual and predicted values. The smaller the values of RMSE and MAPE, the closer are the predicted time series values to the actual values. Note that the error measures are computed after rolling back of the preprocessing step performed, such as deseasonalization and detrending.

5. Experimental results and discussion

5.1. Parameter analysis and selection

In this subsection, we investigate the effect of the smoothing factor parameter on the proposed algorithm for all datasets. From

Table 4
Fivefold cross-validation error comparison results of different models for testing dataset. Bold values indicate the error of a data set.

Models	SH_Children		SH_Adult		BJ		GD	
	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
SARIMA	3302.05	5.19	1340.09	4.86	710.83	29.97	2010.39	11.23
GRNN	1335.79	4.51	1078.59	4.14	436.54	22.81	1041.20	9.65
Wavelet-GRNN	1201.21	3.86	908.62	3.31	311.87	20.63	864.96	8.14
EEMD–GRNN	1151.68	3.40	853.39	3.06	255.66	21.64	786.28	7.86
EEMDAN–GRNN	**1060.41**	**2.23**	**768.35**	**2.24**	**143.29**	**16.26**	**567.28**	**5.61**

Table 5
ANOVA results for hold-out ample.

Data	Measure	ANOVA test	
		Statistics *F*	*p*-value
SH_Children	ARE	7.72	0.000*
SH_Adult	ARE	47.3	0.000*
BJ	ARE	36.8	0.000*
GD	ARE	13.6	0.000*

* Indicates the difference among the five models is significant at the 0.05 level.

the analysis in Section 2.2, a good prediction performance for the GRNN model depends on the smoothing/spread factor. For all data sets, firstly PACF was used to determine the numbers of input lag for the local GRNN predictors for each IMF. The input lags of each IFM are given in Table 2. Figs. 5 and 6 show the average RMSE when the values of smoothing factor varies from 0.01 to 1.0 (in steps of 0.01) for local predictors and fusion predictors, where RMSE is obtained by 5-fold cross-validation.

Figs. 5 and 6 indicate that different behaviors when smoothing factor assumes different values. That is, for the all four datasets, local GRNN predictors and fusion GRNN predictors performs well when smoothing factor assumes small values (<0.1). This range was chosen in pilot tests that indicated that, for all the IFMs, the accuracies tends to become quite stable when smoothing factor is greater than 0.1 (see Figs. 5 and 6). However, GD dataset reveals a slightly different behavior at 7 local predictors. It can be seen from Fig. 5(d) that different local predictors show quite different behavior with different smoothing factor values. This can be due to the different datasets complexities. Table 2 summarizes the values of smoothing factor providing the average cross-validation errors in the testing datasets for different IFM of four datasets. Table 3 report the values of smoothing factor based on the average cross-validation errors in the testing datasets for fusion predictors.

5.2. Models comparison

To evaluate the performance of the proposed approach, the EEMDAN–GRNN model was compared with the EEMD–GRNN,

Table 6
Wilcoxon signed-rank tests results between the five models.

Data	Measure	Ranks of models								
		1		2		3		4		5
SH_Children	RMSE	EEMDAN	<*	EEMD	<	Wavelet	<*	GRNN	<*	ARIMA
	MAPE	EEMDAN	<*	EEMD	<	Wavelet	<*	GRNN	<*	ARIMA
SH_Adult	RMSE	EEMDAN	<*	EEMD	<	Wavelet	<*	GRNN	<*	ARIMA
	MAPE	EEMDAN	<*	EEMD	<	Wavelet	<*	GRNN	<*	ARIMA
BJ	RMSE	EEMDAN	<*	EEMD	<	Wavelet	<*	GRNN	<*	ARIMA
	MAPE	EEMDAN	<*	EEMD	<	Wavelet	<*	GRNN	<*	ARIMA
GD	RMSE	EEMDAN	<*	EEMD	<	Wavelet	<*	GRNN	<*	ARIMA
	MAPE	EEMDAN	<*	EEMD	<	Wavelet	<*	GRNN	<*	ARIMA

EEMDAN corresponds to the EEMDAN–GRNN model, EEMD corresponds to the EEMD–GRNN model and Wavelet corresponds to the Wavelet-GRNN model.
* Indicates the mean difference between the two adjacent models is significant at the 5% significance level.

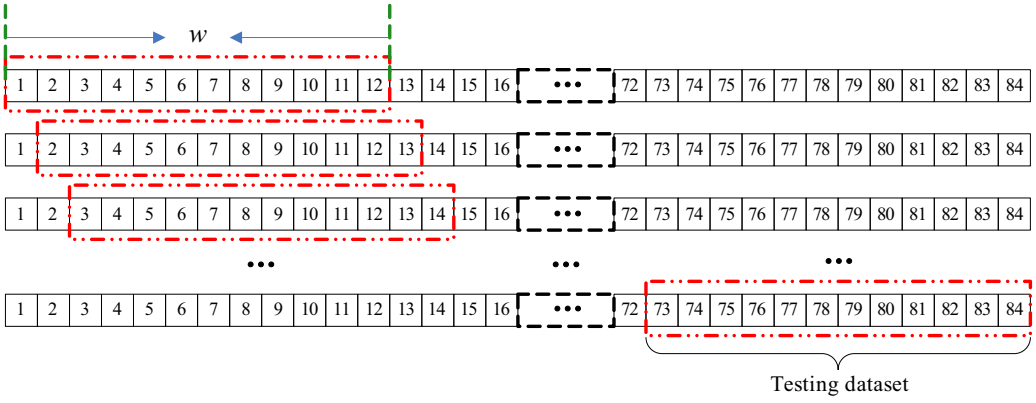

Fig. 7. Schematic illustration of the sliding window (*w* denotes the size of the window).

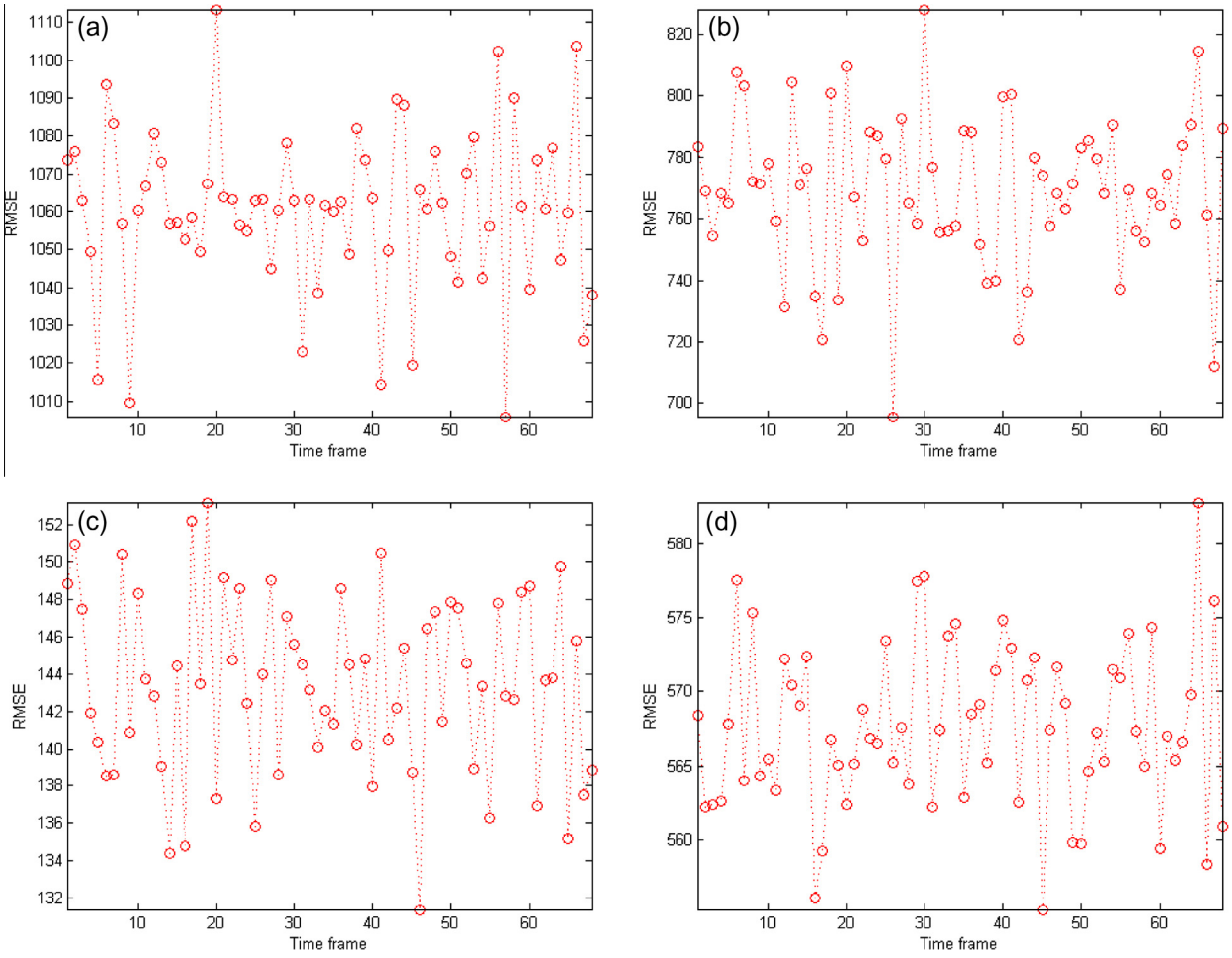

Fig. 8. The EEMDAN–GRNN algorithm behaves in different time frames. (a) SH_Children, (b) SH_Adult, (c) BJ, (d) GD.

Wavelet-GRNN, GRNN, and SARIMA approaches.[4] The average generalization errors (RMSE and MAPE) are calculated through 5-fold cross-validation procedure and the average results based on 5 runs for each data set are depicted in Table 4.

As comparison results presented in Table 4, one can deduce the following observation: (1) the based on time series decomposition prediction models (these are Wavelet-GRNN, EEMD–GRNN and proposed EEMDAN–GRNN) outperform the single global prediction model without time series decomposition methods (that is SARIMA and GRNN) without exception. The main reason could be that the decomposition strategy does effectively improve prediction performance. The Wavelet-GRNN, EEMD–GRNN and proposed EEMDAN–GRNN model adequately makes use of the advantages of the decomposition methods and GRNN algorithm and integrates them well. (2) The proposed EEMDAN–GRNN outperforms the EEMD–GRNN and Wavelet-GRNN. The reason for this may be related to the fact that the decomposition method of EEMDAN is superior to EEMD and Wavelet in terms of contribution to the prediction accuracy. Because EEMDAN effectively overcome the shortcomings existed in EEMD and EEMDAN is more suitable for nonlinear and non-stationary time series analysis. (3) As far as the comparison between the two single prediction models, the SARIMA model mostly ranks the last, while the GRNN can produce far better results. The possible reason is that SARIMA is a typical linear model not suitable for capturing nonlinear patterns hiding in the diarrhoea outpatient visits dataset. In summary it can be concluded that EEMDAN–GRNN has provided more accurate results for both the diarrhoea outpatient visits dataset compared to EEMD–GRNN, Wavelet-GRNN, GRNN and SARIMA models.

5.3. Significance test

In order to further evaluate whether the proposed EEMDAN–GRNN algorithm is superior to the EEMD–GRNN, Wavelet-GRNN, GRNN, and SARIMA algorithms in all four diarrhoea outpatient visits datasets prediction, the Analysis of Variance (ANOVA) procedure is applied for absolute residual errors (ARE) in hold-out sample. Table 5 shows the results of ANOVA test, from which we can see that the all the ANOVA results are significant at the 0.05 significance level, suggesting that there are significant differences among the five models.

To further identify the significant difference between any two models, the Wilcoxon signed-rank test [53] is used to compare all pairwise differences simultaneously here. the Wilcoxon signed-rank test has been widely used to evaluate the predictive capabilities of two different models to determine whether they are significantly different [37,54]. For the details of the Wilcoxon signed rank test, please refer to Zhang et al. [47] and Sudheer et al. [51]. Table 6 shows the results of these multiple comparison tests at 0.05 significance level. For each accuracy measure, we rank order the models from 1 (the best) to 5 (the worst). Table 5 shows that the proposed EEMDAN–GRNN model was significantly different (p-value < 0.05) from the other four models. Because the proposed method can be used to generate the smallest error in the

[4] It should be noted that all models are applied to the same datasets with the same portion of training and test data and the error measures are computed after rolling back of the preprocessing step performed, such as deseasonalization and detrending.

four datasets, we concluded that proposed method is significantly better for predicting four diarrhoea outpatient visits datasets relative to the other four models.

5.4. Algorithm behaves in different time frames

In order to see the generalization performance of the proposed algorithm in different time frames, a sliding window technology was used in this study. The implementation of the sliding window was carried out by splitting the original time series into continuous subsequences. The sliding window size w is set to 12 in this experiment. The dates in window as testing set were utilized to check the performance of the models. The schematic illustration of the sliding window is shown in Fig. 7.

The experiment results are detailed in Fig. 8. In Fig. 8, it can be observed that proposed EEMDAN–GRNN algorithm have better and stable performance in different time frames. This implies that the performance of proposed EEMDAN–GRNN algorithm doesn't depend on specific time frames. The high RMSE values in some time frames are due to the peak values of the datasets.

6. Conclusion

It can be realized from both theoretical and empirical findings that the prediction models based on decomposition-and-ensemble principle are the effective and efficient way to improve prediction performance. This paper proposed an EEMD-based time series decomposition modeling framework for diarrhoea outpatient visits prediction by adding adaptive noise at each stage of the decomposition along with employing data pre-processing and post-processing concepts as reinforcement technologies. Large scale experimental evidences are provided for the purpose of justification. The following are the main conclusions drawn in this paper: (1) Further validate the conclusion that prediction models based on decomposition-and-ensemble principle superior to single global predictors in prediction performances in the diarrhoea outpatient visits prediction problem. The proposed model adequately makes use of the advantages of the time series decomposition methods and muti-local predictor fusion. (2) The GRNN model is potentially a good candidate as predictor for diarrhoea outpatient visits, thanks to several of its unique properties (e.g., single design parameter). (3) The decomposition method of EEMDAN is superior to EEMD and Wavelet-based decomposition methods in terms of contribution to the prediction accuracy for diarrhoea outpatient visits prediction. Therefore EEMDAN–GRNN can be used as a suitable prediction tool for diarrhoea outpatient visits prediction problems.

Our study has the some limitations that need further research. Future studies may aim at exploring the utility of EEMDAN–GRNN in predicting the diarrhoea outpatient visits for other geographical regions and also for multiple-step-ahead prediction and also selecting other prediction models as local predictor such as support vector regression or extreme learning machines.

Acknowledgments

This work was supported by the Shanghai Scientific Development Foundation (Grant No. 13430710100). The authors would like to thank the editor and anonymous reviewers for their valuable suggestions in improving the quality of this paper.

References

[1] World Health Organization, 2013. <http://www.who.int/mediacentre/factsheets/fs330/en/index.html> (accessed 2013).
[2] I.N. Soyiri, D.D. Reidpath, An overview of health forecasting, Environ. Health Prev. Med. 18 (1) (2013) 1–9.
[3] 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.
[4] 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.
[5] N. Zhao, 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: 2010 IEEE Fifth International Conference Bio-Inspired Computing: Theories and Applications (BIC-TA), 2010, pp. 100–103.
[6] 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.
[7] 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.
[8] 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.
[9] R.B. Singh, S. Hales, N.D. 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.
[10] S.J. Lloyd, R.S. Kovats, B.G. Armstrong, Global diarrhea morbidity, weather and climate, Climate Res. 34 (2) (2007) 119–127.
[11] Héctor Allende, Claudio Moraga, Rodrigo Salas, Artificial neural networks in time series forecasting: a comparative analysis, Kybernetika 38 (6) (2002) 685–707.
[12] Tae Yoon Kim et al., Artificial neural networks for non-stationary time series, Neurocomputing 61 (2004) 439–447.
[13] Z. Tang, C.D. Almeida, P.A. Fishwick, Time series forecasting using neural networks vs. Box-Jenkins methodology, Simulation 57 (5) (1991) 303–310.
[14] G.P. Zhang, An investigation of neural networks for linear time-series forecasting, Comput. Oper. Res. 28 (12) (2001) 1183–1202.
[15] Ratnadip Adhikari, R.K. Agrawal, An Introductory Study on Time Series Modeling and Forecasting, 2013, pp. 1302–6613, arXiv preprint.
[16] D.P. Solomatine, A. Ostfeld, Data-driven modelling: some past experiences and new approaches, J. Hydroinform. 10 (1) (2008) 3–22.
[17] A.J.C. Sharkey, Combining Artificial Neural Nets: Ensemble and Modular Multi-Net Systems, Springer-Verlag, 2001.
[18] H. Chris Tseng, Bassam Almogahed, Modular neural networks with applications to pattern profiling problems, Neurocomputing 72 (10) (2009) 2093–2100.
[19] Gasser Auda, Mohamed Kamel, Modular neural networks: a survey, Int. J. Neural Syst. 9 (2) (1999) 129–151.
[20] C.L. Wu, K.W. Chau, Prediction of rainfall time series using modular soft computing methods, Eng. Appl. Artif. Intell. (2012).
[21] Y.Y. Chen, S.W. Wang, N. Chen, X.Q. Long, X.R. Tang, Forecasting Cohesion less Soil Highway Slope Displacement Using Modular Neural Network, Discrete Dynamics in Nature and Society, 2012.
[22] T. Xiong, Y.K. Bao, Z.Y. Hu, Does restraining end effect matter in EMD-based modeling framework for time series prediction? Some experimental evidences, Neurocomputing 123 (2014) 174–184.
[23] J.M. Hu, J.Z. Wang, G.W. Zeng, A hybrid forecasting approach applied to wind speed time series, Renew. Energy 60 (2013) 185–194.
[24] C.F. Chen, M.C. Lai, C.C. Yeh, Forecasting tourism demand based on empirical mode decomposition and neural network, Knowl.-Based Syst. 26 (2012) 281–287.
[25] D. Baratta, G. Cicioni, F. Masulli, L. Studer, Application of an ensemble technique based on singular spectrum analysis to daily rainfall forecasting, Neural Netw. 16 (3) (2003) 375–387.
[26] M. Theodosiou, Disaggregation & aggregation of time series components: a hybrid forecasting approach using generalized regression neural networks and the theta method, Neurocomputing 74 (6) (2011) 896–905.
[27] C.F. Tsai, Y.C. Lin, D.C. Yen, et al., Predicting stock returns by classifier ensembles, Appl. Soft Comput. 11 (2) (2011) 2452–2459.
[28] L. Tang, L. Yu, S. Wang, J. Li, S. Wang, A novel hybrid ensemble learning paradigm for nuclear energy consumption forecasting, Appl. Energy 93 (2012) 432–443.
[29] C.H. Cheng, L.Y. Wei, A novel time-series model based on empirical mode decomposition for forecasting TAIEX, Econ. Modell. 36 (2014) 136–141.
[30] L.A. Yu, S.Y. Wang, K.K. Lai, Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm, Energy Econ. 30 (2008) 2623–2635.
[31] Z. Guo, W. Zhao, H. Lu, J. Wang, Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model, Renew. Energy 37 (2012) 241–249.
[32] Y. Wei, M.C. Chen, Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks, Transp. Res. Part C: Emerg. Technol. 21 (2012) 148–162.
[33] G. Napolitano, F. Serinaldi, L. See, Impact of EMD decomposition and random initialization of weights in ANN hind casting of daily stream flow series: an empirical examination, J. Hydrol. 406 (2011) 199–214.
[34] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, in: Proceedings of the Royal Society of

London Series A-mathematical Physical and Engineering Sciences, Series A, 1998, pp. 903–995.
[35] Z. Wu, N.E. Huang, Ensemble empirical mode decomposition: a noise-assisted data analysis method, Adv. Adapt. Data Anal. 1 (1) (2009) 1–41.
[36] M.E. Torres, M.A. Colominas, G. Schlotthauer, et al., A complete ensemble empirical mode decomposition with adaptive noise, in: 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2011, pp. 4144–4147.
[37] W. Yan, Toward automatic time-series forecasting using neural networks, IEEE Trans. Neural Netw. Learn. Syst. 23 (7) (2012) 1028–1039.
[38] P. Flandrin, G. Rilling, P. Goncalves, Empirical mode decomposition as a filter bank, IEEE Signal Process. Lett. 11 (2) (2004) 112–114.
[39] D. Specht, A general regression neural network, IEEE Trans. Neural Netw. 2 (1991) 568–576.
[40] H. Li, S. Guo, C. Li, et al., A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm, Knowl.-Based Syst. 37 (2013) 378–387.
[41] N.K. Ahmed, A.F. Atiya, N.E. Gayar, H. El-Shishiny, An empirical comparison of machine learning models for time series forecasting, Econ. Rev. 29 (5–6) (2010) 594–621.
[42] D.X. Niu, H.Q. Wang, and Z.H. Gu, Short-term load forecasting using general regression neural network, in: Proc. Int. Conf. Mach. Learn. Cyber, 2005, pp. 4076–4082.
[43] M. Leung, A. Chen, H. Daouk, Forecasting exchange rates using general regression neural networks, Comput. Oper. Res. 27 (2002) 1093–1110.
[44] S.B. Kotsiantis, D. Kanellopoulos, P.E. Pintelas, Data preprocessing for supervised learning, Int. J. Comput. Sci. 1 (2006).
[45] A. Azadeh, M. Saberi, S. Ghaderi, A. Gitiforouz, V. Ebrahimipour, Improved estimation of electricity demand function by integration of fuzzy system and data mining approach, Energy Convers. Manage. 49 (2008) 2165–2177.
[46] Janset Kuvulmaz, Serkan Usanmaz, Seref Naci Engin, Time-series forecasting by means of linear and nonlinear models, MICAI 2005: Advances in Artificial Intelligence, Springer, Berlin, Heidelberg, 2005.
[47] G.Q. Zhang, B. Eddy Patuwo, Michael Y. Hu, Forecasting with artificial neural networks: the state of the art, Int. J. Forecasting 14 (1) (1998) 35–62.
[48] G.J. Bowden, G.C. Dandy, H.R. Maier, Input determination for neural network models in water resources applications. Part 1-background and methodology, J. Hydrol. 301 (1) (2005) 75–92.
[49] G. Zhang, B. Peter, P. Eddy, M.Y. Hu, A simulation study of artificial neural networks for nonlinear time-series forecasting, Comput. Oper. Res. 28 (4) (2001) 381–396.
[50] Ö. Kisi, Constructing neural network sediment estimation models using a data-driven algorithm, Math. Comput. Simul. 79 (1) (2008) 94–103.
[51] K.P. Sudheer, A.K. Gosain, K.S. Ramasastri, A data-driven algorithm for constructing artificial neural network rainfall-runoff models, Hydrol. Process. 16 (6) (2002) 1325–1330.
[52] V. Nourani, M.T. Alami, M.H. Aminfar, A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation, Eng. Appl. Artif. Intell. 22 (2009) 466–472.
[53] F. Wilcoxon, Individual comparisons by ranking methods, Biometrics 1 (6) (1945) 80–83.
[54] J.J. Wang et al., A novel hybrid approach for wind speed prediction, Inform. Sci. 273 (2014) 304–318.
[55] 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.
[56] 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.
[57] Y.M. Wang, J. Li, J.Z. Gu, Z.L. Zhou, Z.J. Wang, Artificial neural networks for infectious diarrhea prediction using meteorological factors in Shanghai (China), Appl. Soft Comput. 35 (2015) 280–290.
[58] K.D. Lafferty, The ecology of climate change and infectious diseases, Ecology 90 (4) (2009) 888–900.