Industrial Management & Data Systems

Thailand tourism forecasting based on a hybrid of discrete wavelet decomposition and NARX neural network

Ratree Kummong Siriporn Supratid

Received 12 November 2015
Revised 22 January 2016
Accepted 12 February 2016

Thailand tourism forecasting based on a hybrid of discrete wavelet decomposition and NARX neural network

Ratree Kummong and Siriporn Supratid
College of Information and Communication Technology, Rangsit University, Pathumthani, Thailand

Abstract
Purpose – Accurate forecast of tourist arrivals is crucial for Thailand since the tourism industry is a major economic factor of the country. However, a nonstationarity, normally consisted in nonlinear tourism time series can seriously ruin the forecasting computation. The purpose of this paper is to propose a hybrid forecasting method, namely discrete wavelet decomposition (DWD)-NARX, which combines DWD and the nonlinear autoregressive neural network with exogenous input (NARX) to cope with such nonstationarity, as a consequence, improve the effectiveness of the demand-side management activities.
Design/methodology/approach – According to DWD-NARX, wavelet decomposition is executed for efficiently extracting the hidden significant, temporal features contained in the nonstationary time series. Then, each extracted feature set at a particular resolution level along with a relative price as an exogenous input factor are fed into NARX for further forecasting. Finally, the forecasting results are reconstructed. Forecasting performance measures rely on mean absolute percentage error, mean absolute error as well as mean square error. Model overfitting avoidance is also considered.
Findings – The results indicate the superiority of the DWD-NARX over other efficient related neural forecasters in the cases of high forecasting performance rate as well as competently coping with model overfitting.
Research limitations/implications – The scope of this study is confined to Thailand tourist arrivals forecast based on short-term projection. To resolve such limitations, future research should aim to apply the generalization capability of DWD-NARX on other domains of managerial time series forecast under long-term projection environment. However, the exogenous input factor is to be empirically revised on domain-by-domain basis.
Originality/value – Few works have been implemented either to handle the nonstationarity, consisted in nonlinear, unpredictable time series, or to achieve great success on finding an appropriate and effective exogenous forecasting input. This study applies DWD to attain efficient feature extraction; then, utilizes the competent forecaster, NARX. This would comprehensively and specifically deal with the nonstationarity difficulties at once. In addition, this study finds the effectiveness of simply using a relative price, generated based on six top-ranked original tourist countries as an exogenous forecasting input.
Keywords Discrete wavelet decomposition,
Nonlinear autoregressive neural network with exogenous input, Nonstationary time series
Paper type Research paper

1. Introduction
Tourism is vital for several countries, as a result of a large consumption of money for businesses with their goods and services as well as the opportunity for employment in the service industries, including airline transportation and accommodation services. As a consequence, tourism both directly and indirectly contributes a significant share to nation's gross domestic product (GDP). Unlike manufacturing, retail trade or

Industrial Management & Data Systems
Vol. 116 No. 6, 2016
pp. 1242-1258

0263-5577
DOI 10.1108/IMDS-11-2015-0463

construction, tourism industrial product and services, e.g., hotel rooms, airline seats and car rentals have a perishable nature; they cannot be inventoried. Hence, to avoid the financial cost of excess capacity or the opportunity costs of unfulfilled demand, tourism management planning is crucial. In this process, the accurate forecast of tourism volume in the form of arrivals is especially important because it is an indicator of precise future demand, thereby providing basic information for subsequent managerial planning as well as decision and policy making. Management sectors, including private as well as government ones can use such basic information to plan their future operations and to foresee the need for facilities and infrastructure development.

The purpose of this study is to propose a novel forecasting model, a hybrid between discrete wavelet decomposition (DWD) (Mallat, 1989) and a nonlinear autoregressive neural network with exogenous factors (NARX) (Chen *et al.*, 1990; Narendra and Parthasarathy, 1990), namely DWD-NARX. The proposed method can be used by hospitality managers as well as forecast practitioners to produce accurate forecasts of tourist flows to Thailand. NARX, one of proficiently forecasting tools exploits recurrent neural architecture (Elman, 1990). As opposed to other recurrent neural networks (RNN), it has limited feedback architectures that come only from the output neuron instead of from hidden neurons. It has been reported that such learning architecture can yield more effective results in NARX model than in other recurrent architectures with hidden states (Horne and Giles, 1995). In order to enhance NARX forecasting performance, a relative price is employed here as an exogenous input as it economically influences tourists' destination-decision choice in some significance. However, nonstationarity usually existed in nonlinear tourism time series possibly severely degrades the forecasting performance. DWD, therefore is employed here for effectively coping with such nonstationarity problem. With respect to the proposed method, DWD is executed to achieve efficient feature extraction for an individual level of frequency resolution in the tourism time series environment. Then, the extracted feature set at a particular frequency resolution level along with a relative price as an exogenous input factor are fed into NARX for further forecasting. At last, the forecasting outputs from all the resolution levels are reconstructed. An inbound tourism demand to Thailand, which is one of the main tourist destinations in Asia is focussed in this paper. DWD-NARX is tested with Thailand tourism monthly time series over January 1999 to December 2013. Performance evaluation relies on mean absolute percentage errors (MAPE), mean absolute error (MAE) as well as mean square error (MSE). The remaining parts of the paper are organized as follows. In Sections 1.1 and 1.2 motivation for Thailand tourism forecast and literature review are respectively mentioned. NARX and wavelet decomposition are briefly reviewed in Section 2. The proposed DWD-NARX is described in Section 3. Section 4 delineates experimental comparisons among the proposed method and some related classical neural forecasters. Managerial implications are stated in Section 5. The overall conclusion is drawn in the last section.

1.1 Motivation for Thailand tourism forecast

Thailand is worldwide known as the "Land of Smiles." *Time* magazine reported in 2013 (Quan, 2013) that Bangkok, the capital of Thailand was identified as one of the most visited city in the world by the global destination cities index; while Suvarnabhumi Airport and Siam Paragon Shopping Mall were the world's most geotagged location on Instagram. In addition, a famous tourism destination in Thailand is represented by an elegant, attractive decorated Wat Phra Kaew, also known as the Temple of the Emerald Buddha or the Grand Palace (HikerBays, n.d.). Moreover, there are not only several popular dream beaches at Phuket, Kho Samui as well as Pattaya, but also many great places with perfect Thai food, ranging from street stall to luxurious restaurants (DiscoveryThailand, 2015). The importance of tourism to Thai economics as a significant contribution to the GDP is noticeable. Based on the annual research by the World Travel and Tourism Council (2014), the total contribution of the tourism industry to Thailand GDP in 2013 was accounted for 20.2 percent of GDP or about THB2,401.10 billion; it supported 15.4 percent of the total employment with about 6.01 million jobs. In regarding to its direct contribution, the tourism contributed 9.0 percent of total GDP or about THB1,074.0 billion and supported more than two million jobs or 6.6 percent of total employment. Such figures indicate that tourism has become a significant economic factor of Thailand. This main reason, in turn, has caused an increasing interest in further advanced precise forecasting approaches.

1.2 Literature review

The most commonly used time series forecasting techniques have pertained to ARIMA, GARCH (Babu and Reddy, 2015; Lim and McAleer, 2002). However, these models, or even a specialized demand heuristic forecasting one (Petropoulos *et al.*, 2013) adopt piecewise linear function as basic elements of prediction model. The functional form for the problem has to be specified by users. It could take a lot of time to experiment with the different possible function relations and algorithms to obtain proper models. In fact, fitting a nonlinear model to time series data is slightly more involved than fitting a linear one. On the other side, artificial neural networks (ANNs) (Rumelhart *et al.*, 1986), nonlinear forecasting techniques have explicit flexibility; and can be used to estimate the nonlinear of input-output relationship without the limits of traditional linear time series models. The outperformance of ANNs over the classical time series model has been shown in Zhang *et al.*, (2013). Backpropagation neural network (BPNN) (Rumelhart *et al.*, 1986) is a multilayer mapping network that minimizes an error backward while information is transmitted forward during training procedures. BPNN has been utilized in forecasting applications such as inflation and software tool adoption forecasts (Aiken, 1999; Morris *et al.*, 2004). According to Teixeira and Fernandes (2014) and Claveria and Torra (2014), BPNN has also been applied on tourism forecast. Unlike the BPNN, generalized regression neural network (GRNN) (Specht, 1991) does not need an iterative training procedure. Only the widths of hidden-to-output weights, based on radial basis function need to be learned. GRNN has been employed for tourism forecast in Hu *et al.* (2012) and Çuhadar *et al.* (2014).

Nevertheless, both BPNN and GRNN are executed on a basis of static-based learning. It was indicated in El-Shafie *et al.* (2012) the advantage of time series forecasting accuracy based on dynamic over static neural network model during both training and testing stages. This supported a consistent level of accuracy regarding seen and unseen data. A powerful class of dynamic nonlinear systems can be referenced as nonlinear autoregressive neural network with exogenous factors (NARX) (Chen *et al.*, 1990; Narendra and Parthasarathy, 1990). The NARX model has been extensively used in various applications (Çoruh *et al.*, 2014; Sheremetov *et al.*, 2014; Chang *et al.*, 2014). In Xu *et al.* (2014) and Babu and Pachiyappan (2014), superior time series forecasting results were achieved using NARX over BPNN. In addition, NARX yielded the best results compared to BPNN, GRNN and Elman's RNN as reported in Mitrea *et al.*, (2009). One of the important characteristics of the NARX model is that the evolution of a phenomenon can be explained by its previous behavior as well as the effect of exogenous factors when they exist. According to the past researches

(Song *et al.*, 2003; Wong *et al.*, 2006), cost of living, referenced as a relative price variable has been determined as the exogenous factor that mainly contributes to the tourism demand forecast. Such a relative price is defined as a weighted average ratio of the consumer price indices between the destination and the origin countries, adjusted by the bilateral exchange rate. A lower consumer price index along with a higher exchange rate in favor of the origin country's currency can result in more tourists visiting the destination from that origin country. The relative price variable is also used to measure the impacts of inflation and exchange rate between the relative countries.

Even if an efficient dynamical nonlinear forecasting model is employed, different kinds of nonstationarities should be bore in mind. Such nonstationarities can emerge from seasonal fluctuations and irregular influences, existing in tourism time series. As such time series have often been analyzed in time domain, this does not provide any information with respect to the frequency components. Distinction between low-, medium- and high-frequency fluctuations of time series can be better studied by transforming data to frequency domain. Indeed, traditional Fourier (1822) spectral analysis, has been used to identify the different frequency components of a time series, e.g. trends, cycles, seasonalities, noise; and to quantify their respective importance based on sine wave. As it is based on a stationary assumption, thereby seems to be too restrictive because the tourism as well as other types of economic time series are often subject to regime shifts, jumps, volatility clustering, outliers or long-term trends. In order to overcome this drawback, a short-time Fourier transform (Gabor, 1946), also known as Gabor or windowed Fourier transform has been developed. In order to estimate the frequencies, it uses a time period or window less than the number of observations N. Then, the sample signal is split into subsamples; and the Fourier transform is computed on these subsamples. The existing problem relates to the right choice of the window and its constancy over time. The DWD (Mallat, 1989), taking its roots from Fourier transform can overcome the limitation of both traditional and short-time Fourier analysis, as DWD can combine information from both time and frequency domain with no requirement of stationarity. It extracts the different frequencies that drive the concerned variable in the time domain by decomposing into its time scale components. Each of these scale components reflects the evolution of the signal through time at a particular frequency resolution level. The scales, from the shortest to the largest consecutively represent high-, medium- and low-frequency fluctuations. This denotes the innovation of the wavelet decomposition that its scale window can be adjusted automatically to various levels of frequency by employing time compression or dilatation rather than a variation of frequency in the modulated signal. Thus, wavelet analysis is capable of revealing particular aspects of data such as trends, breakdown points, discontinuities in higher derivatives and self-similarity, hidden in nonstationary time series. Therefore, DWD has been employed to decompose time series data into wavelet frequency levels for further forecasting. The related forecasting applications includes stock index, financial time series, oil price and short-term power load forecasting (Kao *et al.*, 2013; Zhang *et al.*, 2001; Yousefi *et al.*, 2005; Amjady and Keynia, 2009).

2. A brief review of NARX neural network and DWD

NARX forecasting neural network has been applied for modeling discrete-time nonlinear dynamical systems (Leontaritis and Billings, 1985; Norgaard *et al.*, 2000). In this network, the next value of the dependent output signal at time t+1 is regressed on previous values of the output signal and exogenous input signal. Thus, the NARX model can be mathematically represented as:

$$y(t+1) = f\left[y(t), y(t-1), \ldots,\ y\left(t-d_y\right); u(t), u(t-1), \ldots, u(t-d_u)\right] \tag{1}$$

According to (1), the input of the NARX network is formed by two types of regressors, one is the exogenous input regressor $u(t) \in \mathbb{R}$, here represented by the relative price variable, and the other one is the output regressor $y(t) \in \mathbb{R}$, here referred as the tourist arrival amount at time t. $d_u \in \mathbb{Z}$, and $d_y \in \mathbb{Z}$, consecutively are the *lags* of the exogenous input and output regressors of the system, where $d_v \geqslant d_u \geqslant 1$. The nonlinear mapping $f(\cdot)$ is generally an unknown smooth function; and can be approximated by a standard multilayer perceptron network. Here, the output of the NARX network is estimated under a series-parallel (SP) concept, where the output's regressor is formed by actual values of the system's output, as shown as follows:

$$\hat{y}(t+1) = \hat{f}\left[y(t), y(t-1), \ldots,\ y\left(t-d_y\right); u(t), u(t-1), \ldots, u(t-d_u)\right] \tag{2}$$

where $\hat{y}(t+1)$ and $\hat{f}(\cdot)$ are an estimators of $y(t+1)$ and $f(\cdot)$, respectively. To enhance NARX forecasting capability, DWD can be adopted, since it provides powerful support to analyze a given tourist arrival time series, $y(t)$ at various frequency fluctuation levels. The DWD applies recursively a succession of the low- and high-pass filters or wavelet filtering basis functions to time series $y(t)$ which allows separating its high-frequency components from the low-frequency ones. The decomposition can be illustrated as a form of dyadic tree. Figure 1 delineates a dyadic tree of third-level DWD.

First $y(t)$ is decomposed into the trend or the approximate, low-frequency components at the first-level scale decomposition, $A_1(t)$ and the deviations from the trend or the detail, high-frequency components at the first-level scale decomposition, $D_1(t)$. $A_1(t)$ is successively decomposed into low- and high-frequency components at the second-level, denoted as $A_2(t)$ and $D_2(t)$ consecutively, and so on. Based on this dyadic tree, the tourist arrivals time series $y(t)$ can be reconstructed by combining $A_3(t)$, $D_3(t)$, $D_2(t)$ and $D_1(t)$, that refers to the inverse wavelet transform.

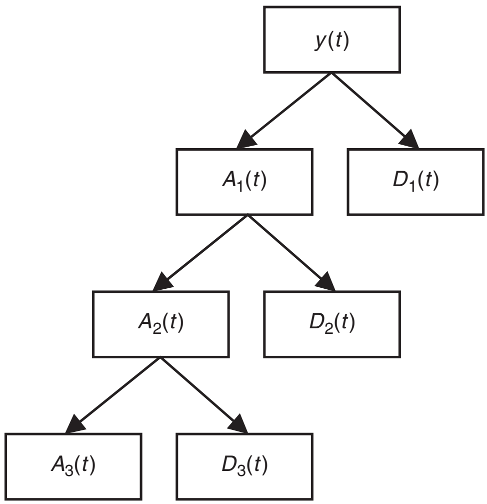


Figure 1.
Three-level discrete wavelet decomposition dyadic tree

3. The proposed DWD-NARX

The proposed hybrid forecasting method, DWD-NARX serves as a combination of DWD and the SP-architecture NARX. However, one of the key issues for the success of tourism forecasting is a suitable selection of an appropriate exogenous input factor. With regard to tourism forecasting, the exogenous input factor is carefully constructed based on economic theory; it should influence the way in which tourists make their decision regarding the destination choice. By this reason, an exogenous input factor at time t is defined, here as a relative price variable $u(t)$ which is a ratio that measures the weighted average consumer price index in Thailand relative to that in some particular number of original countries, adjusted by the relevant exchange rates (Song *et al.*, 2003). One to seven top-ranked original tourist countries have been considered. Those seven top-ranked countries sequentially include Malaysia, China, Japan, Korea, UK and USA and Singapore, in descending order. The computation of such relative price $u(t)$ is shown as follows:

$$u(t) = \sum_{i=1}^{K} \frac{CPI_{Th}(t)/EX_{Th}(t)}{CPI_i(t)/EX_i(t)} \cdot w_i(t) \tag{3}$$

where $CPI_{Th}(t)$ and $CPI_i(t)$ (Trading Economics, 2015; FRED, 2015) are monthly consumer price index regarding Thailand and original country i, respectively. $EX_{Th}(t)$ and $EX_i(t)$ (OANDA, 2015) successively are the exchange rate indices for Thailand and original country i. $K = 1,2, \ldots, 7$ represents the number of considered top-ranked original countries. The exchange rates refer to the monthly market rates of the local currency against the US dollar. The weight $w_i(t)$ is calculated by:

$$w_i(t) = \frac{TA_i(t)}{\sum_{j=1}^{K} TA_j(t)} \tag{4}$$

where $TA_i(t)$ is the number of tourist arrivals from original country i. Pearson's correlation coefficient (r) values of the relative prices $u(t)$ based on one to seven top-ranked original countries in respect to Thailand tourist arrivals amount are calculated, resulting as 0.62, 0.43, 0.56, 0.62, 0.68, 0.73, 0.64. Among these values, the best competent correlation degree, r in respect to Thailand tourist arrivals amount is yielded from the $u(t)$ based on six top-ranked original countries. This verifies the explanatory efficiency of the exogenous input to predict the tourism time series in a certain level regarding the limited scope here.

With the use of $u(t)$ based on six top-ranked original countries, an overall framework of the proposed DWD-NARX is presented in Figure 2. It can be explained in terms of training and testing phases as follows:

(1) Training phase:

- A set of tourist arrivals training signal is decomposed by DWD into $A_3(t)$, $D_3(t)$, $D_2(t)$ and $D_1(t)$.
- Each component is separately fed along with the corresponding exogenous input, $u(t)$ to an individual NARX for model training purpose. This indicates that not only the approximate part or the trend of the signal, but also the detail part or the deviation at various fluctuation types and levels are taken into the forecasting analysis.

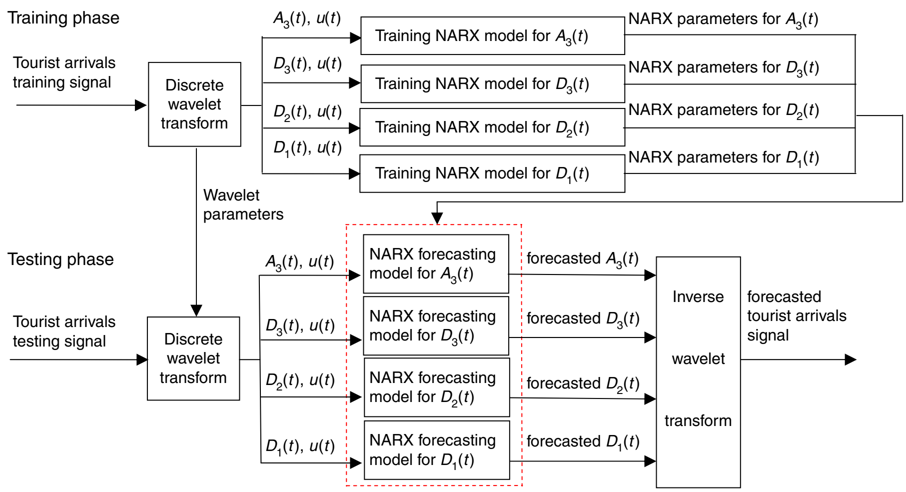


Figure 2.
An overall framework of the proposed DWD-NARX

- At the end of this learning phase, the wavelet as well as NARX trained parameters for each component are obtained for further using in testing phase.

(2) Testing phase:

- A set of tourist arrival testing signal is decomposed into $A_3(t)$, $D_3(t)$, $D_2(t)$ and $D_1(t)$.
- Each component is separately fed along with the corresponding exogenous input, $u(t)$ to an individual NARX for forecasting purpose, using the NARX trained parameters obtained from the training phase.
- All the forecasted components are supplied to inverse wavelet transform process for finally reconstructing the forecasted tourist arrivals signal.

It is noticed that $A_3(t)$, $D_3(t)$, $D_2(t)$ and $D_1(t)$ represent the hidden features extracted by DWD. $A_3(t)$ is the low-frequency approximate components; whereas $D_3(t)$, $D_2(t)$ and $D_1(t)$ are the high-frequency detail components, ordered from coarsest to finest. Based on DWD, $A_3(t)$ as well as $D_3(t)$, $D_2(t)$ and $D_1(t)$ are extracted respectively by:

$$A_3(t) = \sum_{\forall n} a_{3,n} \varphi_{3,n}(t) \tag{5}$$

$$D_j(t) = \sum_{\forall n} d_{j,n} \psi;_{j,n}(t) \tag{6}$$

where $j = 1, 2, 3$ and $n \geqslant 0$ refers to time interval which is an integer type. From (5), one can say that $A_3(t)$ is a linear combination of the wavelet low-pass filtering basis function, $\varphi(t)$. In a similar manner, from (6) $D_3(t)$, $D_2(t)$ and D_1(t) are linear combination of the wavelet high-pass filtering basis function, $\psi(t)$. DWD finds the coefficients $a_{3,n}$ and $d_{j,n}$ by projecting the tourist arrival signal, $y(t)$ onto the wavelet low- and high-pass filtering functions consecutively at different time interval and particular frequency resolution levels.

DWD-NARX was developed using MATLAB2015a. The achievement of the forecast is measured in terms of MAPE, MAE and MSE, as follows:

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{y(t)-\hat{y}(t)}{y(t)}\right| \tag{7}$$

$$MAE = \frac{1}{N}\sum_{t=1}^{N}\left|y(t)-\hat{y}(t)\right| \tag{8}$$

$$MSE = \frac{1}{N}\sum_{t=1}^{N}(y(t)-\hat{y}(t))^2 \tag{9}$$

where $y(t)$ and $\hat{y}(t)$ respectively are actual and forecasted amount of tourist arrivals in month t; N is the forecasting time horizon, depending on training or testing phase. Three lags of tourist arrivals amount together with three lagged observations of explanatory exogenous variable, referred to the relative price variable are employed as inputs to the forecasting model. This is expressed as follows:

$$\hat{f}[y(t-1),\ y(t-2),\ y(t-3),\ u(t-1),\ u(t-2),\ u(t-3)] \rightarrow \hat{y}(t) \tag{10}$$

where $\hat{f}(\cdot)$ is an estimator of nonlinear mapping function, which is approximated by the forecasting method.

4. Experimental results

Thailand tourism monthly dataset over January 1999 to December 2013 (Ministry of Tourism and Sports, 2014) is focussed in the experiments. In total, 80 and 20 percent of total 180 data are successively processed in training and testing phases. To evaluate the forecasting performance, tenfold cross-validation is incorporated in the experiment using rolling forecasting window (Hyndman, 2010) that moves the training window and keeps it of fixed length in forecasting time horizon. Figure 3 demonstrates an example with respect to the structure of training and testing datasets for each fold of cross-validations; where gray and white boxes consecutively represent training and testing interval periods.

Necessary parameters used in the comparative forecasters are declared in Table I. Such parameters are selected so as to achieve the effective results regarding each individual forecasting method, using a few number of working neurons. With respect to the proposed method DWD-NARX, wavelet parameters are chosen depending on a competency to evaluate local and global behavior of signal. Experimental comparison relies on popular, effective wavelet low- and high-pass filtering functions: Daubechies (db), Coiflet (coif) as well as Symlet (sym), under various scale decomposition levels.

Based on Table I, DWD-NARX using 6-tab Daubechies with third-level wavelet decomposition provides the best forecasting results during training phase. Relying on the best of tenfold cross-validation with regard to DWD-NARX, Figures 4 and 5 exhibit approximate and detail parts of the tourist arrival signal after wavelet decomposition and after NARX forecasting in training and testing phases sequentially. It is seen especially in the first- and second-level fine detail components the decrease of sudden changes or discontinuities in the signal amplitude at certain instants of time, as a results of NARX forecasting; whereas long-term trend, represented by the third-level approximate and the third-level coarse-detail components remain almost the same.

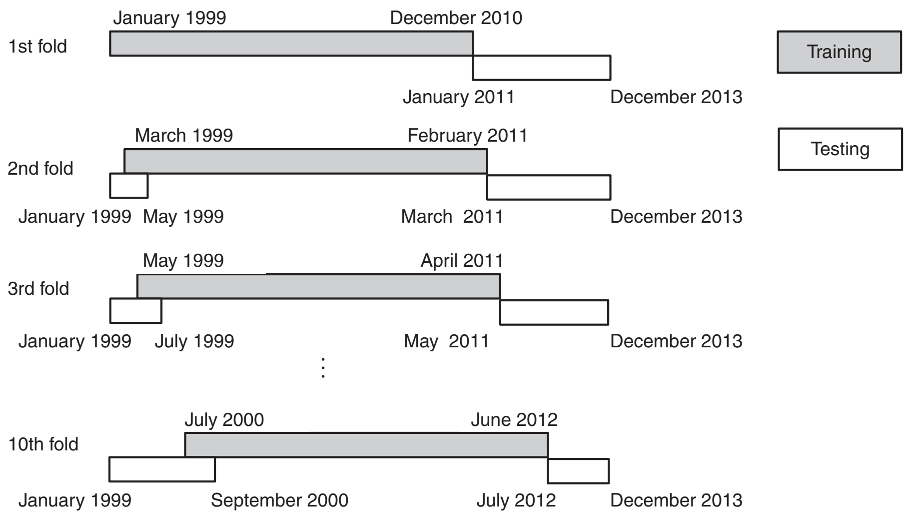


Figure 3. An example with respect to the structure of training and testing sets for each fold of cross-validations

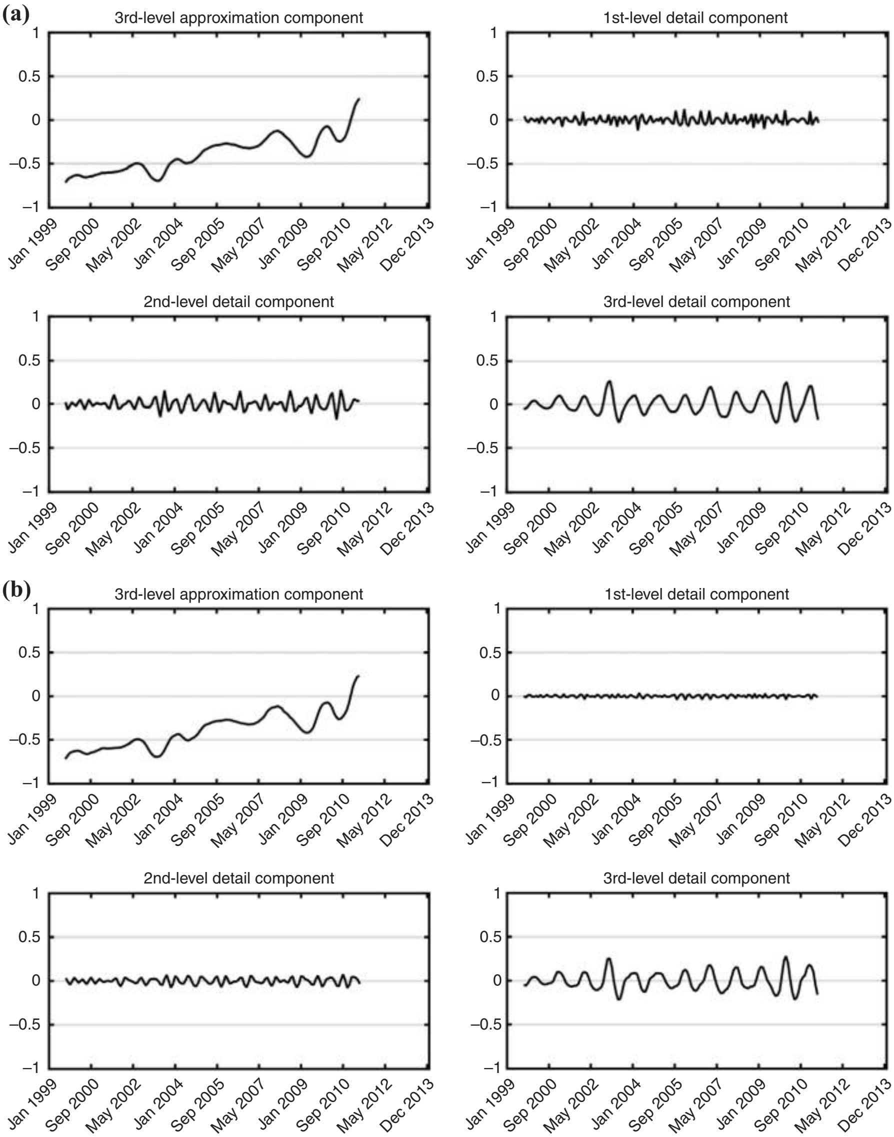
Table I. Parameters used in the forecasting methods

Forecasting methods	Parameters used
DWD-NARX	3 hidden neurons, wavelet decomposition level 3, db6
NARX, BPNN	3 hidden neurons
RNN	1 hidden neuron
GRNN	3 hidden neurons, 1.0087 spread

With respect to the comparative forecasting methods, the reconstructed training and testing outputs relying on the best of tenfold cross-validation are depicted along with the actual data in Figure 6. This figure reveals a significant matching between the prediction and the actual data, yielded by DWD-NARX compared to the other forecasting methods. The evidence of the significant matching as well as a few deviation between the forecasted and the actual data is obviously pointed in Table II, where DWD-NARX generates the least of the best as well as median values and few standard deviation of forecasting errors based on MAPE, MAE and MSE using training and testing cases for ten fold cross-validation.

With regard to the median of forecasting errors, 54.44, 53.45 and 77.86 percent of MAPE, MAE and MSE improvement based on DWD-NARX over NARX for the training cases are respectively yielded; whereas 63.60, 61.97 and 81.71 percent improvement are consecutively resulted for the testing ones. In total, 30.60, 32.44 and 53.71 percent of MAPE, MAE and MSE improvement based on NARX over BPNN are respectively calculated for the training cases; while 26.40, 18.62 and 30.77 percent are successively provided for the testing ones. Relying on Wilcoxon's rank sum test with 99 percent confidence interval, the better performance of the proposed DWD-NARX than NARX can be confirmed by *p*-value of one with respect to all the three types of error measure. Such confirmation insists the powerful capability of the wavelet analysis as a feature extraction for improving the forecasting performance.

Furthermore, Table III displays the highest, median values and standard deviation of absolute Pearson's correlation coefficient between the forecasted and the actual



Notes: (a) Approximate and detail parts of the tourist arrival signal after wavelet decomposition in training phase, from August 1999 to April 2011; (b) approximate and detail parts of the tourist arrival signal after NARX forecasting in training phase, from August 1999 to April 2011

Figure 4. Approximate and detail parts of the tourist arrival signal with regard to training phase, after wavelet decomposition (a) and after NARX forecasting (b)

tourist arrivals for the testing cases, with respect to tenfold cross-validation. It is apparently seen that the DWD-NARX provides the most reliable and precise forecast, compared to the others; whereas the least correlations between the prediction and actual values, given by RNN indicate the weak forecasting performance.

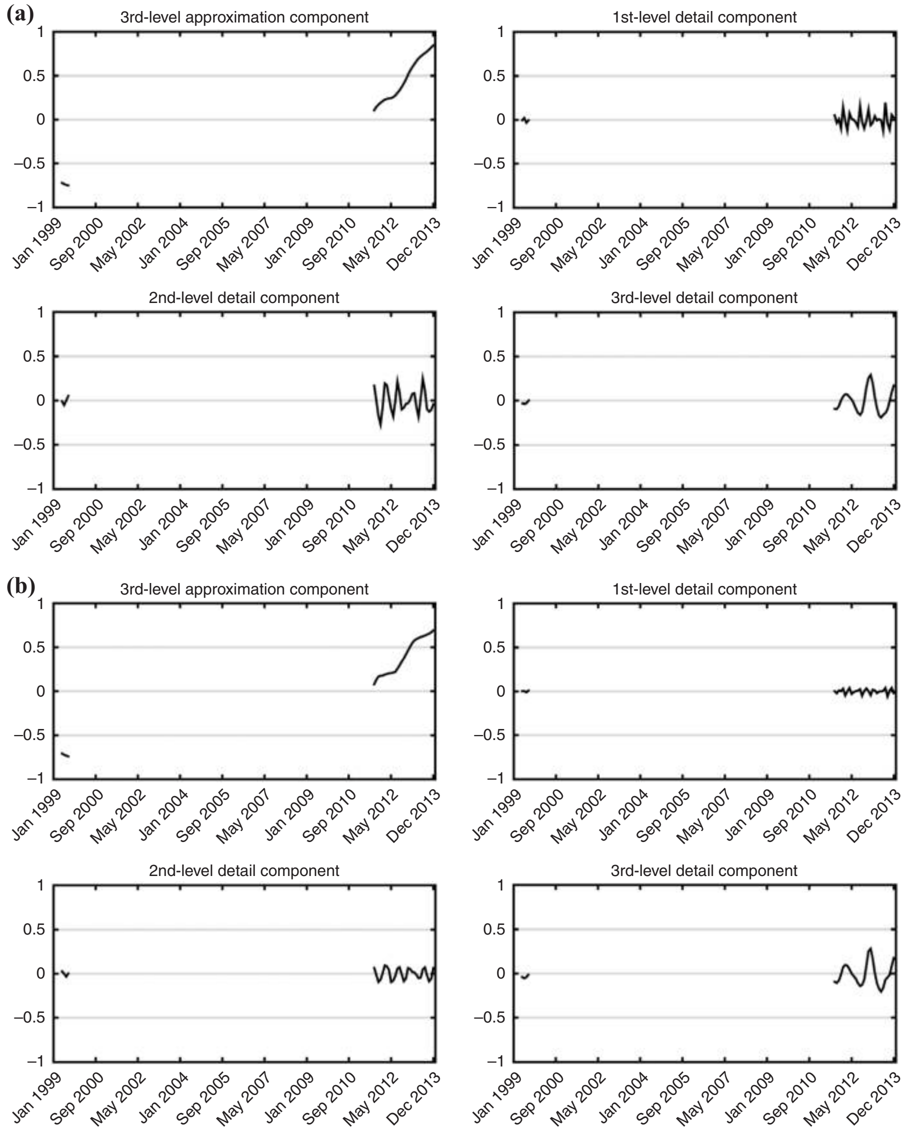
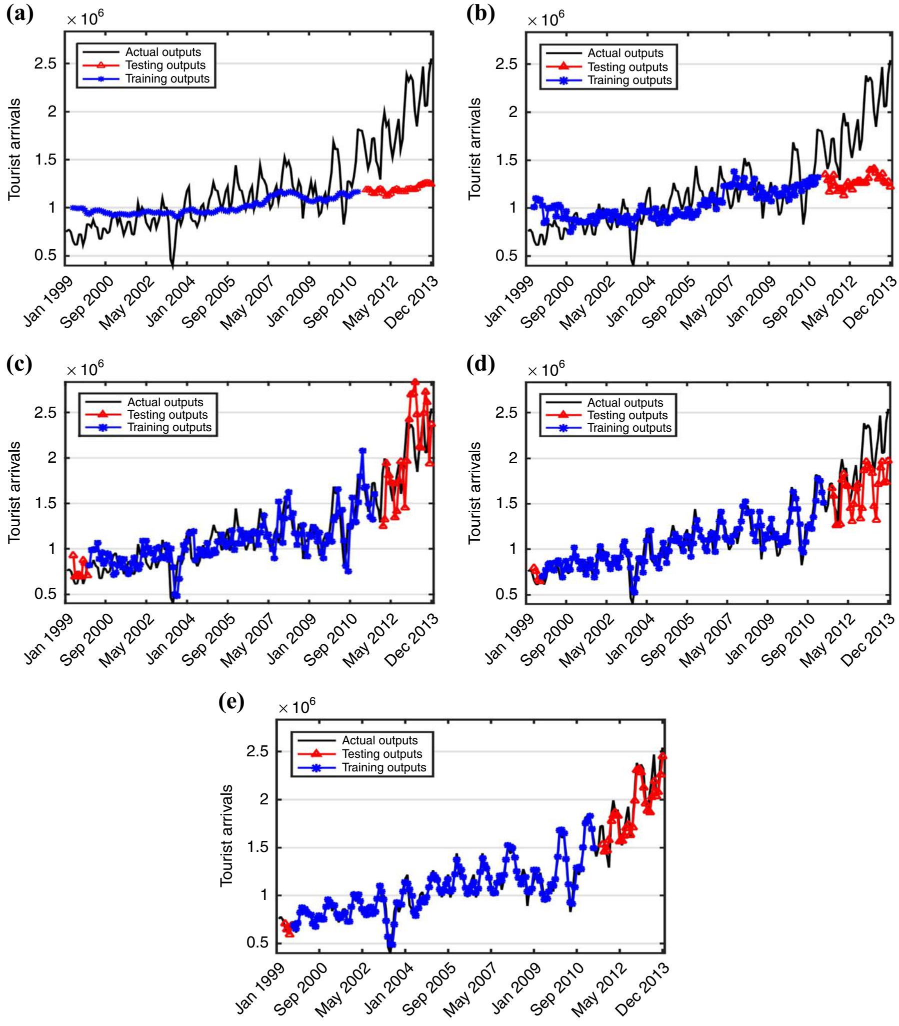


Figure 5.
Approximate and detail parts of the tourist arrival signal with regard to testing phase, after wavelet decomposition (a) and after NARX forecasting (b)

Notes: (a) Approximate and detail parts of the tourist arrival signal after wavelet decomposition in testing phase, from April 1999 to July 1999 and from August 2011 to December 2013; (b) approximate and detail parts of the tourist arrival signal after NARX forecasting in testing phase, from April 1999 to July 1999 and from August 2011 to December 2013

According to Tables II and IV, the model overfitting is evaluated by comparing the best and median forecasting errors based on MAPE, MAE and MSE for seen and unseen, different sets of training and testing cases, with regard to tenfold cross-validation. As aforementioned, Table II explicitly reports the evidences that DWD-NARX yields the



Notes: (a) GRNN; (b) RNN; (c) BPNN; (d) NARX; (E) DWD-NARX

Figure 6. The actual data, along with the model training and testing outputs, achieved from the best of tenfold cross-validation for each forecaster

best performance in terms of forecasting errors, relying on MAPE, MAE and MSE for training and testing cases. Table IV displays difference percentages of forecasting errors between training and testing cases, which can also be accounted as one of criteria measurements for overfitting appraisal. The less difference of forecasting errors between training and testing ones is, the fewer the overfitting of the model would be. It is seen from Table IV that DWD-NARX generates the least overfitting for nearly all types of errors.

5. Managerial implications

The proposed DWD-NARX, which combines DWD with NARX forecaster is accounted as a bridge for the gap between computer-science/engineering theory and industrial management practice. Wavelet analysis, here referred as DWD represents

Table II. Best, median values and standard deviation of forecasting errors

Methods	MAPE						MAE						MSE					
	Training			Testing			Training			Testing			Training			Testing		
	Best	Median	SD	Best	Median	SD	Best	Median	SD	Best	Median	SD	Best	Median	SD	Best	Median	SD
GRNN	15.94	17.39	1.16	35.51	40.90	3.20	1.62E+05	1.77E+05	1.26E+04	5.67E+05	6.26E+05	8.17E+04	4.34E+10	5.32E+10	7.84E+09	4.25E+11	4.83E+11	1.15E+11
RNN	12.80	13.61	1.11	30.95	37.71	13.06	1.29E+05	1.40E+05	9.37E+03	4.57E+05	6.94E+05	2.44E+05	2.99E+10	3.32E+10	3.11E+09	2.64E+11	5.90E+11	7.48E+11
BPNN	11.57	12.81	0.97	17.62	23.22	7.86	1.13E+05	1.31E+05	1.07E+04	2.45E+05	3.49E+05	1.40E+05	2.07E+10	2.83E+10	4.49E+09	8.35E+10	1.69E+11	1.81E+11
NARX	8.71	8.89	0.10	15.52	17.09	2.21	8.49E+04	8.85E+04	3.54E+03	2.02E+05	2.84E+05	7.29E+04	1.17E+10	1.31E+10	1.50E+09	5.76E+10	1.17E+11	6.52E+10
DWD-NARX	3.81	4.05	0.15	5.12	6.22	3.48	3.75E+04	4.12E+04	2.11E+03	6.47E+04	1.08E+05	8.14E+04	2.43E+09	2.90E+09	4.02E+08	6.38E+09	2.14E+10	4.45E+10

computer-science/engineering-oriented theory; whereas the nonlinear autoregressive neural network using relative price as an exogenous input factor, referred here as NARX can be regarded as an efficient forecasting tool for managerial forecasting practice. Without DWD, NARX alone can yield over 26 percent better forecasting performance based on the median of MAPE than BPNN, RNN and GRNN; on the other side, NARX with the use of DWD provides more than 63 percent MAPE-based forecasting improvement over NARX alone for the unseen, different sets of testing cases, as seen in the previous section. The powerful capability of feature extraction, belonging to DWD causes both low forecasting error rate and model overfitting. Such high performance along with few overfitting forecast play a key role in enhancing demand-side tourism management. This leads to more suitably consolidated management planning, which in turn precedes high-performing managerial decision making. Moreover, the generalization potentiality of DWD to extract and analyze trends as well as various types and levels of changes from the trend, hidden in nonstationary time series contributes to some important implication. Such implication indicates that the proposed model is not only limited to Thailand tourism forecast, but is also possibly applied with nonstationary time series forecast, related to other management domains. However, the exogenous input factor is to be empirically reconsidered on domain-by-domain or case-by-case basis, for the sake of effective forecast.

6. Conclusion

This paper presents a hybrid forecasting method, namely DWD-NARX for coping with nonstationarity problem, usually existed in nonlinear tourism time series that highly possibly degrade the forecasting performance. DWD is capable to handle high-level nonstationary time series; it has been demonstrated to be a powerful tool for analyzing particular aspects of data such as trends, breakdown points, discontinuities in higher derivatives, hidden in the time series data. Based on DWD, such hidden features can be

Table III. Highest, median values and standard deviation of absolute Pearson's correlation coefficient between the forecasted and the actual tourist arrivals

	$\lvert r \rvert$		
Methods	Highest	Median	SD
GRNN	0.9442	0.9206	0.1342
RNN	0.9004	0.8127	0.1523
BPNN	0.9401	0.8591	0.2174
NARX	0.9815	0.9489	0.1279
DWD-NARX	0.9961	0.9828	0.0508

Table IV. Difference percentages of forecasting errors between training and testing

	MAPE		MAE		MSE	
Methods	Best	Median	Best	Median	Best	Median
GRNN	122.77	135.19	250.00	253.67	879.26	807.89
RNN	141.80	177.08	254.26	395.71	749.50	1,677.11
BPNN	52.29	81.26	116.81	166.41	303.38	497.17
NARX	78.19	92.24	137.93	220.90	392.31	793.13
DWD-NARX	34.38	53.58	72.53	162.14	162.55	637.93

extracted in time domain in the form of approximate as well as fine and coarse detail components at various frequency resolution levels. In this work, the potent capability of the wavelet analysis as an effective feature extraction tool for enhancing forecasting performance is insisted. Each extracted feature set at a specific resolution level along with a particular type of exogenous explanatory input are fed into the nonlinear autoregressive of neural network with exogenous factors (NARX), a powerful class of dynamic nonlinear forecasting systems. The exogenous variable, here defined as the relative price with respect to six top-ranked original tourist countries is taken into an account. Within the scope of this work, such relative price has been found to yield the most well-suited correlation associated with the tourist arrivals amount. This confirms, in a certain level the explanatory efficiency of the exogenous input to predict the tourism data.

With the proposed DWD-NARX, the prediction performance can be enhanced significantly in terms of error reduction as well as model overfitting avoidance, compared to NARX alone along with some efficient, classical neural models such as GRNN, Elman's RNN and BPNN. The results denote that the DWD-NARX is proficiently competent to improve the effectiveness of the demand-side management activities. Performance assessment relies on MAPE, MAE and MSE. The limitation of this work relates that the proposed DWD-NARX is expressed to be apparently suitable for short-term projection in complex and uncertain environment, focussing on Thailand tourist arrivals forecasting application. However, the generalization of DWD capability allows the proposed DWD-NARX to be applied with other domains of time series forecasting applications. Thereby, future works may focus on further enhancing such restriction, aiming to explore the proposed forecasting model based on long-term projection under other domains of time series environment.

References

Aiken, M. (1999), "Using a neural network to forecast inflation", *Industrial Management & Data Systems*, Vol. 99 No. 7, pp. 296-301.

Amjady, N. and Keynia, F. (2009), "Short-term load forecasting of power systems by combination of wavelet transform and neuro-evolutionary algorithm", *Energy*, Vol. 34 No. 1, pp. 46-57.

Babu, C.N. and Reddy, B.E. (2015), "Prediction of selected Indian stock using a partitioning-interpolation based ARIMA-GARCH model", *Applied Computing and Informatics*, Vol. 11 No. 2, pp. 130-143.

Babu, N.R. and Pachiyappan, A. (2014), "Dynamic neural network based very short-term wind speed forecasting", *Wind Engineering*, Vol. 38 No. 2, pp. 120-128.

Chang, F.-J., Chen, P.-A., Lu, Y.-R., Huang, E. and Chang, K.-Y. (2014), "Real-time multi-step-ahead water level forecasting by recurrent neural networks for urban flood control", *Journal of Hydrology*, Vol. 517, September, pp. 836-846.

Chen, S., Billings, S.A. and Grant, P.M. (1990), "Non-linear system identification using neural networks", *International Journal of Control*, Vol. 51 No. 6, pp. 1191-1214.

Claveria, O. and Torra, S. (2014), "Forecasting tourism demand to Catalonia: neural networks vs time series models", *Economic Modelling*, Vol. 36, January, pp. 220-228.

Çoruh, S., Geyikçi, F., Kılıç, E. and Çoruh, U. (2014), "The use of NARX neural network for modeling of adsorption of zinc ions using activated almond shell as a potential biosorbent", *Bioresource Technology*, Vol. 151, January, pp. 406-410.

Çuhadar, M., Cogurcu, I. and Kukrer, C. (2014), "Modelling and forecasting cruise tourism demand to Izmir by different artificial neural", *International Journal of Business and Social Research*, Vol. 4 No. 3, pp. 12-28.

DiscoveryThailand (2015), "Travel guide Thailand destinations", available at: www.discoverythailand.com/destination.asp (accessed April 16, 2015).

Elman, J.L. (1990), "Finding structure time", *Cognitive Science*, Vol. 14 No. 2, pp. 179-221.

El-Shafie, A., Noureldin, A., Taha, M., Hussain, A. and Mukhlisin, M. (2012), "Dynamic versus static neural network model for rainfall forecasting at Klang River Basin, Malaysia", *Hydrology and Earth System Sciences*, Vol. 16 No. 4, pp. 1151-1169.

Fourier, J.B. (1822), *Théorie Analytique de la Chaleur*, Chez Firmin Didot, père et fils, Paris.

FRED (2015), "Main economic indicators", available at: http://research.stlouisfed.org/fred2/release?rid=205 (accessed January 23, 2015).

Gabor, D. (1946), "Theory of communication", *Journal of the Institute of Electrical Engineers*, Vol. 93 No. 26, pp. 429-457.

HikerBays (n.d.), "Thailand tourist board and the most important informations for travelers", available at: http://hikersbay.com/asia/thailand?gclid=CJmrqJaH-8QCFRcOjgodGZEA_w (accessed April 16, 2015).

Horne, B.G. and Giles, C.L. (1995), "An experimental comparison of recurrent neural networks", *Proceedings of the Conference Neural Information Processing Systems 1994, MIT Press, Denver*, pp. 697-704.

Hu, Y., Xie, R. and Zhang, W. (2012), "Tourists flow prediction by clustering-based GRNN", *Proceedings of the 9th International Forum on Digital TV and Wireless Multimedia Communication in Shanghai, Springer, Chennai, November 9-10*, pp. 396-402.

Hyndman, R.J. (2010), "Why every statistician should know about cross-validation", available at: http://robjhyndman.com/hyndsight/crossvalidation/ (accessed August 23, 2015).

Kao, L.-J., Chiu, C.-C., Lu, C.-J. and Chang, C.-H. (2013), "A hybrid approach by integrating wavelet-based feature extraction with MARS and SVR for stock index forecasting", *Decision Support Systems*, Vol. 54 No. 3, pp. 1228-1244.

Leontaritis, I.J. and Billings, S.A. (1985), "Input-output parametric models for non-linear systems part I: deterministic non-linear systems", *International Journal of Control*, Vol. 41 No. 2, pp. 303-328.

Lim, C. and McAleer, M. (2002), "Time series forecasts of international travel demand for Australia", *Tourism Management*, Vol. 23 No. 4, pp. 389-396.

Mallat, S.G. (1989), "A theory for multiresolution signal decomposition: the wavelet representation", *IEEE Transactions on Pattern Recognition and Machine Intelligence*, Vol. 11 No. 7, pp. 674-693.

Ministry of Tourism and Sports,Thailand (2014), "Statistic tourist", available at: www.tourism.go.th/home/listcontent (accessed March 1, 2014).

Mitrea, C.A., Lee, K.M. and Wu, Z. (2009), "A comparison between neural networks and traditional forecasting methods: a case study", *International Journal of Engineering Business Management*, Vol. 1 No. 2, pp. 19-24.

Morris, S.A., Greer, T.H., Hughes, C. and Clark, W.J. (2004), "Prediction of CASE adoption: a neural network approach", *Industrial Management & Data Systems*, Vol. 104 No. 2, pp. 129-135.

Narendra, K.S. and Parthasarathy, K. (1990), "Identification and control of dynamical systems using neural networks", *IEEE Transactions*, Vol. 1 No. 1, pp. 4-27.

Norgaard, M., Ravn, O., Poulsen, N. and Hansen, L. (2000), *Neural Networks for Modelling and Control of Dynamic Systems*, Springer, Berlin.

OANDA (2015), "Average exchange rates", available at: www.oanda.com/currency/average (accessed January 23, 2015).

Petropoulos, F., Nikolopoulos, K., Spithourakis, G.P. and Assimakopoulos, V. (2013), "Empirical heuristics for improving intermittent demand forecasting", *Industrial Management & Data Systems*, Vol. 113 No. 5, pp. 683-696.

Quan, K. (2013), "And the world's no. 1 tourist destination is … Bangkok is now the most-visited city in world", available at: http://newsfeed.time.com/2013/06/01/bangkok-claims-the-worlds-no-1-tourist-destination-title/ (accessed April 16, 2015).

Rumelhart, D.E. and McClelland, J.L., Corporate PDP Research Group (1986), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 1, MIT Press, Cambridge MA.

Sheremetov, L., Cosultchi, A., Martínez-Muñoz, J., Gonzalez-Sánchez, A. and Jiménez-Aquino, M. (2014), "Data-driven forecasting of naturally fractured reservoirs based on nonlinear autoregressive neural networks with exogenous input", *Journal of Petroleum Science and Engineering*, Vol. 123 No. SI, pp. 106-119.

Song, H., Wong, K.K. and Chon, K.K. (2003), "Modelling and forecasting the demand for Hong Kong tourism", *International Journal of Hospitality Management*, Vol. 22 No. 4, pp. 435-451.

Specht, D.F. (1991), "A general regression neural networks", *IEEE Transactions*, Vol. 2 No. 6, pp. 568-567.

Teixeira, J. and Fernandes, P. (2014), "Tourism time series forecast with artificial neural networks", *Review of Applied Management Studies*, Vol. 12 Nos 1-2, pp. 26-36.

Trading Economics (2015), "Consumer price index (CPI)", available at: www.tradingeconomics.com/ (accessed January 22, 2015).

Wong, K.K., Song, H. and Chon, K.S. (2006), "Bayesian models for tourism demand forecasting", *Tourism Management*, Vol. 27 No. 5, pp. 773-780.

World Travel and Tourism Council (2014), "Annual research: key facts", available at: www.wttc.org/-/media/files/reports/economic%20impact%20research/country%20reports/thailand2014.pdf (accessed March 3, 2015).

Xu, Y.L., Chen, H.X., Guo, W. and Zhu, Q.Y. (2014), "A comparison of NARX and BP neural network in short-term building cooling load prediction", *Applied Mechanics and Materials*, Vols 513-517, February, pp. 1545-1548.

Yousefi, S., Weinreich, I. and Reinarz, D. (2005), "Wavelet-based prediction of oil prices", *Chaos, Solitons and Fractals*, Vol. 25 No. 2, pp. 265-275.

Zhang, B.-L., Coggins, R., Jabri, M. and Dersch, D. (2001), "Multiresolution forecasting for futures trading using wavelet decompositions", *IEEE Transactions on Neural Networks*, Vol. 12 No. 4, pp. 765-775.

Zhang, X., Liu, Y., Yang, M., Zhang, T., Young, A.A. and Li, X. (2013), "Comparative study of four time series methods in forecasting typhoid fever incidence in China", *Plos One*, Vol. 8 No. 5, pp. 1-11.

Corresponding author
Siriporn Supratid can be contacted at: siri_sup1@hotmail.com