

Forecasting Patient Visits to Hospitals using a WD&ANN-based Decomposition and Ensemble Model

Lean Yu ^{1*}, Geye Hang ¹, Ling Tang ¹, Yang Zhao ², K.K. Lai ³

¹ School of Economics and Management, Beijing University of Chemical Technology, Beijing 100029, CHINA

² Institute of Science and Development, Chinese Academy of Sciences, Beijing 100190, CHINA

³ Department of Management Sciences, City University of Hong Kong, Kowloon, HONG KONG

Received 23 May 2017 • Revised 9 August 2017 • Accepted 27 September 2017

ABSTRACT

Forecasting the number of patient visits to hospitals has aroused an increasingly large interest from both theoretic and application perspectives. To enhance the accuracy of forecasting hospital visits, this paper proposes a hybrid approach by coupling wavelet decomposition (WD) and artificial neural network (ANN) under the framework of "decomposition and ensemble". In this model, the WD is first employed to decompose the original monthly data of the number of patient visits to hospitals into several components and one residual term. Then, the ANN as a powerful prediction tool is implemented to fit each decomposed component and generate individual prediction results. Finally, all individual prediction values are fused into the final prediction output by simple addition method. For illustration and verification, four sets of monthly series data of the number of patient visits to hospitals are used as the sample data, and the results show that the proposed model can obtain significantly more accurate forecasting results than all considered popular forecasting techniques.

Keywords: patient visits forecasting, time series data, decomposition and ensemble, wavelet decomposition, artificial neural networks

INTRODUCTION

The high efficiency of hospital management depends to some degree on appropriate allocation of material resources and proper physician and nurse staffing because of the limit of those resources and hospital budget pressure. Forecasting the number of patient visits to hospitals can be helpful in allocating limited human and material resources of hospitals (Hadavandi et al., 2012). For instance, forecasting of short-term hospital census may result in improvement of inpatient bed allocation and decrease in the incidence of overstaffing and understaffing (Littig et al., 2007). Nowadays, for its great help in hospitals' resource allocation, forecasting the number of patient visits to hospitals has been paid more and more attention to and achieved a significant status in hospital management (Safar & Alkhezzi, 2016). More accurate hospital visits prediction can contribute to higher efficiency of hospital management without any doubt.

In the existing literature, a lot of studies have been done in this field and traditional forecasting techniques, especially autoregressive integrated moving average (ARIMA) modeling, are quite popular. For instance, Milner (1997) modeled attendances at accident and emergency departments by one-off original ARIMA. Abdel-Aal and Mangoud (1998) applied an ARIMA model in forecasting the monthly patient volume at a health care clinic. Diaz et al. (2001) used ARIMA to forecast emergency hospital admissions. Zibners (2006) used Box-Jenkins ARIMA to predict patient visits to an academic pediatric emergency department (ED). Friede et al. (2009) utilized ARIMA to predict daily trauma admissions considering the effect of weather, weekday and other variables. Abraham et al. (2009) applied a seasonal ARIMA to forecast emergency occupancy. Sun et al. (2009) established ARIMA models to predict daily attendance at an ED. Kam (2010) used seasonal ARIMA to predict daily number of patient visits to ED. ARIMA was constructed to predict ED visitor volume in Chen et al.'s study (2011). Kadri et al. (2014) used ARIMA to forecast daily patient attendances. Exponential smoothing model, as a common time series forecasting

© **Authors.** Terms and conditions of Creative Commons Attribution 4.0 International (CC BY 4.0) apply.

✉ yulean@amss.ac.cn (*Correspondence) ✉ rainontheway@sina.com ✉ tangling_00@126.com

✉ 404446053@qq.com ✉ msklai@outlook.com

Contribution of this paper to the literature

- WD-FNN-ADD learning paradigm can effectively improve the forecasting precision, in terms of RMSE and MAPE criteria, since WD-FNN-ADD model outperforms all the benchmark models. The proposed WD-FNN-ADD model is a very promising tool in forecasting complex time series data of hospital visits.
- The proposed WD-FNN-ADD approach can also be utilized to address other difficult forecasting tasks, especially for complex, volatile data with multiple data characteristics.
- How to select the most appropriate wavelet in wavelet decomposition is extremely crucial, which exerts an influence on the decomposition results and accordingly has an impact on forecasting results.

tool, is developed by Bergs (2014) to model monthly ED visits. Kovats et al. (2004) applied autoregressive Poisson to model daily emergency hospital admissions. Koestler et al. (2013) introduced a seasonality adjusted Poisson Autoregressive (PAR) model to hospital census forecasting.

Other traditional methods applied in forecasting hospital visits are presented as follows. Dan and Qualls (1994) compared five types of prediction models including raw observations, moving averages, mean values with moving averages, seasonal indicators with moving averages, and ARIMA in predictions of ED volume, length of stay, and acuity. And it turned out that simpler models performed best in their study. Rotstein et al. (1997) formulated a General Linear Model (GLM) to describe ED patient volume, which implied the model's usefulness in short-range forecasting of patient volume. Batal et al. (2001) developed prediction equations for daily patient volume via stepwise linear regression analysis. Reis and Mandl (2003) constructed a trimmed mean seasonal model in expected number of daily patient visits to an ED. Brillman et al. (2005) established a first order cyclical regression model with fixed-width sine and cosine harmonics as the seasonal component and a hierarchical model with scalable Gaussian function as the seasonal component for ED daily respiratory chief complaints. Flottesmesch et al. (2007) established a mathematical model of ED census which could be utilized as a forecasting method. In Boyle et al.'s study (2008), a general linear regression model formed with 11 dummy variables in predicting monthly patient admissions performed best. Jones (2009) utilized Vector autoregression (VAR) to model and forecast demand in the emergency department. Au-Yeung et al. (2009) predicted patient arrivals to an accident and emergency department via a structural time series model.

Among so many traditional approaches, the prediction efficacy differs. Morzuch and Allen (2006) compared an unobserved components model (UCM) and double exponential smoothing in prediction of hourly hospital ED arrivals, and the latter outperformed the former. Three forecasting models including hourly historical average, seasonal ARIMA and sinusoidal with an autoregression-structured error term were used in ED bed occupancy by Schweigler et al. (2009). Marcilio et al. (2013) used time-series methods including generalized linear models, generalized estimating equations and seasonal ARIMA to forecast daily ED visits. Kim et al. (2014) applied exponential smoothing, ARIMA, seasonal ARIMA, and generalized autoregressive conditional heteroscedasticity (GARCH) methods in patient volume forecasting. Although these traditional techniques are popular in forecasting hospital visits, they are not good at dealing complexity in hospital visits data.

Another approach for hospital visits forecasting is Artificial Intelligence (AI) such as Artificial Neural Network (ANN) (Pan, 2017). Although ANN doesn't seem as prevalent as those traditional methods in this field, it is widely used in other fields of prediction as a powerful forecasting tool. However, there are still a few studies in this field with AI involved. For example, Jones et al. (2008) compared seasonal ARIMA, exponential smoothing, time series regression and ANN in daily ED patient volume prediction to linear regression and the former four methods didn't bring consistent good results in forecasting out of sample though they all show better in sample fitting. Aladag and Aladag (2012) explored modeling the number of outpatient visits by ANN with different activation functions. Xu et al. (2013) modeled daily patient arrivals at ED via ANN. However, though superior to traditional methods, AI methods also have their own shortcomings. For example, the local minimum and overfitting problems sometimes occur while using ANN.

Since both single traditional and AI methods mentioned above seem to come across a bottleneck in improving the accuracy of predicting patient visits to hospitals, some hybrid models have been formulated to dig more deeply in this field. Those hybrid models are advantageous in dealing with data complexity, which can lead to better forecasting results. Cheng et al. (2008) coupled a new fuzzy time series method based on weighted-transitional matrix with an expectation method and a grade-selection method as forecasting methods to predict the number of outpatient visits. Garg et al. (2012) proposed a model based on fuzzy time series to forecast the number of outpatient visits. Hadavandia et al. (2012) developed a hybrid AI model using genetic algorithm for outpatient visits forecasting. These hybrid models developed a new platform for hospital visits prediction.

To address the shortcomings of single traditional AI models, this study proposes a hybrid decomposition-and-ensemble learning paradigm incorporating wavelet decomposition (WD), artificial neural network (ANN) and the simple addition (ADD) to predict the number of patient visits to Chinese hospitals based on the principle of

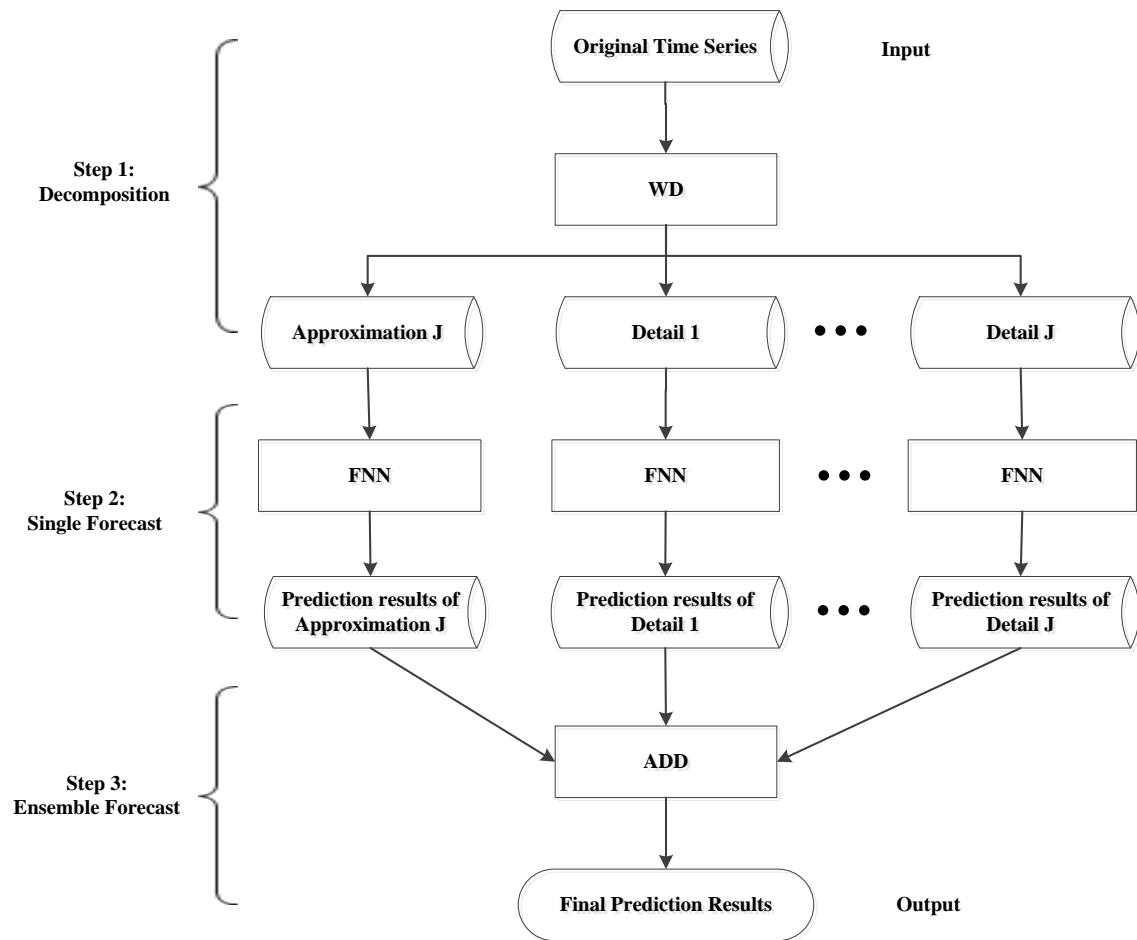


Figure 1. The overall process of the WD&ANN-based decomposition-and-ensemble learning paradigm

“decomposition and ensemble” (Yu et al., 2008; Wang et al., 2011; Wang, Tang, Yu, 2011; Tang et al., 2012; Yu et al., 2015; Tang et al., 2015). In this model, the WD is first employed to decompose the original monthly data of the number of patient visits to hospitals into several details and one approximation. Then, ANN, as a powerful prediction tool, is implemented to model each decomposed component and generate individual prediction results. Finally, all individual prediction values are fused into the final prediction output by the simple addition (ADD).

Generally speaking, due to the high irregularity and volatility of the number of patient visits to hospitals, the principle of “decomposition and ensemble” is introduced to cope with difficulties in modeling hospital visits. Consequently, a hybrid decomposition-and-ensemble learning paradigm integrating WD, ANN and ADD is proposed. In this proposed paradigm, WD, as a decomposition tool, is first performed to divide nonlinear, complicated hospital visits data into several relatively simple subsets of data, which are more easily processed further and thus bringing better forecasting performance.

The main motivation of this study is to propose a hybrid ensemble learning paradigm integrating WD, ANN and ADD to predict the number of patient visits to Chinese hospitals and compare its prediction performance with some extant popular forecasting techniques. The rest of the paper is organized as follows. Section 2 describes the proposed hybrid ensemble learning paradigm in detail. In Section 3, forecasting results and effectiveness of the proposed methodology are discussed. Finally, conclusions are drawn in Section 4.

METHOD

Due to volatility of hospital visits, a novel WD & ANN-based decomposition-and-ensemble approach is introduced to forecast the number of patient visits to hospitals. The general framework of the proposed hybrid learning paradigm can be illustrated in Figure 1. Generally speaking, there are three main steps involved. In this section, WD is first presented in Section 2.1 and ANN is then described in Section 2.2 Finally, the WD&ANN-based decomposition-and-ensemble algorithm is formulated in Section 2.3.

Wavelet Decomposition (WD)

Wavelet decomposition is proposed to overcome the drawbacks of Fourier transform. In Fourier transform, a signal is required to be periodic and frequencies are assumed not to evolve in time. To resolve the stationary assumption, short-time Fourier transform is proposed by applying a single fixed window, which leads to more cycles captured in a window with frequency increasing. Unlike short-time Fourier transform, the striking feature of wavelet transform lies in its adjustable windows. In wavelet transform, long window is used for low frequency while short window is for high frequency (Benhmad, 2011). In general, wavelet transform is superior in representing local details of a signal and capturing transient mutation in a signal, which results in its popularity in signal analysis, image processing, medical imaging and diagnosis, et al. For example, Geva and Kerem (1998) decomposed the electroencephalogram (EEG) signal via fast wavelet transform (FWT) to forecast generalized epileptic seizures. In Sierra et al.'s study (1997), an electrocardiogram signal was processed by wavelet multi-scale decomposition for further risk prediction.

Original time series data x_t ($t = 1, 2, \dots, T$) can be decomposed into several components by discrete wavelet transform at a given scale j , which can be expressed as

$$x_t = \sum_k c_k^j \Phi_k^j(t) + \sum_{j=1}^J \sum_k d_k^j \Psi_k^j(t) \tag{1}$$

where the former term is a low-frequency approximation containing scaling functions Φ_k^j (k is determined by data length and j .) with coefficients c_k^j . The other term consisting of mother wavelet functions $\Psi_k^j(t)$ with coefficients d_k^j in Eq.(1) is composed of a series of high-frequency details.

Artificial Neural Network (ANN)

With flexible function designs and powerful ability of self-learning, artificial intelligence (AI) is superior to traditional forecasting methods. In addition, artificial neural network (ANN) may be one of the most effective AI tools and as a typical learning paradigm, it is widely employed in forecasting of complex data. In this study, a standard three-layer feed-forward neural network (FNN) (Hornik et al., 1989; White, 1990), based on error back-propagation algorithm, is applied for modeling the number of visits in hospitals.

Usually, a FNN-based forecasting model can be trained by in-sample dataset and then applied to out-of-sample dataset for prediction. The model parameters (connection weights and node biases) are adjusted iteratively by a process of minimizing the forecasting error function. Basically, the final output of the FNN-based forecasting model can be represented as

$$f(x) = a_0 + \sum_{j=1}^q w_j \varphi \left(a_j + \sum_{i=1}^p w_{ij} x_i \right) \tag{2}$$

where x_i ($i = 1, 2, \dots, p$) represents the input patterns, $f(x)$ is the output, a_j ($j = 0, 1, 2, \dots, q$) is a bias on the j th unit, w_{ij} ($i = 1, 2, \dots, p; j = 1, 2, \dots, q$) and w_j are the connection weights between layers, $\varphi(\cdot)$ is the transfer function of the hidden layer, p is the number of input nodes, and q is the number of hidden nodes. Actually, the FNN model in Eq. (2) performs a nonlinear functional mapping from past observations $(x_{t-1}, x_{t-2}, \dots, x_{t-p})$ to future value x_{t-l} , i.e.

$$x_{t+l} = \Phi(x_{t-1}, x_{t-2}, \dots, x_{t-p}, W) + \xi_t \tag{3}$$

where l is the horizon, W is the vector of all parameters, and $\Phi(\cdot)$ is the function trained by FNN.

WD & ANN-based Decomposition-and-Ensemble Paradigm

Based on the techniques above, a hybrid WD&ANN-based decomposition-and-ensemble paradigm for patient volume is formulated as shown in [Figure 1](#).

With a given time series x_t ($t = 1, 2, \dots, T$), a multi-step-ahead prediction is performed in the form as follows:

$$\hat{x}_{t+m} = f(x_t, x_{t-1}, \dots, x_{t-(l-1)}) \tag{4}$$

where \hat{x}_{t+m} is the predicted value in period $t + m$, x_t is the actual value, l represents the lag order and m is the horizon.

There are three main steps in the proposed paradigm, as illustrated in [Figure 1](#).

Step 1: The original time series x_t ($t = 1, 2, \dots, T$) is decomposed into one low-frequency approximation and J high-frequency details by discrete wavelet transform (DWT).

Step 2: The components obtained after Step 1 are modeled respectively using FNN, a powerful forecasting tool, to produce corresponding predictions.

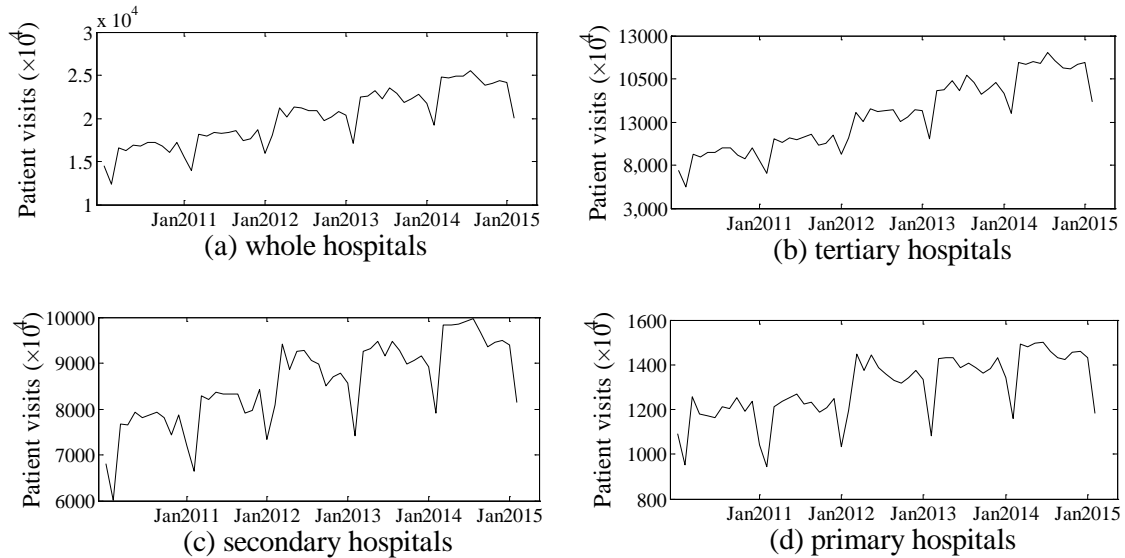


Figure 2. The original monthly number of visits to the (a) whole hospitals (b) tertiary hospitals (c) secondary hospitals (d) primary hospitals in China

Step 3: All the prediction results of wavelet decomposition components are fused into the final prediction for the original time series x_t via simple addition (ADD). Since WD decomposes the original data x_t into several components, the individual prediction results can be added up to produce the final result.

To summarize, the proposed WD&ANN-based decomposition-and-ensemble learning paradigm can be abbreviated as a “WD (decomposition)-FNN (prediction)-ADD (ensemble)” hybrid learning approach.

In order to verify the effectiveness of the proposed WD&ANN-based decomposition-and-ensemble learning paradigm, four datasets of visits in hospitals in China are targeted for testing in the next section.

EXPERIMENTAL STUDY

To verify the effectiveness of this proposed hybrid paradigm, four sets of data of Chinese hospital visits are utilized as sample data, and some benchmark models are used to compare with the proposed approach. This part, Section 3, comprises five major parts, including data description, evaluation criteria, benchmark models, parameter settings and results analyses.

Data Description

In this study, four different monthly series datasets of the number of patient visits to Chinese hospitals which are aggregates from the whole hospitals, tertiary hospitals, secondary hospitals and primary hospitals in China are chosen. These data are obtained from Wind Info. Each dataset ranges from January 2010 to February 2015 with 57 observations excluding December in every year as shown in **Figure 2**. The data from January 2010 to January 2013 are used as the training set (accounting for about 80% with 44 observations), and the remaining data are used as the testing set (with 13 observations). To test the forecasting capability of different models, one-, two- and three-step-ahead predictions are performed.

Evaluation Criteria

To evaluate forecasting accuracy, two criteria are applied and they are the root mean squared error (RMSE) and the mean absolute percent error (MAPE):

$$RMSE = \sqrt{\frac{1}{M} \sum_{t=1}^M (\hat{x}_t - x_t)^2} \tag{5}$$

$$MAPE = \frac{1}{M} \sum_{t=1}^M \left| \frac{x_t - \hat{x}_t}{x_t} \right| \tag{6}$$

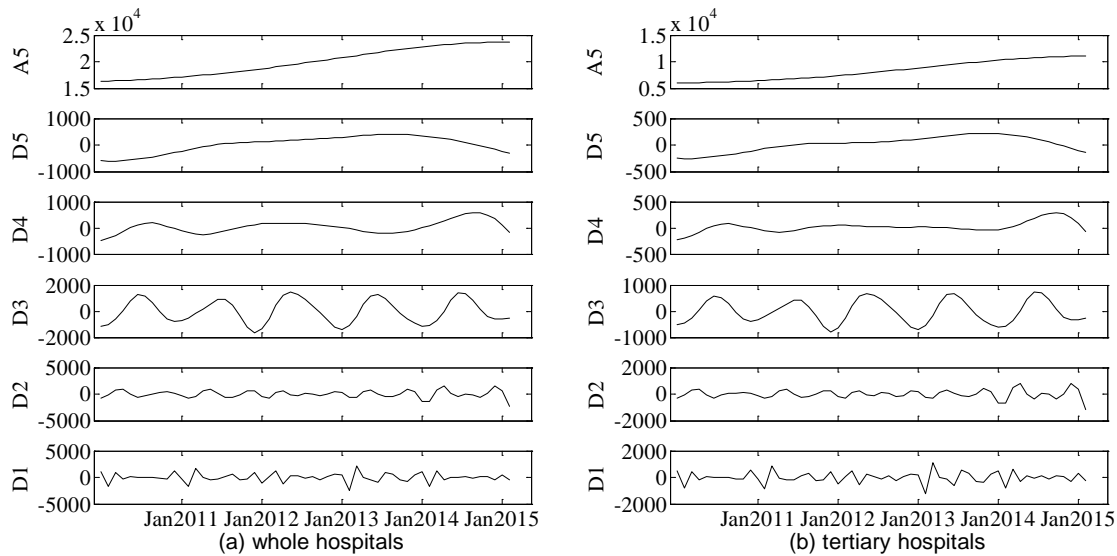


Figure 3. Decomposition result for the monthly number of patient visits to the(a) whole hospitals (b) tertiary hospitals via WD

where x_t is the actual value, \hat{x}_t is the predicted value and M is the number of observations in the testing dataset.

Benchmark Models

In order to test the effectiveness of the proposed WD&ANN-based decomposition-and-ensemble forecasting model, two types of traditional forecasting models and an artificial intelligence (AI) method are formulated as benchmarks.

For traditional forecasting models, auto-regressive integrated moving average (ARIMA) is quite popular (Box et al., 2011), and has been widely utilized in predicting hospital visits. Another traditional forecasting model, i.e., exponential smoothing (ES), is a kind of important forecasting technique and also shows up in hospital visits forecasting (Bergs et al., 2014; Morzuch, 2006). So the two methods mentioned above are implemented.

For artificial intelligence (AI), the most typical AI technique, i.e., feed-forward neural network (FNN), is embraced in this study as a benchmark model.

In general, for the proposed WD-FNN-ADD decomposition-and-ensemble model, two traditional methods (i.e., ARIMA and exponential smoothing) and one artificial intelligence (AI) method (i.e., FNN) are formulated to be compared with.

Parameter Settings

For traditional methodology, the ARIMA models as a benchmark in this study are determined based on autocorrelation and partial correlation analysis. Parameters of exponential smoothing models are set based on the Standard Error minimization principle. For AI technique, each FNN built in this study uses seven nodes in the hidden layer, one output neuron and I input neurons, where I is the lag order determined by autocorrelation analysis and finally set to 11. FNNs run iteratively 10000 times using the training subsets.

For the WD&ANN-based hybrid model, the first step is to decompose the original data into an approximation and a series of details via wavelet decomposition (WD) where Daubechies wavelet $db8$ is used. In accordance with the length of each dataset, the WD (Daubechies, 1992; Mallat, 1989) decomposition level is set to 5, which means WD decomposes the data into one approximation and five details, as shown respectively in Figures 3-4. The second step of the hybrid model is to forecast the approximation and all the details by FNN, where for consistency purpose FNNs are formed in the similar way as those applied as benchmark models. In the third step, the application of the ADD ensemble method, simply summing together those individual prediction results obtained in the second step, contributes to the production of the final results.

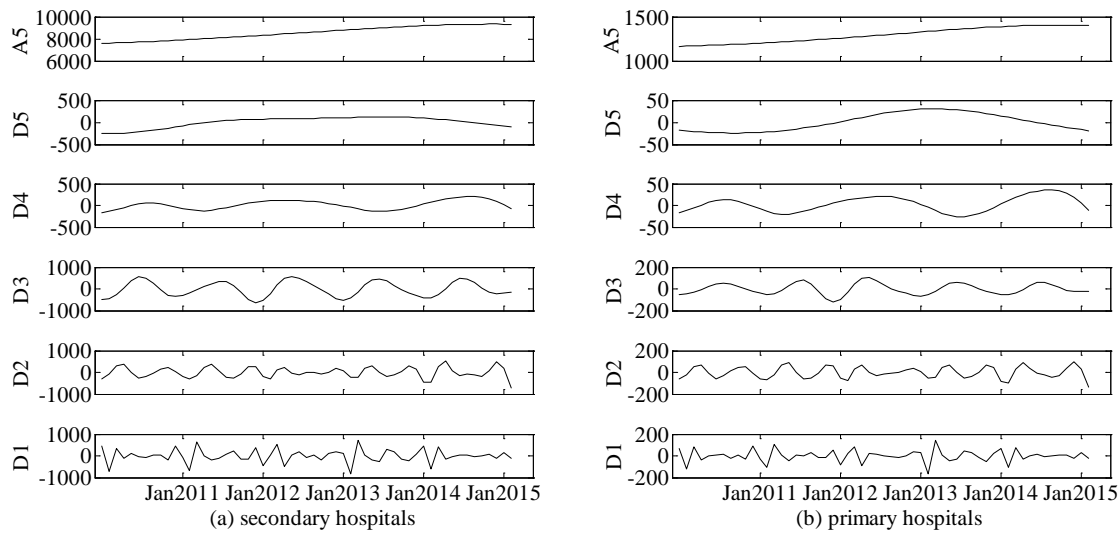


Figure 4. Decomposition result for the monthly number of patient visits to the (a) secondary hospitals (b) primary hospitals via WD

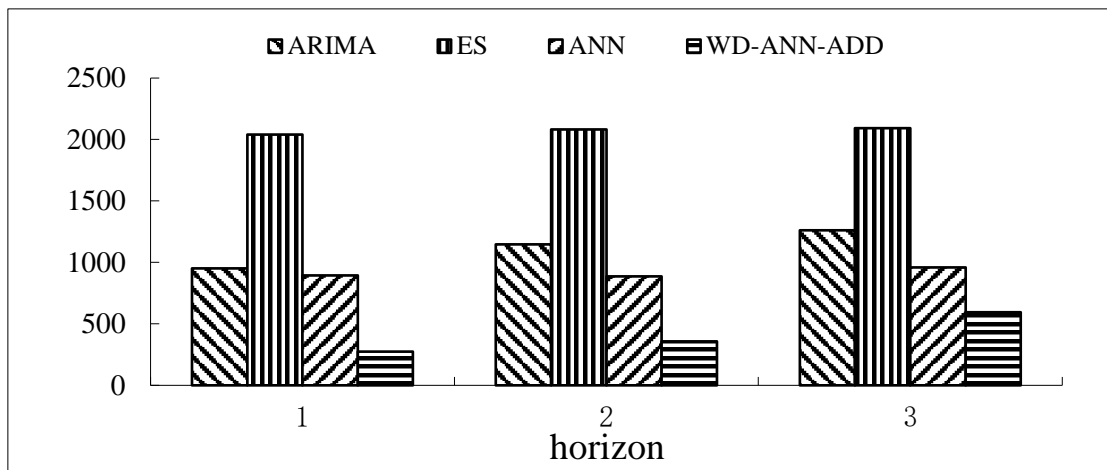


Figure 5. RMSE comparison of different models for the number of patient visits to the whole hospitals in China

Results Analyses

The prediction results of the four models (i.e., two traditional models ARIMA and ES, one artificial technique FNN, and the proposed hybrid model WD-FNN-ADD) for four diverse sets of hospital visits data are illustrated in **Figures 5-12**. **Figures 5-6, 7-8, 9-10** and **11-12** respectively show the forecasting performance comparison results for the whole, tertiary, secondary and primary hospital visits from the aspects of RMSE and MAPE criteria. From the results, it is obvious that the proposed method WD-ANN-ADD outperforms all the three benchmark models in all cases, in terms of RMSE and MAPE.

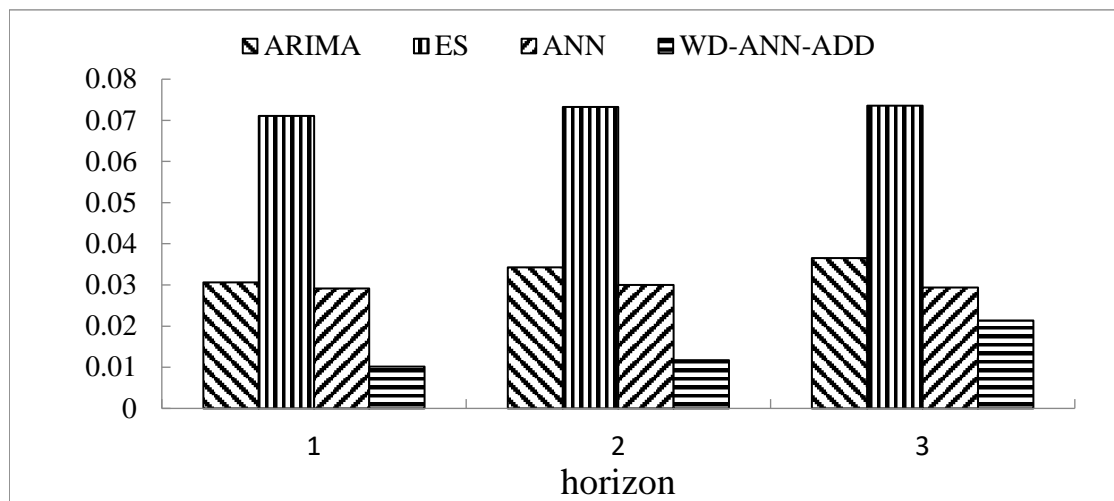


Figure 6. MAPE comparison of different models for the number of patient visits to the whole hospitals in China

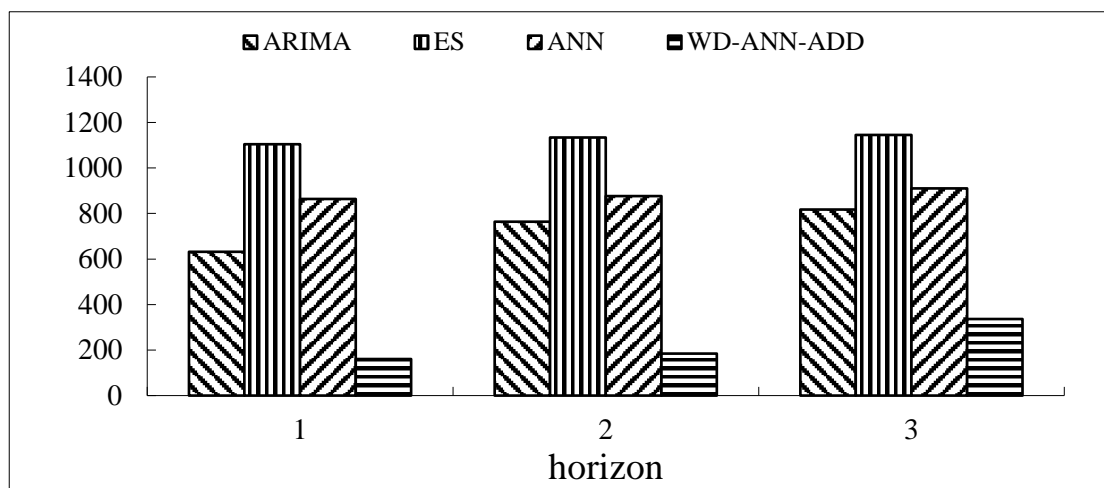


Figure 7. RMSE comparison of different models for the number of patient visits to the tertiary hospitals in China

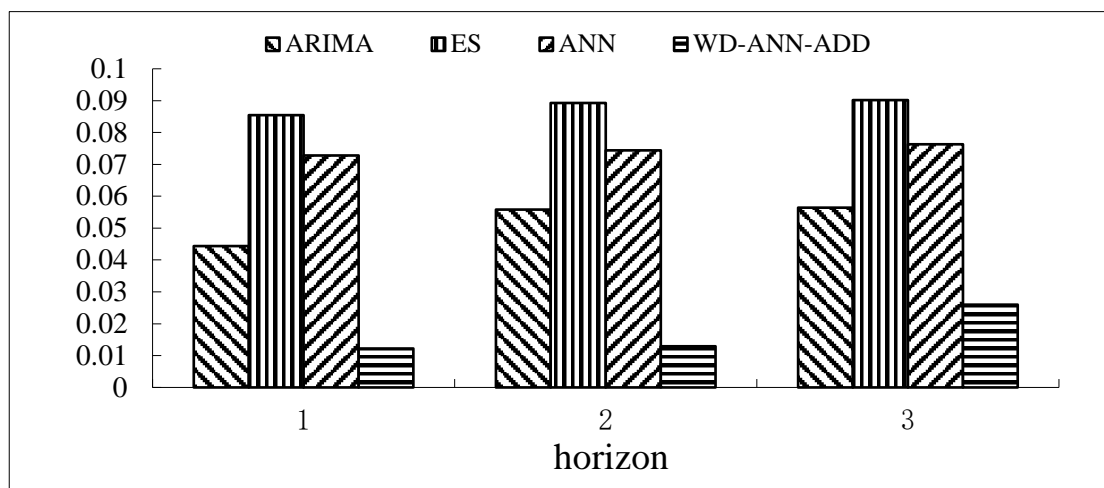


Figure 8. MAPE comparison of different models for the number of patient visits to the tertiary hospitals in China

The results of RMSE and MAPE criteria for four separate types of hospital visits reveal that the proposed WD-ANN-ADD model performs the best in all cases compared to the other three models while the exponential smoothing (ES) method always shows the worst prediction accuracy without exception. As for ARIMA and ANN, It is hard to define whose performance in forecasting is better for the situation differs when the experimental dataset renews and more details are described in the following content.

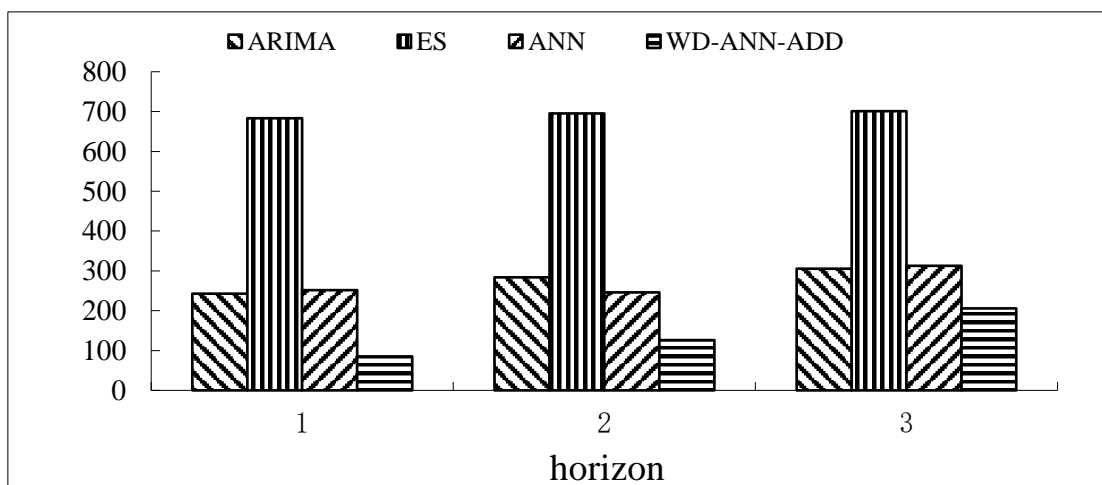


Figure 9. RMSE comparison of different models for the number of patient visits to the secondary hospitals in China

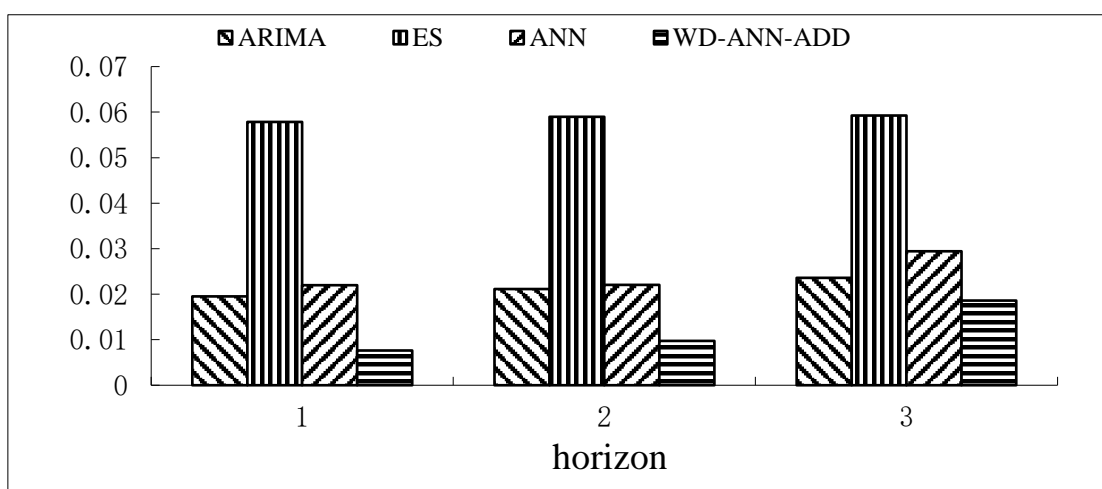


Figure 10. MAPE comparison of different models for the number of patient visits to the secondary hospitals in China

The fact that the hybrid WD-FNN-ADD model strikingly surpasses all the other three benchmarks in multi-step-ahead prediction accuracy, in terms of RMSE and MAPE criteria, displays the superiority of the “decomposition and ensemble” strategy in prediction accuracy improvement and also confirms the effectiveness of this novel paradigm in improving forecasting accuracy. In the light of the hybrid method’s (i.e., WD-FNN-ADD) absolute excellent performance in prediction, a conclusion can be drawn: wavelet composition (WD) does play an important role in the whole experimental process. Wavelet decomposition as a common signal processing tool can decompose the original complex data into several components that can be more easily processed further, which explains the importance of wavelet decomposition in the hybrid model, i.e., WD-FNN-ADD.

Both the RMSE and MAPE of the exponential smoothing model (ES) prediction in the context of four different sets of hospital visits data are much higher than the others’ (i.e., ARIMA, ANN and WD-ANN-ADD model). Due to volatility of the monthly number of patient visits to hospitals and inconsistency of data structure, it is hard to model each dataset exactly via an exponential smoothing method with a sole parameter.

In this study, which one of ARIMA and ANN has better performance in modeling the number of patient visits hospitals can’t be hastily determined. ANN outperforms ARIMA a little, in terms of RMSE and MAPE criteria, in forecasting the number of visits in the whole Chinese hospitals as shown in Figures 5-6. From Figures 7-8 and Figures 11-12, it is easy to find that ARIMA demonstrates better results for both the Chinese tertiary and secondary hospital visits prediction than ANN does. Here comes the Chinese secondary hospital visits forecasting. From Figures 9-10, almost all the values of RMSE and MAPE of ARIMA prediction are lower than ANN’s except the only one RMSE value of the two-step-ahead Chinese secondary hospital visits prediction as shown in Figure 9. The hidden reason may lie in data characteristics and this should be discussed in the near future.

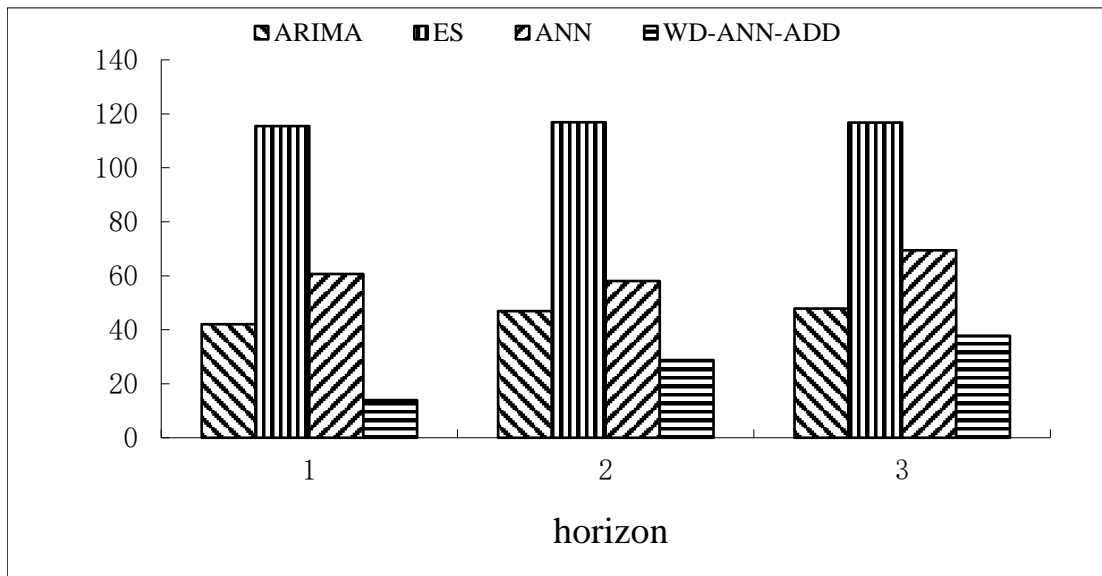


Figure 11. RMSE comparison of different models for the number of patient visits to the primary hospitals in China

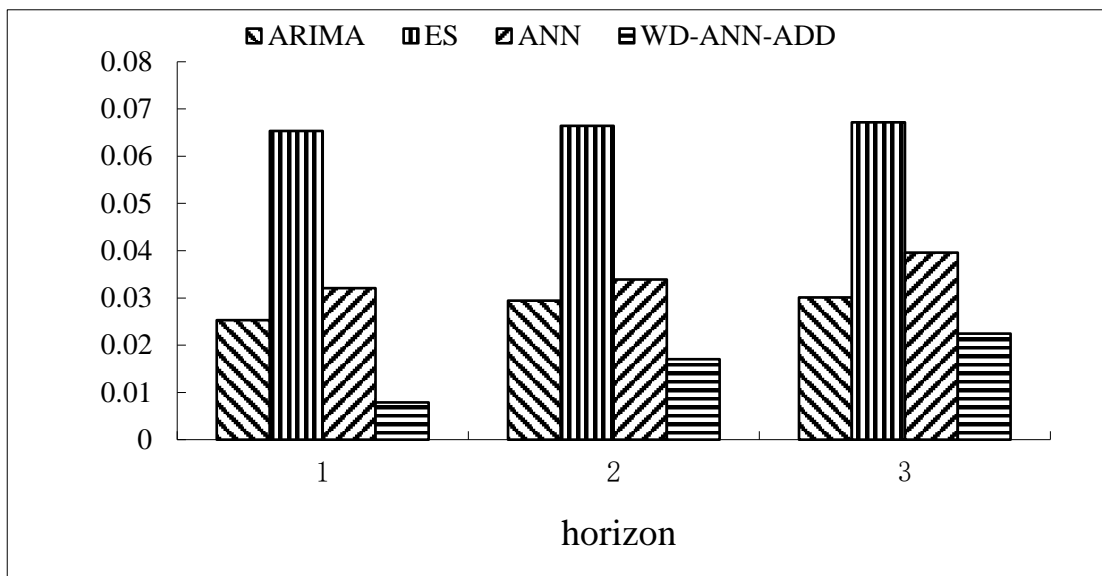


Figure 12. MAPE comparison of different models for the number of patient visits to the primary hospitals in China

Generally, four main conclusions can be drawn according to the analyses above. (1) The proposed WD-based FNN hybrid decomposition-and-ensemble model, i.e., WD-FNN-ADD, is significantly superior to all the benchmark models, i.e., ARIMA, exponential smoothing and ANN, in terms of RMSE and MAPE comparison of one-, two- and three-step-ahead forecasting. (2) The proposed hybrid approach remarkably outperforms all the other models listed in this study in all cases of the whole, tertiary, secondary and primary hospital visits prediction, revealing that the principle of “decomposition and ensemble” can effectively improve forecasting precision in the context of hospital visits and requiring more emphasis in the significant status of wavelet decomposition in the hybrid model in the frame of “decomposition and ensemble”. (3) That general exponential smoothing method performs the worst among all the methods in this study indicates the efficacy loss of this method in predicting mutable and irregular data. (4) The proposed WD-based FNN decomposition-and-ensemble forecasting approach, with the framework of “decomposition and ensemble” and the powerful WD decomposition technique, is a rather effective and promising model for forecasting the number of visits in hospitals.

CONCLUSIONS

In order to address the difficulties of complex hospital visits data, a hybrid decomposition-and-ensemble learning paradigm is proposed, by integrating wavelet decomposition (WD), feed-forward neural network (FNN) and simple addition (ADD), i.e., WD-FNN-ADD. The main contribution of the proposed hybrid approach can be

summered into two aspects. First, different from other traditional models, a much more powerful method with its merits of handling the problem of data complexity, i.e., WD-FNN-ADD, is employed in hospital visits forecasting. Second, to the best of our knowledge, although the principle of “decomposition and ensemble” has been widely applied in time series forecasting, e.g., crude oil price (Garg et al., 2012; Tang et al., 2012; Yu et al., 2014; Yu et al., 2015), hydropower consumption (Yu et al., 2011; Wang et al., 2011) and nuclear consumption forecasting (Wang et al., 2011), such a hybrid decomposition-and-ensemble model, i.e., WD-FNN-ADD, hasn't been introduced to hospital visits analysis or forecasting research. Therefore, the gap is filled via this study applying WD-FNN-ADD to the field of hospital visits forecasting.

With the four datasets of the monthly number of patient visits to the whole, tertiary, secondary and primary hospitals in China as sample data, the experimental results indicate that the proposed hybrid WD-FNN-ADD learning paradigm can effectively improve the forecasting precision, in terms of RMSE and MAPE criteria, since WD-FNN-ADD model outperforms all the benchmark models. The proposed WD-FNN-ADD model is a very promising tool in forecasting complex time series data of hospital visits.

Besides predicting the number of patient visits to hospitals, the proposed WD-FNN-ADD approach can also be utilized to address other difficult forecasting tasks, especially for complex, volatile data with multiple data characteristics. Moreover, how to select the most appropriate wavelet in wavelet decomposition is extremely crucial, which exerts an influence on the decomposition results and accordingly has an impact on forecasting results. We will look into these issues in the near future.

ACKNOWLEDGEMENTS

This work is supported by grants from the National Natural Science Foundation of China (NSFC Nos. 71622011, 71433001 and 71301006), the National Program for Support of Top Notch Young Professionals, Beijing Advanced Innovation Center for Soft Matter Science and Engineering, and the Theme Based Research Grant, RGC of Hong Kong (TRS Project 8770001).

REFERENCES

- Abdel-Aal, R. E., & Mangoud, A. M. (1998). Modeling and forecasting monthly patient volume at a primary health care clinic using univariate time-series analysis. *Computer Methods & Programs in Biomedicine*, 56(3), 235–247.
- Abraham, G., Byrnes, G. B., & Bain, C. (2009). Short-term forecasting of emergency inpatient flow. *Information Technology in Biomedicine, IEEE Transactions on*, 13(3), 380–388.
- Aladag, C. H., & Aladag, S. (2012). Forecasting the number of outpatient visits with different activation functions. *Advances in Time Series Forecasting*, 26.
- Au-Yeung, S. W., Harder, U., McCoy, E. J., Knottenbelt, W. J. (2009). Predicting patient arrivals to an accident and emergency department. *Emergency Medicine Journal*, 26(4), 241–244.
- Batal, H., Tench, J., Mcmillan, S., Adams, J., & Mehler, P. S. (2001). Predicting patient visits to an urgent care clinic using calendar variables. *Academic Emergency Medicine Official Journal of the Society for Academic Emergency Medicine*, 8(1), 48–53.
- Benhmad, F. (2011). A wavelet analysis of oil price volatility dynamic. *Economics Bulletin*, 31(1), 792–806.
- Bergs, J., Heerinx, P., & Verelst, S. (2014). Knowing what to expect, forecasting monthly emergency department visits: A time-series analysis. *International Emergency Nursing*, 22(2), 112–115.
- Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2011). *Time series analysis: forecasting and control*. John Wiley & Sons.
- Boyle, J., Wallis, M., Jessup, M., Crilly, J., Lind, J., Miller, P., & Fitzgerald, G. (2008). Regression forecasting of patient admission data. *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*. IEEE, 3819–3822.
- Brillman, J. C., Burr, T., Forslund, D., Joyce, E., Picard, R., & Umland, E. (2005). Modeling emergency department visit patterns for infectious disease complaints: results and application to disease surveillance. *BMC Medical Informatics & Decision Making*, 5:4(1), 1–14.
- Chen, C. F., Ho, W. H., Chou, H. Y., Yang, S. M., Chen, I. T., & Shi, H. Y. (2011). Long-term prediction of emergency department revenue and visitor volume using autoregressive integrated moving average model. *Computational & Mathematical Methods in Medicine*, 2011(3), 713–750.
- Cheng, C. H., Wang, J. W., & Li, C. H. (2008). Forecasting the number of outpatient visits using a new fuzzy time series based on weighted-transitional matrix. *Expert Systems with Applications*, 34(4), 2568–2575.
- Dan, T., & Qualls, C. (1994). Time series forecasts of emergency department patient volume, length of stay, and acuity. *Annals of Emergency Medicine*, 23(2), 299–306.

- Daubechies, I. (1992). *Ten lectures on wavelets*. Philadelphia: Society for industrial and applied mathematics.
- Diaz, J., Alberdi, J. C., & Pajares, M. S. (2001). A model for forecasting emergency hospital admissions: effect of environmental variables. *Journal of Environmental Health*, 64(3), 9-15.
- Flottemesch, T. J., Gordon, B. D., Jones, S. S. (2007). Advanced Statistics: Developing a formal model of emergency department census and defining operational efficiency. *Academic Emergency Medicine Official Journal of the Society for Academic Emergency Medicine*, 14(9), 799-809.
- Friede, K. A., Osborne, M. C., Erickson, D. J., Roesler, J. S., Azam, A., Croston, J. K., McGonigal, M. D., & Ney, A. L. (2009). Predicting trauma admissions: the effect of weather, weekday, and other variables. *Minnesota medicine*, 92(11), 47-49.
- Garg, B., Beg, M. M. S., & Ansari, A. Q. (2012). A new computational fuzzy time series model to forecast number of outpatient visits. Fuzzy Information Processing Society (NAFIPS), 2012 Annual Meeting of the North American. IEEE, 1-6.
- Geva, A. B., Kerem, D. H. (1998). Forecasting generalized epileptic seizures from the EEG signal by wavelet analysis and dynamic unsupervised fuzzy clustering. *Biomedical Engineering, IEEE Transactions on*, 45(10), 1205-1216.
- Hadavandi, E., Shavandi, H., Ghanbari, A., Abbasian-Naghneh, S. (2012). Developing a hybrid artificial intelligence model for outpatient visits forecasting in hospitals. *Applied Soft Computing*, 12(2), 700-711.
- Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feed-forward networks are universal approximators. *Neural Networks*, 2(89), 359-366.
- Jones, S. S. (2009). A multivariate time series approach to modeling and forecasting demand in the emergency department. *Journal of Biomedical Informatics*, 42(1), 123-139.
- Jones, S. S., Thomas, A., Evans, R. S., Welch, S. J., Haug, P. J., & Snow, G. L. (2008). Forecasting daily patient volumes in the emergency department. *Academic Emergency Medicine*, 15(2), 159-170.
- Kadri, F., Harrou, F., Chaabane, S., & Tahon, C. (2014). Time series modelling and forecasting of emergency department overcrowding. *Journal of Medical Systems*, 38(9), 490-491.
- Kam, H. J., Sung, J. O., & Park, R. W. (2010). Prediction of daily patient numbers for a regional emergency medical center using time series analysis. *Healthcare Informatics Research*, 16(3), 158-165.
- Kim, K., Lee, C., O'Leary, K., Rosenauer, S., & Mehrotra, S. (2014). *Predicting Patient Volumes in Hospital Medicine: A Comparative Study of Different Time Series Forecasting Methods*. Tech. rep., Northwestern University.
- Koestler, D. C., Ombao, H., & Bender, J. (2013). Ensemble-based methods for forecasting census in hospital units. *BMC Medical Research Methodology*, 13(6), 454-462.
- Kovats, R. S., Hajat, S., & Wilkinson, P. (2004). Contrasting patterns of mortality and hospital admissions during hot weather and heat waves in Greater London, UK. *Occupational & Environmental Medicine*, 61(11), 893-898.
- Littig, S. J., & Isken, M. W. (2007). Short term hospital occupancy prediction. *Health Care Management Science*, 10(1), 47-66.
- Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 11(7), 674-693.
- Marcilio, I., Hajat, S., & Gouveia, N. (2013). Forecasting daily emergency department visits using calendar variables and ambient temperature readings. *Academic Emergency Medicine*, 20(8), 769-777.
- Milner, P. C. (1997). Ten-year follow-up of ARIMA forecasts of attendances at accident and emergency departments in the Trent region. *Statistics in Medicine*, 16(18), 2117-2125.
- Morzuch, B. J., & Allen, P. G. (2006). Forecasting hospital emergency department arrivals. *26th Annual Symposium on Forecasting*, Santander, Spain.
- Pan, W. T. (2017). A Newer Equal Part Linear Regression Model: A Case Study of the Influence of Educational Input on Gross National Income. *Eurasia Journal of Mathematics, Science & Technology Education*, 13(8), 5765-5773.
- Reis, B. Y., & Mandl, K. D. (2003). Time series modeling for syndromic surveillance. *BMC Medical Informatics & Decision Making*, 3(1), 1-11.
- Rotstein, Z., Wilf-Miron, R., Lavi, B., Shahar, A., Gabbay, U., & Noy, S. (1997). The dynamics of patient visits to a public hospital ED: a statistical model. *American Journal of Emergency Medicine*, 15(6), 596-599.
- Safar, A., & Alkhezzi, F. (2016). Students' perspectives of the impact of online streaming media on teaching and learning at the college of education at Kuwait University. *Eurasia Journal of Mathematics, Science & Technology Education*, 12(12).
- Schweigler, L. M., Desmond, J. S., McCarthy, M. L., Bukowski, K. J., Ionides, E. L., & Younger, J. G. (2009). Forecasting models of emergency department crowding. *Academic Emergency Medicine*, 16(4), 301-308.

- Sierra, G., Reinhardt, L., Fetsch, T., Martinez-Rubio, A., Makijarvi, M., Yli-Mayry, S., Montonen, J., Katila, T., Borggreffe, M., & Breithardt, G. (). Risk stratification of patients after myocardial Infarction based on wavelet decomposition of the signal-averaged Electrocardiogram. *Annals of Noninvasive Electrocardiology*, 2(1), 47-58.
- Sun, Y., Heng, B. H., Seow, Y. T., et al. Forecasting daily attendances at an emergency department to aid resource planning. *BMC Emergency Medicine*, 9(1), 1.
- Tang, L., Dai, W., Yu, L., & Wang, S. (). A novel CEEMD-based EELM ensemble learning paradigm for crude oil price forecasting. *International Journal of Information Technology & Decision Making*, 14(01), 141-169.
- Tang, L., Yu, L., Wang, S., Li, J., & Wang, S. (). A novel hybrid ensemble learning paradigm for nuclear energy consumption forecasting. *Applied Energy*, 93, 432-443.
- Wang, S., Tang, L., & Yu, L. (2011). SD-LSSVR-based decomposition-and-ensemble methodology with application to hydropower consumption forecasting. *Proceedings of the 2011 Fourth International Joint Conference on Computational Sciences and Optimization*. IEEE Computer Society, 603-607.
- Wang, S., Yu, L., Tang, L., & Wang, S. (2011). A novel seasonal decomposition based least squares support vector regression ensemble learning approach for hydropower consumption forecasting in China. *Fuel & Energy Abstracts*, 36(11), 6542-6554.
- White, H. (1990). Connectionist nonparametric regression: multilayer feedforward networks can learn arbitrary mappings. *Neural Networks*, 3(90), 535-549.
- Xu, M., Wong, T. C., & Chin, K. S. (2013). Modeling daily patient arrivals at Emergency Department and quantifying the relative importance of contributing variables using artificial neural network. *Decision Support Systems*, 54(3), 1488-1498.
- Yu, L., Dai, W., Tang, L. (2015). A novel decomposition ensemble model with extended extreme learning machine for crude oil price forecasting. *Engineering Applications of Artificial Intelligence*. <http://dx.doi.org/10.1016/j.engappai.2015.04.016>
- Yu, L., Wang, S., & Lai, K. K. (2008). Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. *Energy Economics*, 30(5), 2623-2635.
- Yu, L., Zhao, Y., & Tang, L. (2014). A compressed sensing based AI learning paradigm for crude oil price forecasting. *Energy Economics*, 46(C), 236-245.
- Zibners, L. M. (2006). Local weather effects on emergency department visits: a time series and regression analysis. *Pediatric Emergency Care*, 22(2), 104-106.

<http://www.ejmste.com>