Original Research Paper

Hydrological Forecasting Using Hybrid Data-Driven Approach

[1]Youngmin Seo and [2]Sungwon Kim

[1]*Department of Constructional Environmental Engineering, Kyungpook National University, Sangju, 37224, South Korea*
[2]*Department of Railroad and Civil Engineering, Dongyang University, Yeongju, 36040, South Korea*

Abstract: This study develops a hybrid model, EEMD-FANN, coupling feed Forward Artificial Neural Network (FANN) and Ensemble Empirical Mode Decomposition (EEMD) for improving the accuracy of daily river stage forecasting. An original river stage data is broken down into a residue and Intrinsic Mode Functions (IMFs) using the EEMD and different FANNs are developed as forecasting models for the decomposed IMFs and residue, respectively. The final forecasted time series is produced by the ensemble aggregation of the forecasted IMFs and residue. The efficiency of EEMD-FANN model is assessed based on the comparison with that of single Adaptive Neuro-Fuzzy Inference System (ANFIS) and FANN to demonstrate the applicability of the hybrid approach in daily river stage forecasting. As a result, it is found that the EEMD-FANN model utilizing time series decomposition by the EEMD and ensemble aggregation produces better performance than the single ANFIS and FANN models using original river stage time series as inputs. The results of this study also signify that the approach coupling the EEMD and FANN can significantly enhance the forecasting ability of the single FANN model and can be utilized as an effective modeling methodology to forecast river stage precisely.

Keywords: River Stage Forecasting, Adaptive Neuro-Fuzzy Inference System, Feed Forward Artificial Neural Network, Ensemble Empirical Mode Decomposition

Introduction

Forecasting river stage precisely plays an important role for enhancing hydrologic practices such as dam operation, water supply, river management and flood and drought prevention. For decades, the attention on data-driven approaches has been gained for estimating hydrological variables including rainfall-runoff, river discharge, soil moisture, evaporation, reservoir water level, etc. Especially, Support Vector Machine (SVM), ANFIS and FANN have been received attention as effectual approaches to analyze complicated and nonlinear hydrologic phenomena (Kisi, 2007; Kim and Kim, 2008; Wu *et al*., 2009; Othman and Naseri, 2011; Kim *et al*., 2012; 2013; Kaltech, 2015; Seo *et al*., 2015b).

The development of hybrid models combining various statistical approaches and data-driven models has been increased for improving the efficiency of conventional forecasting models. Especially, time series decomposition utilizing wavelet and wavelet packet transforms has been known to further improve the forecasting ability of conventional data-driven models (Amiri and Asadi, 2009; Adamowski and Sun, 2010; Gokhale and Khanduja, 2010; Kisi *et al*., 2011; Nourani *et al*., 2012; Ravikumar and Tamilselvan, 2014; Seo, 2015; Seo *et al*., 2015a).

Recently, nonlinear data analysis and hybrid model development utilizing Empirical Mode Decomposition (EMD)-based approaches have been successfully performed in various fields. The EMD, which is a self-adaptive and empirical technique to break down a time series, can be utilized to examine nonlinear and nonstationary meteorological and hydrologic data (Tang *et al*., 2012). The EEMD is a data processing technique based on noise addition (Wu and Huang, 2009) which is devised to improve the EMD. Huang *et al*. (2009) applied EMD and signal analysis based on Hilbert transform (Hilbert spectral analysis) to investigate the characteristics of nonlinear river flow time series. Karthikeyan and Kumar (2013) assessed the forecasting ability of nonstationary time series applying forecasting models which are based on wavelet and EMD. Kisi *et al*. (2014) presented a nonparametric

method which is to build up a model combining Artificial Neural Network (ANN) and EMD to forecast monthly river stage. Huang *et al*. (2014) examined a conjunction model which combines modified EMD and SVM in monthly stream flow prediction. Shabri and Samsudin (2015) proposed a hybrid modeling methodology utilizing Least Square Support Vector Machine (LSSVM) combined with EMD to predict water demand. Wang *et al*. (2015) suggested a time series forecasting model combining autoregressive integrated moving average and EEMD to predict runoff on annual basis. Zhao and Chen (2015) proposed a novel hybrid model based on EEMD and autoregressive forecasting model for annual runoff forecasting. Wang *et al*. (2013) presented an approach which includes the decomposition of annual rainfall series utilizing EEMD and rainfall-runoff modeling utilizing SVM. Guo *et al*. (2016) applied EEMD to examine the intrinsic multi-scale properties which are inherent in the variability of precipitation. Ouyang *et al*. (2016) proposed a hybrid modeling approach using EEMD, Support Vector Regression (SVR) and phase-space reconstruction for monthly rainfall forecasting. Xu *et al*. (2016) presented a conjunction model integrating EEMD, Back-Propagation (BP)-based ANN (BPANN) and nonlinear regression equation for annual runoff simulation.

This study presents a hybrid modeling approach, EEMD-FANN, coupling EEMD and FANN to enhance the model performance in river stage forecasting. For accessing the efficiency of EEMD-FANN, the model accuracy was evaluated utilizing some favorite statistical performance indexes. The comparative analysis of EEMD-FANN and single forecasting models, ANFIS and FANN, was also implemented for assessing the applicability of EEMD-FANN.

Materials and Methods

Used Data

The Andong Dam watershed, located in the upper Nakdong river basin of the southeastern inland region of South Korea, was selected as a study area to forecast daily river stage utilizing EEMD-FANN model. The time series data used for developing the model were gathered from two hydrological observatories (Socheon and Dosan) in the watershed (Fig. 1) utilizing the Water Management Information System (WAMIS) which is an internet-based portal system providing the water resources information of South Korea.

Figure 1 depicts the study area andong Dam watershed, including the structure of stream network and the location of hydrological observatories. The time series data of daily river stage were collected between 2002 and 2013 (12 years). For model training and testing, they were split into two sub-periods, data for 2002-2010 (9 years, 75%) and data for 2011-2013 (3 years, 25%).
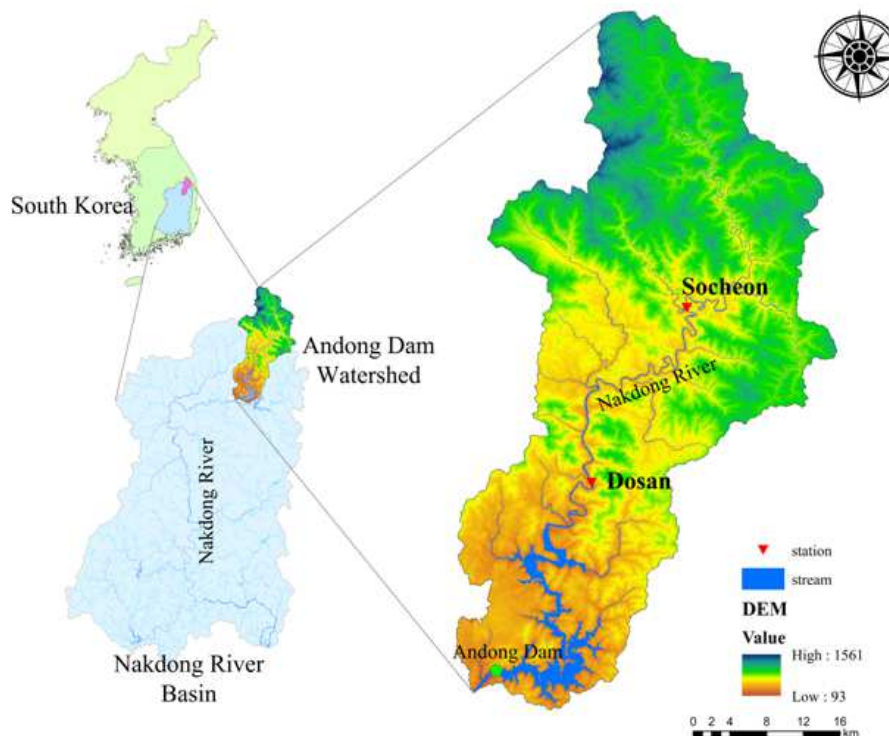

Fig. 1. Andong Dam watershed and hydrological observatories

EMD

EMD, which was originally suggested by Huang *et al.* (1998), is a data analysis technique breaking down a nonlinear and nonstationary signal into several components. Unlike other approaches such as wavelet and Fourier transforms, the technique includes empirical and self-adaptive properties (Tang *et al.*, 2012). The key concept of the EMD is to break down an input data series into two components: A residue and Intrinsic Mode Functions (IMFs) (Yu *et al.*, 2008). Based on the concept proposed by Huang *et al.* (1998), the IMFs should comply with the following requirements:

- In data set, the number of zero-crossings and local extreme values (local maximum and minimum values) should be identical each other, or the difference of the number should be one at most
- The mean values of two envelopes, which are created by interpolating local maximum and minimum values, respectively, should be zero at all data points

The EMD algorithm, which is also called the shifting procedure, is summarized as the following key steps (Huang *et al.*, 1998; 2003):

- Detect local minimal and maximal values in a time series, $x(t)$
- Interpolate the local minimal values for creating the lower envelope, $L(t)$ and the local maximal values for upper envelopes, $U(t)$. In this step, interpolation method such as cubic spline can be applied for creating the envelope lines
- Compute the mean, $M(t)$, of the envelope lines
- Calculate the detail component, $D(t)$, by subtracting $M(t)$ from $x(t)$
- Check the conditions of $D(t)$: (i) if $D(t)$ satisfies the two requirements of IMF, an IMF is determined and $x(t)$ is substituted with the residue. The residue, $R(t)$, is obtained by subtracting $D(t)$ from $x(t)$; (ii) if $D(t)$ does not meet the requirements, $x(t)$ is replaced with $D(t)$
- Perform the phases (1)-(5) repeatedly, until the IMF cannot be derived any more as $R(t)$ has a local extreme point or becomes monotonic.

Once an EMD process is completed, the sum of the final residue, $R_k(t)$ and IMFs, $C_j(t)$, can produce the original time series as in Equation 1:

$$x(t) = \sum_{j=1}^{k} C_j(t) + R_k(t) \tag{1}$$

where, k denotes the number of IMFs. For detailed information on EEMD, readers can refer to Huang *et al.* (1998; 2003).

EEMD

EEMD, which is a data analysis technique proposed by Wu and Huang (2009), is developed to resolve the problem of mode mixing which is an obvious disadvantage of EMD. The mode mixing indicates that an IMF is made up of signal covering the broad bandwidth of frequency, or signals in a similar frequency band are contained in one or more IMF (Ren *et al.*, 2015). The key point of EEMD is signal decomposition utilizing noise addition. The EEMD is based on the concept that a signal is formed by adding Additive White Gaussian Noise (AWGN) to true value and the ensemble mean of decomposed time series with different white noise series yields better estimates of true time series. According to Wu and Huang (2009), the EEMD algorithm is comprised of four key phases:

- Generate an AWGN series and add it to original time series
- Break down the time series obtained from step (1) into IMFs and a residue using EMD
- Perform the phases (1)-(2) repeatedly, but use different AWGN at each iteration
- Calculate the ensemble means of residue and IMFs for each component to obtain the final decomposed time series

For more detailed information on EEMD, readers can refer to Wu and Huang (2009).

FANN

FANN is a data-driven modeling system which is mathematically emulated based on the architecture and function of the neural system of human brain. Multilayer Perceptron (MLP) type of FANN is generally comprised of input, output and hidden layers. For example, FANN architecture with five input neurons, 13 hidden neurons and an output neuron can be depicted schematically as in Fig. 2. According to Günther and Fritsch (2010), the MLP composed of a hidden layer with k neurons calculates the output based on the following equation:

$$O(\mathbf{x}) = f\left(w_0 + \sum_{j=1}^{k} w_j f\left(w_{0j} + \sum_{i=1}^{n} w_{ij} x_i \right) \right)$$
$$= f\left(w_0 + \sum_{j=1}^{k} w_j f\left(w_{0j} + \mathbf{w}_j^T \mathbf{x} \right) \right) \tag{2}$$

where, w_0 and w_{0j} are respectively the intercepts for the neuron of output layer and the jth neuron of hidden layer, w_j is the connection strength between the jth neuron of hidden layer and the output neuron of output layer, $\mathbf{w}_j = (w_{1j}, \cdots, w_{nj})$ is the set of connection strengths between the neurons of input layer and the jth neuron of hidden layer and $x = (x_1, \cdots, x_n)$ is the input vector.
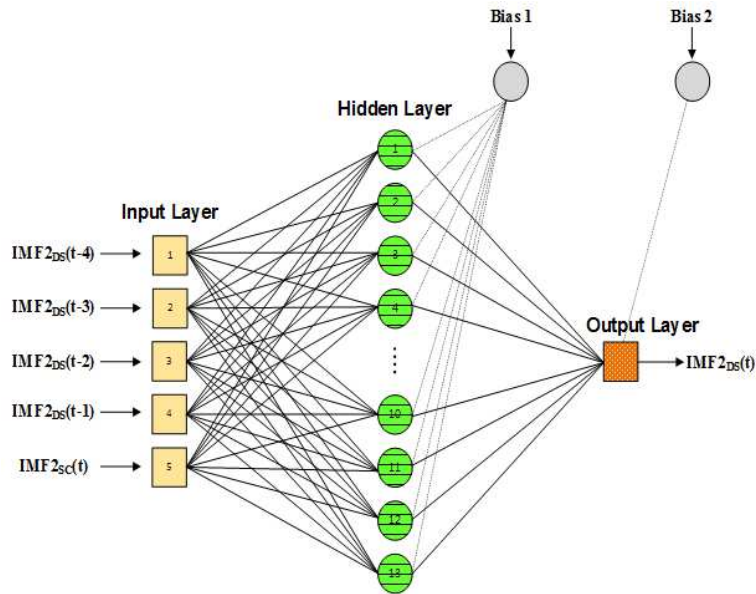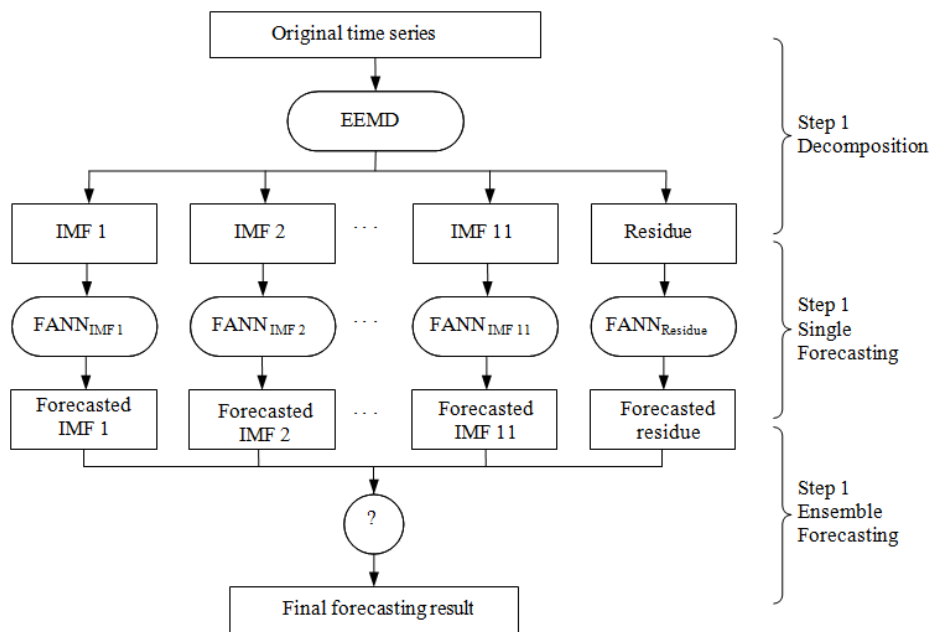
Fig. 2. FANN architecture for IMF2

Fig. 3. Flowchart for EEMD-FANN

For further details on the FANN, readers can refer to Günther and Fritsch (2010).

EEMD-FANN

EEMD-FANN is a hybrid model coupling EEMD and FANN. As depicted in Fig. 3, the EEMD-FANN approach is comprised of three key steps: Decomposition, single forecasting and ensemble forecasting. The main algorithm of EEMD-FANN approach is comprised of three key phases:

- Decomposition: Original river stage time series is broken down into a residue and n IMFs utilizing EEMD as described in previous section
- Single forecasting: The FANN is utilized as a forecasting model for the residue and IMFs. The FANN models built up for the IMFs and residue perform one-day-ahead forecasting, respectively. The selection of input variables for each FANN is based on the optimal lag time determined by partial autocorrelation and cross correlation functions

- Ensemble forecasting: The final forecasted time series is determined through the aggregation of the single forecasting for the residue and IMFs obtained from phase 2

Results and Discussion

Analysis

Original time series gathered from two hydrological observatories was broken down utilizing EEMD to build up EEMD-FANN model. In the decomposition process, the ensemble sample size and the standard deviation of AWGN are respectively set up to 100 and 0.2, based on the previous studies (Tang *et al*., 2012; Xu *et al*., 2016).

In daily river stage forecasting utilizing EEMD-FANN approach, one of the significant phases is to choose the efficient input variables. The input variables of FANN models for IMFs and a residue were selected based on the optimal lag time determined by statistical correlation analysis utilizing cross-correlation function (also known as sliding inner-product) and partial autocorrelation function according to the previous studies (Sudheer *et al*., 2002; Shabri and Samsudin, 2015). Table 1 summarizes the model configuration for the IMFs and residue.

This study employed MLP type of FANN model for IMF and residue forecasting. In the MLP modeling, the number of hidden neurons was optimized by an iterative approach which examines RMSE values depending on different number of hidden neurons. The operation in neurons was implemented utilizing the logistic sigmoid activation function and the MLP model was trained utilizing the most popular Back Propagation (BP) learning algorithm. The training and testing data were normalized to [0, 1] for enhancing the efficiency of BP algorithm (Dawson and Wilby, 2001).

The performance of EEMD-FANN approach was assessed quantitatively utilizing seven model efficiency indexes and compared with that of single ANFIS and FANN models which were investigated by Seo *et al*. (2015a). The indexes applied in this study are as follows:

- Dimensionless indexes: Coefficient of determination (r^2), index of agreement (IA) and Coefficient of Efficiency (CE)
- Residual error-based indexes: Mean Absolute Error (MAE), Mean Higher Order Error (MS4E), Root Mean Squared Error (RMSE) and Mean Squared Relative Error (MSRE)

For specific information on the indexes, readers can refer to Dawson and Wilby (2001).

Evaluation of Model Performance

Table 2 summarizes the model efficiency indexes of EEMD-FANN, FANN and ANFIS models in daily river stage forecasting. From Table 2, it was observed that the EEMD-FANN model yielded larger values of dimensionless indexes (IA, r^2 and CE) and smaller values of residual error-based indexes (MAE, RMSE, MS4E and MSRE), compared with the ANFIS and FANN models. Larger values of dimensionless indexes and smaller values of residual error-based indexes indicate better model efficiency (Seo *et al*., 2015a; Dawson and Wilby, 2001). The result demonstrated that the EEMD-FANN was superior to the ANFIS and FANN models. From this result, it was found that the EEMD-FANN model coupling EEMD and FANN produced better forecasting performance than the ANFIS and FANN, in terms of the indexes. The result also signified that the EEMD can further boost the forecasting ability of single FANN model.

Table 1. Input and output variables for FANN modeling of IMFs and a residue

Data	Inputs	Outputs
IMF1	$IMF1_{SC}(t)$	$IMF1_{DS}(t)$
IMF2	$IMF2_{DS}(t-4)$, $IMF2_{DS}(t-3)$, $IMF2_{DS}(t-2)$, $IMF2_{DS}(t-1)$, $IMF2_{SC}(t)$	$IMF2_{DS}(t)$
IMF3	$IMF3_{DS}(t-4)$, $IMF3_{DS}(t-3)$, $IMF3_{DS}(t-2)$, $IMF3_{DS}(t-1)$, $IMF3_{SC}(t)$	$IMF3_{DS}(t)$
IMF4	$IMF4_{DS}(t-3)$, $IMF4_{DS}(t-2)$, $IMF4_{DS}(t-1)$, $IMF4_{SC}(t)$	$IMF4_{DS}(t)$
IMF5	$IMF5_{DS}(t-2)$, $IMF5_{DS}(t-1)$, $IMF5_{SC}(t)$	$IMF5_{DS}(t)$
IMF6	$IMF6_{DS}(t-2)$, $IMF6_{DS}(t-1)$, $IMF6_{SC}(t)$	$IMF6_{DS}(t)$
IMF7	$IMF7_{DS}(t-2)$, $IMF7_{DS}(t-1)$, $IMF7_{SC}(t)$	$IMF7_{DS}(t)$
IMF8	$IMF8_{DS}(t-2)$, $IMF8_{DS}(t-1)$, $IMF8_{SC}(t)$	$IMF8_{DS}(t)$
IMF9	$IMF9_{DS}(t-1)$, $IMF9_{SC}(t)$	$IMF9_{DS}(t)$
IMF10	$IMF10_{DS}(t-1)$, $IMF10_{SC}(t)$	$IMF10_{DS}(t)$
IMF11	$IMF11_{DS}(t-1)$, $IMF11_{SC}(t)$	$IMF11_{DS}(t)$
Residue	$RES_{DS}(t-1)$, $RES_{SC}(t)$	$RES_{DS}(t)$

Note. DS, Dosan; SC, Socheon; RES, Residue

Table 2. Performance evaluation during testing period

Models	CE	IA	r^2	MAE (m)	RMSE (m)	MS4E ($10^{-4}m^4$)	MSRE (10^{-6})
EEMD-FANN	0.988	0.997	0.989	0.0285	0.0464	1.704	0.0747
FANN	0.965	0.991	0.967	0.0525	0.0787	11.909	0.2154
ANFIS	0.980	0.995	0.982	0.0361	0.0595	2.138	0.1233

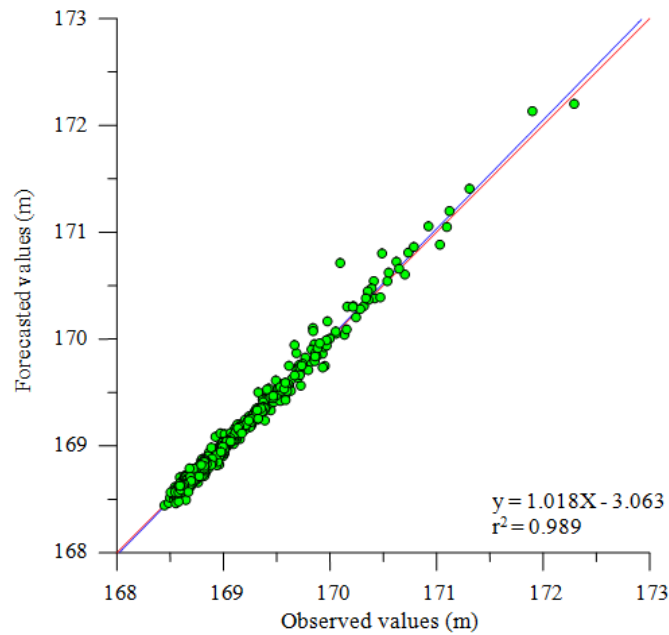Note. Values of model efficiency indexes for FANN and ANFIS from Seo *et al*. (2015a)
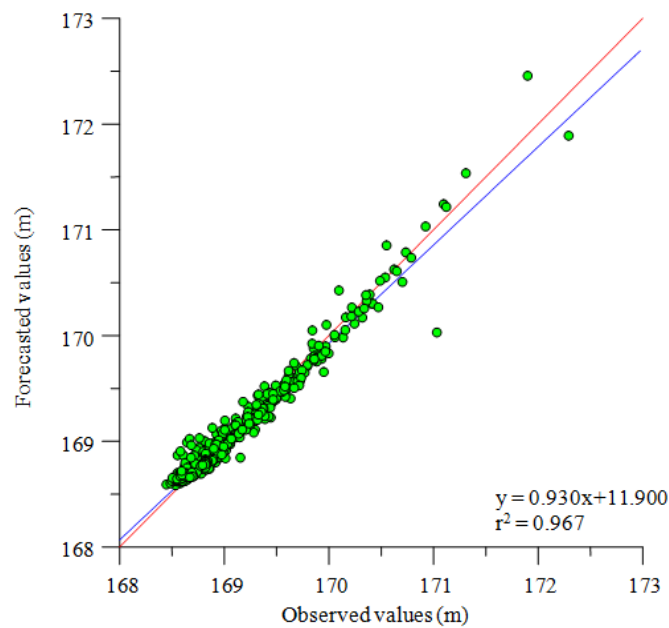
Fig. 4. Scatter chart for EEMD-FANN

$y = 1.018X - 3.063$
$r^2 = 0.989$

Fig. 5. Scatter chart for FANN

$y = 0.930x + 11.900$
$r^2 = 0.967$

Figures 4-6 show scatter charts for the three compared models: EEMD-FANN, ANFIS and FANN. It was seen from Fig. 4-6 that the degree of dispersion around the 45-degree slope line (red line) for the EEMD-FANN model is smaller than that for the ANFIS and FANN models. When straight lines (blue lines), $y = ax + b$, fitted for the scatter points were examined, it was also observed that the values of slope (*a*) and intercept (*b*) for the EEMD-FANN were closer to the values of 1 and 0, respectively. From the figures, it was apparent that the values forecasted by the EEMD-FANN get closer to the observed river stage values and the error values were smaller, in comparison with the single ANFIS and FANN. From the graphical comparison, the EEMD-FANN was found to provide more excellent forecasting ability, in comparison with the single ANFIS and FANN. The results also indicated that the EEMD can further elevate the forecasting efficiency of the single FANN model in daily river stage forecasting.
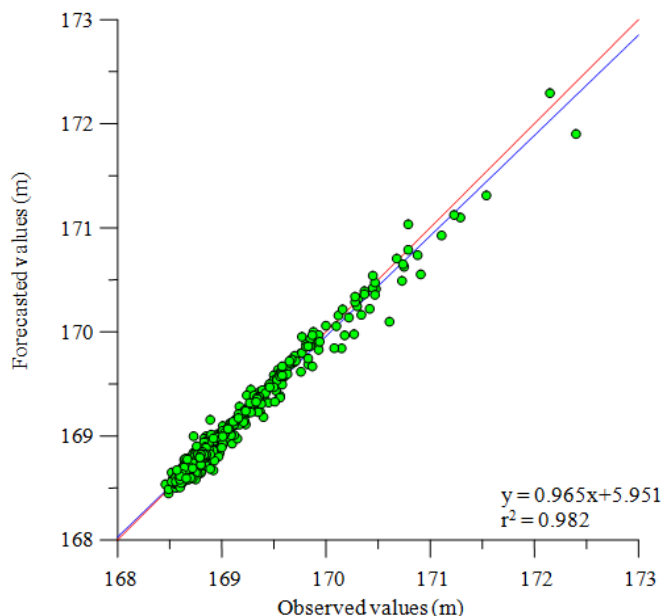
Fig. 6. Scatter chart for ANFIS

Conclusion

This research investigates the efficiency of a hybrid data-driven approach, EEMD-FANN, which integrates EEMD and FANN for forecasting daily river stage. The detailed purposes are to build up the hybrid data-driven model for improving the accuracy of daily river stage forecasting in the Andong Dam watershed located in the eastern inland region of South Korea and assess the model applicability based on comparison with the performance of single ANFIS and FANN models. The efficiency of EEMD-FANN, ANFIS and FANN models is evaluated utilizing dimensionless indexes (IA, r^2 and CE) and residual error-based indexes (MAE, RMSE, MS4E and MSRE). As a result, the EEMD-FANN model produces more excellent efficiency than the single ANFIS and FANN models, in terms of the model efficiency indexes and graphical comparison. The results indicate that the hybrid data-driven approach coupling EEMD and FANN model can further boost the forecasting ability of single FANN model and can be an effective hydrological forecasting approach.

Author's Contributions

All authors made substantial contributions to this paper.

Youngmin Seo: Contributed to research design, literature review, model development and manuscript preparation and revision.

Sungwon Kim: Contributed to conceptualization, manuscript review and revision and research supervision.

Ethics

The authors confirm that this work is original and has not been published elsewhere.

References

Adamowski, J. and K. Sun, 2010. Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds. J. Hydrol., 390: 85-91. DOI: 10.1016/j.jhydrol.2010.06.033

Amiri, G.G. and A. Asadi, 2009. Comparison of different methods of wavelet and wavelet packet transform in processing ground motion records. Int. J. Civil Eng., 7: 248-257.

Dawson, C.W. and R.L. Wilby, 2001. Hydrological modelling using artificial neural networks. Progress Phys. Geography, 25: 80-108. DOI: 10.1177/030913330102500104

Gokhale, M.Y. and D.K. Khanduja, 2010. Time domain signal analysis using wavelet packet decomposition approach. Int. J. Commun. Netw. Syst. Sci., 3: 321-329. DOI: 10.4236/ijcns.2010.33041

Günther, F. and S. Fritsch, 2010. Neuralnet: Training of neural networks. R J., 2: 30-38.

Guo, B., Z. Chen, J. Guo, F. Liu and C. Chen *et al.*, 2016. Analysis of the nonlinear trends and non-stationary oscillations of regional precipitation in Xinjiang, northwestern China, using ensemble empirical mode decomposition. Int. J. Environ. Res. Public Health, 13: 345-345. DOI: 10.3390/ijerph13030345

Huang, N.E., Z. Shen, S.R. Long, M.C. Wu and H.H. Shih *et al*., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. Royal Society London Math. Phys. Eng. Sci., 454: 903-995. DOI: 10.1098/rspa.1998.0193

Huang, N.E., M.L. Wu, S.R. Long, S.S.P. Shen and W. Qu *et al*., 2003. A confidence limit for the empirical mode decomposition and Hilbert spectral analysis. Proc. Royal Society London Math. Phys. Eng. Sci., 459: 2317-2345. DOI: 10.1098/rspa.2003.1123

Huang, S., J. Chang, Q. Huang and Y. Chen, 2014. Monthly streamflow prediction using modified EMD-based support vector machine. J. Hydrol., 511: 764-775. DOI: 10.1016/j.jhydrol.2014.01.062

Huang, Y., F.G. Schmitt, Z. Lu and Y. Liu, 2009. Analysis of daily river flow fluctuations using empirical mode decomposition and arbitrary order Hilbert spectral analysis. J. Hydrol., 373: 103-111. DOI: 10.1016/j.jhydrol.2009.04.015

Kaltech, A.M., 2015. Wavelet genetic algorithm-support vector regression (Wavelet GA-SVR) for monthly flow forecasting. Water Resources Manage., 29: 1283-1293. DOI: 10.1007/s11269-014-0873-y

Karthikeyan, L. and D.N. Kumar, 2013. Predictability of nonstationary time series using wavelet and EMD based ARMA models. J. Hydrol., 502: 103-119. DOI: 10.1016/j.jhydrol.2013.08.030

Kim, S. and H.S. Kim, 2008. Neural networks and genetic algorithm approach for nonlinear evaporation and evapotranspiration modeling. J. Hydrol., 351: 299-317. DOI: 10.1016/j.jhydrol.2007.12.014

Kim, S., J. Shiri and O. Kisi, 2012. Pan evaporation modeling using neural computing approach for different climatic zones. Water Resources Manage., 26: 3231-3249. DOI: 10.1007/s11269-012-0069-2

Kim, S., J. Shiri, O. Kisi and V.P. Singh, 2013. Estimating daily pan evaporation using different data-driven methods and lag-time patterns. Water Resources Manage., 27: 2267-2286. DOI: 10.1007/s11269-013-0287-2

Kisi, O., 2007. Streamflow forecasting using different artificial neural network algorithms. J. Hydrol. Eng., 12: 532-539. DOI: 10.1061/(ASCE)1084-0699(2007)12:5(532)

Kisi, O., L. Latifoğlu and F. Latifoğlu, 2014. Investigation of empirical mode decomposition in forecasting of hydrological time series. Water Resources Manage., 28: 4045-4057. DOI: 10.1007/s11269-014-0726-8

Kisi, O., J. Shiri and A.H. Nazemi, 2011. A wavelet-genetic programming model for predicting short-term and long-term air temperatures. J. Civil Eng. Urbanism, 1: 25-37.

Nourani, V., M. Parhizkar, T.R. Khanghah, A.H. Baghanam and E. Sharghi, 2012. Wavelet-based feature extraction of rainfall-runoff process via self-organizing map. Proceedings of the 12th WSEAS International Conference on Applied Computer Science, (ACS' 12), Greece, pp: 101-106.

Othman, F. and M. Naseri, 2011. Reservoir inflow forecasting using artificial neural network. Int. J. Phys. Sci., 6: 434-440. DOI: 10.5897/IJPS10.649

Ouyani, Q., W. Lu, X. Xin, Y. Zhang and W. Cheng *et al*., 2016. Monthly rainfall forecasting using EEMD-SVR based on phase-space reconstruction. Water Resources Manage., 30: 2311-2325. DOI: 10.1007/s11269-016-1288-8

Ravikumar, K. and S. Tamilselvan, 2014. On the use of wavelets packet decomposition for time series prediction. Applied Math. Sci., 8: 2847-2858.

Ren, Y., P.N. Suganthan and N. Srikanth, 2015. A comparative study of empirical mode decomposition-based short-term wind speed forecasting methods. IEEE Trans. Sustainable Energy, 6: 236-244. DOI: 10.1109/TSTE.2014.2365580

Seo, Y., 2015. River stage forecasting model combining wavelet packet transform and artificial neural network. J. Environ. Sci. Int., 24: 1023-1036. DOI: 10.5322/JESI.2015.24.8.1023

Seo, Y., S. Kim, O. Kisi and V.P. Singh, 2015a. Daily water level forecasting using wavelet decomposition and artificial intelligence techniques. J. Hydrol., 520: 224-243. DOI: 10.1016/j.jhydrol.2014.11.050

Seo, Y., S. Kim and V.P. Singh, 2015b. Physical Interpretation of River Stage Forecasting using Soft Computing and Optimization Algorithms. In: Harmony Search Algorithm, J.H. Kim and Z.W. Geem (Eds.), Springer, Heidelberg, ISBN-10: 3662479265, pp: 259-266.

Shabri, A. and R. Samsudin, 2015. Empirical mode decomposition-least squares support vector machine based for water demand forecasting. Int. J. Adv. Soft Comput. Applic., 7: 38-53.

Sudheer, K.P., A.K. Gosain and K.S. Ramasastri, 2002. A data-driven algorithm for constructing artificial neural network rainfall-runoff models. Hydrol. Proc., 16: 1325-1330. DOI: 10.1002/hyp.554

Tang, L., L. Yu, S. Wang, J. Li and S. Wang, 2012. A novel hybrid ensemble learning paradigm for nuclear energy consumption forecasting. Applied Energy, 93: 432-443. DOI: 10.1016/j.apenergy.2011.12.030

Wang, W., K. Chau, D. Xu and X.Y. Chen, 2015. Improving forecasting accuracy of annual runoff time series using ARIMA based on EEMD decomposition. Water Resources Manage., 29: 2655-2675. DOI: 10.1007/s11269-015-0962-6

Wang, W., D. Xu, K. Chau and S. Chen, 2013. Improved annual rainfall-runoff forecasting using PSO–SVM model based on EEMD. J. Hydroinformat., 15: 1377-1390. DOI: 10.2166/hydro.2013.134

Wu, C.L., K.W. Chau and Y.S. Li, 2009. Predicting monthly streamflow using data-driven models coupled with data-preprocessing techniques. Water Resources Res. DOI: 10.1029/2007WR006737

Wu, Z.H. and N.E. Huang, 2009. Ensemble empirical mode decomposition: A noise-assisted data analysis method. Adv. Adaptive Data Analy., 1: 1-41. DOI: 10.1142/S1793536909000047

Xu, J., Y. Chen, L. Bai and Y. Xu, 2016. A hybrid model to simulate the annual runoff of the Kaidu River in northwest China. Hydrol. Earth Syst. Sci., 20: 1447-1457. DOI: 10.5194/hess-20-1447-2016

Yu, L., S. Wang and K.K. Lai, 2008. Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. Energy Econom., 30: 2623-2635. DOI: 10.1016/j.eneco.2008.05.003

Zhao, X. and X. Chen, 2015. Auto regressive and ensemble empirical mode decomposition hybrid model for annual runoff forecasting. Water Resources Manage., 29: 2913-2936. DOI: 10.1007/s11269-015-0977-z