Amirkabir Journal of Civil Engineering

Amirkabir J. Civil Eng., 52(10) (2021) 649-652
DOI: 10.22060/ceej.2019.16043.6109

Monthly precipitation prediction improving using the integrated model based on kernel-wavelet and complementary ensemble empirical mode decomposition

Kiyoumars Roushangar[1,*] , Roghayeh Ghasempour[2]

[1] Professor, Faculty of Civil Engineering, University of Tabriz, Tabriz, Iran
[2] PhD Candidate, Faculty of Civil Engineering, University of Tabriz, Tabriz, Iran

ABSTRACT: Estimates of monthly rainfall are important for various purposes such as flood estimation, drought, irrigation planning, and river basin management. In the present study, the monthly rainfall of Tabriz station was investigated using the intelligent Gaussian Process Regression (GPR) method based on Complementary Ensemble Empirical Mode Decomposition (CEEMD) and Wavelet Transform (WT). Different models were defined based on teleconnection patterns and climatic elements, and the impact of different input parameters was assessed. The obtained results proved high capability and efficiency of the applied method in predicting the monthly precipitation. The results showed that time series decomposition based on wavelet transformation led to more accurate outcomes compared to the complementary ensemble empirical mode decomposition. The best evaluation of test series using wavelet transform decomposition was obtained for the state of modeling based on teleconnection patterns and climatic elements with the values of DC=0.889, R=0.961 and RMSE=0.036. Also, based on the sensitivity analysis, Pt-3 was found to be the most effective parameter in modeling.

Review History:

Received: 2019-03-23
Revised: 2019-04-29
Accepted: 2019-09-17
Available Online: 2019-09-29

Keywords:

Climatic elements
Empirical mode
GPR, Precipitation
Wavelet Transform

1. INTRODUCTİON

Precipitation plays an important role in determining the climate of a region. Precise estimation of precipitation is required to manage and plan water resources, as well as other related applications such as hydrology, climatology, meteorology and agriculture. So far, numerous precipitation prediction methods have been proposed in literature, including time series models, regression models, probabilistic models, machine learning models, physical models, and a host of hybrid models such as Tofani et al [1] and Soltani et al., [2]. In recent years, the Meta model approaches such as Artificial Neural Networks (ANNs), Neuro-Fuzzy models (NF), Genetic Programming (GP), Support Vector Machine (SVM), and Gaussian Process Regression (GPR) have been applied in investigating the hydraulic and hydrologic complex phenomena [3]. Hybrid models involving signal decomposition have also been shown to be effective in improving prediction accuracy of time series prediction methods, as indicated in [4]. Complementary Ensemble Empirical Mode Decomposition analysis is one of the widely used signal decomposition methods for hydrological time series prediction. Decomposition of time series reduces the difficulty of forecasting, thereby improving forecasting accuracy. Due to the complexity of the precipitation phenomenon and the effect of various parameters on its prediction, in this study, the capability of GPR as a kernel based approach and also integrated CEEMD-GPR and DWT-GPR models were assessed for precipitation modeling in Tabriz station during the period time of 1978-2017. Different input combinations were considered using climatic data such as precipitation, monthly temperature and relative humidity with one to four time lag and telecommunication patterns. Considering two different types of modeling based on original time series data (without data decomposition) and time series decomposition data, the capability of used methods was investigated. Also, sensitivity analysis was performed to determine the effective parameters in monthly precipitation prediction process.

2. METHODOLOGY

In the current study, monthly rainfall data of the Tabriz station was used during the period of 1978-2017. GPR model are based on the assumption that adjacent observations should convey information about each other. Gaussian processes are a way of specifying a prior directly over function space. This is a natural generalization of the Gaussian distribution, whose mean and covariance are a vector and matrix, respectively. The Gaussian distribution is over vectors, whereas the Gaussian process is over functions. Thus, due to prior knowledge about the data and functional dependencies, no validation process is required for generalization, and GP regression models are able to understand the predictive distribution corresponding to the test input. A GP is defined as a collection of random variables, any finite number of which has a joint multivariate Gaussian distribution. The WT is a popular method and a very

*Corresponding author's email: kroshangar@yahoo.com

Copyrights for this article are retained by the author(s) with publishing rights granted to Amirkabir University Press. The content of this article is subject to the terms and conditions of the Creative Commons Attribution 4.0 International (CC-BY-NC 4.0) License. For more information, please visit https://www.creativecommons.org/licenses/by-nc/4.0/legalcode.

Table 1. Statistical parameters results for Test series

Model	Performance criteria										
	R	DC	RMSE		R	DC	RMSE		R	DC	RMSE
Based on precipitation data											
GPR					CEEMD- GPR				DWT- GPR		
P(I)	0.378	0.219	0.236		0.623	0.484	0.119		0.725	0.525	0.099
P(II)	0.536	0.333	0.222		0.766	0.548	0.090		0.906	0.816	0.057
P(III)	0.605	0.508	0.193		0.822	0.721	0.081		0.914	0.835	0.054
P(IV)	0.587	0.420	0.195		0.805	0.699	0.086		0.913	0.828	0.062
Based on precipitation and teleconnection patterns data											
CEEMD- GPR					DWT- GPR						
I(I)	0.638	0.508	0.106		0.708	0.588	0.091				
I(II)	0.834	0.735	0.069		0.926	0.851	0.046				
I((III)	0.845	0.742	0.064		0.938	0.859	0.042				
I(IV)	0.822	0.726	0.075		0.912	0.841	0.050				
I(V)	0.818	0.724	0.080		0.908	0.838	0.053				
I(VI)	0.827	0.732	0.072		0.918	0.848	0.048				
I(VII)	0.866	0.768	0.054		0.961	0.889	0.036				
I(VIII)	0.471	0.351	0.250		0.523	0.407	0.166				

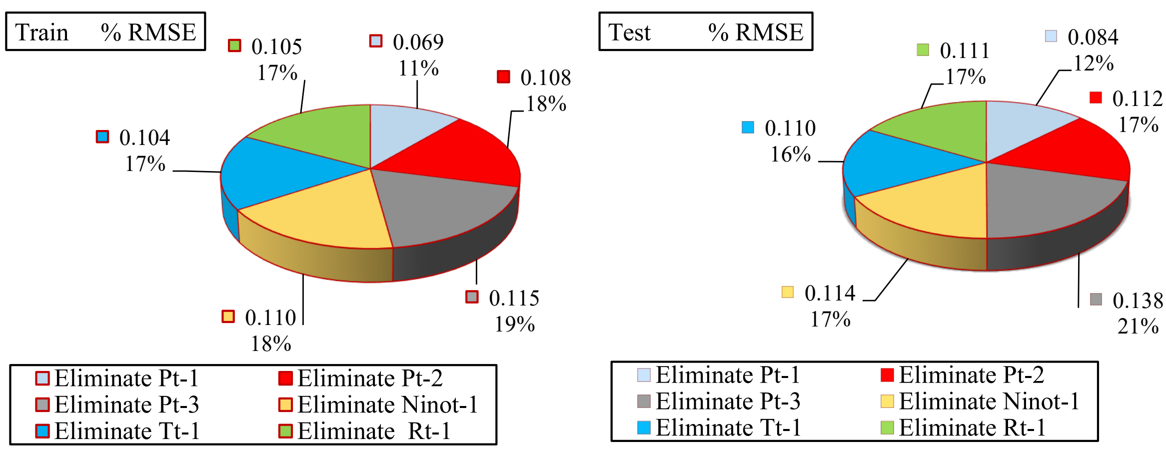

Fig. 1. Relative significance of each of input parameters of the best model

precise method for time series processing. While the general theory behind WT is quite analogous to that of the short time Fourier transform, WT allows for a completely flexible window function (called the mother wavelet), which can be changed over time based on the shape and compactness of the signal. Given this property, WT can be used to analyze the time–frequency characteristics of any kind of time series. CEEMD was proposed to solve the mode mixing issue of empirical mode decomposition (EMD) which specifies the true IMF as the mean of an ensemble of trials [5]. Each trial consists of the decomposition results of the signal plus a white noise of finite amplitude. EMD can be used to decompose any complex signal into finite intrinsic mode functions and a residue, resulting in subtasks with simpler frequency components and stronger correlations that are easier to analyze and forecast. Another important feature of empirical mode decomposition is that it can be used for noise reduction of noisy time series, which can be effective in improving the accuracy of model predictions.

3. RESULTS AND DİSCUSSİON

In order to evaluate and review the performance of the developed models and determine the accuracy of the selected models, three performance criteria named Correlation Coefficient (R), Determination Coefficient (DC), and Root Mean Square Errors (RSME) were used according to Table 1. According to the results, it could be seen that the accuracy of single GPR model did not yield to accurate prediction. However, both integrated WT-GPR and EEMD-GPR models led to higher accuracy than the GPR model. The use of these two methods deacresed the error criteria approximatly between 25 to 35%. It was observed that in prediction of precipitation, climatic elements including monthly average temperature, relative humidity, and previous months precipitation as well as teleconnection patterns were effective in prediction process.

According to the Fig. 1, sensitivity analysis was performed to determine the most significant parameters in modeling process. It was observed that P_{t-3} is the most effective parameter in precipitation modeling.

4. CONCLUSİONS

The comparison of the developed models' accuracy revealed that both integrated CEEMD-GPR and DWT-GPR models had better performance compared with the GPR model in predicting the monthly precipitation. The use of these two methods deacresed the error criteria 25 to 35%. Also, based on the sensitivity analysis, P_{t-3} was found to be the most effective parameter in modeling process.

REFERENCES

[1] P. Tofani, E. Mosaedi, A. Fakheri Fard, Precipitation forecast using wavelet theory, Water and Soil Journal (Agriculture Sciences and Technology), 25(5) (2011) 1217-1226. [in Persian]

[2] A. S. Soltani, A. Saberi, M. Gheisouri, Determination of the best time series model for forecasting annual rainfall of selected stations of Western Azerbaijan province, Researches in Geographical Sciences, 17(44) (2017) 87-105.

[3] K. Roushangar, R. Ghasempour, The study of the performance of classical and artificial intelligence methods in the estimation of roughness coefficients in pontoons, Irrigation and Drainage Journal of Iran, 12(4) (2019) 811-822. [in Persian]

[4] K.M. Lau, H.Y. Weng, Climate signal detection using wavelet transform, How to make time series sing, Bull Am Meteorol Soc, 76 (1995) 2391-2402.

[5] Y. Amirat, M. Benbouzidb, T. Wang, K. Bacha, G. Feld, EEMD-based notch filter for induction machine bearing faults detection, Applied Acoustics, 133 (2018) 202–209.

HOW TO CITE THIS ARTICLE

K. Roushangar, R. Ghasempour, Monthly precipitation prediction improving using the integrated model based on kernel-wavelet and complementary ensemble empirical mode decomposition, Amirkabir J. Civil Eng., 52(10) (2021) 649-652.

DOI: 10.22060/ceej.2019.16043.6109