Monthly Runoff Prediction Using Wavelet Transform and Fast Resource Optimization Network (Fron) Algorithm

Jiahai Lu and Aiyun Li

College of Water Resources Science and Engineering, Taiyuan University of Technology, No.79, West Yingze Avenue, Wanbailin District, Taiyuan City, Shanxi Province, People’s Republic of China.
Email: liaiyun@tyut.edu.cn

Abstract. In this study, Discrete Wavelet Transform (DWT) and a Fast Resource Optimization Network (FRON) Algorithm were integrated to forecast monthly runoff and improve the accuracy of forecasting models. The original monthly runoff sequence is reconstructed by DWT processing into two components of deterministic and stochastic components, then these two components were inputted into two different FRON model respectively, and finally the prediction results of the two models were summed up as the final forecasts of monthly runoff. The WFRON1 model is compared with the WFRON2 model and the FRON model without discrete wavelet transform. The mean absolute error (MAE), the deterministic coefficient (DC) and the correlation coefficient (R) are used as the model evaluation index. The model is applied to the prediction of monthly runoff at the Pantuo hydrological station in Shangyuan River. The results show that: The evaluation index of WFRON2 model is better than WFRON1, and both models predict the effect better than FRON model.

1. Introduction
The accurate of medium- and long-term hydrological forecasting plays an important role in the optimal management watershed water resources as well as water-related operation management [1-2]. Accurate simulation and prediction of the runoff has always been one of the most important issues for domestic hydrologists [3]. The process of runoff forecasting is influenced by many factors such as rainfall, underlying surface and human activities. The variation of runoff is complicated and highly nonlinear [4]. Therefore, the runoff forecasting can be a difficult task because of its high non-linearity and extreme temporal variability.

In hydrology and water resources systems, the annual runoff variation can be approximated as stationary series and the monthly runoff variation are usually non-stationary series. There has been an increased interest in wavelet-based analysis in hydrology and water resources topics over the past decade. Wavelet transform method can solve the limited accuracy of non-stationary series [5-8]. Since most of the runoff variation are nonlinear and non-stationary, many time-frequency analysis methods have been applied to hydrology sequence for their ability to simultaneously provide local information both in time and frequency domains. The internal structure and variation characteristics of the hydrology sequence can be obtained by analyzing the wavelet coefficients. The WGRNN model is obtained by combining two methods, Discrete Wavelet Transform (DWT) and Generalized Regression Neural Network (GRNN), for one-month-ahead streamflow forecasting, the comparison results revealed that the WGRNN performs better than the GRNN and FFNN models in monthly streamflow prediction. The coupling DWT and ANN model had been used for seasonal river runoff forecasting. The WGEP model, is developed by integrating the DWT with the Gene Expression Programming...
(GEP) models, mainly forecast the runoff using rainfall data. The study showed that the GEP based rainfall-runoff models can be considered as an alternative to the ANN based models. The SVM model based on wavelet transform is used to simulate and forecast the runoff [9-12]. The results show that the model will use the time series of wavelet transform as input, and give better and more robust prediction accuracy. The radial basis function neural network (RBFNN) has the best approximation characteristic and strong non-linear mapping ability, which is suitable for solving various nonlinear problems. The coupling model of wavelet transform and RBFNN has few applications in hydrology sequence prediction. This paper attempts to couple the RBFNN with DWT. Based on the analysis of the internal structure and the variation characteristics of the hydrology sequence. This paper proposes a fast resource optimization network (FRON) for hidden layer nodes of RBF network.

2. Model and Numerical Results

Wavelet transform can express the time-frequency information of time series. Wavelet function is a kind of function with oscillatory characteristics. It can calculate the individual spectral components by stretching transformation, so it can analyze non-stationary data. Using the multi-resolution analysis of wavelet, the high frequency detail of the original data is removed. The singular point of white noise and signal with completely different properties under multi-scale wavelet transform, the white noise can be removed.

The DWT-FRON model is obtained by combining two methods, DWT and FRON. In this paper, the original time series are decomposed into a certain number of sub-time series components (Ds) and approximate sequences (A) by using the Mallat DWT algorithm with orthogonal asymmetric Daubechies, the deterministic and stochastic components of the original series were analyzed. The DWT-FRON model is used to predict the deterministic and stochastic components separately. The predictive values of the known runoff series are obtained by linearly superposing the predictive values of deterministic and stochastic components. The method is defined as WFRON1 model. However, random components are highly nonlinear. When using FRON model to predict, although the effect is better than the traditional prediction method, but in general, there are still low accuracy problems. In this paper, we attempt to decompose and reconstruct known runoff sequences, eliminate invalid sub-sequences that have bad correlation with known runoff sequences, select effective sequences and add them to the approximate sequences to obtain new sequences. The new sequence data is taken as the FRON model input and measured monthly runoff data as the output of the FRON model which constitutes WFRON2 model. The mean absolute error (MAE), the deterministic coefficient (DC) and the correlation coefficient (R) are used as the model evaluation parameters. The runoff forecasting model with the highest prediction accuracy is selected. WFRON1 model and WFRON2 model structure shown in Figure 1.

![Figure 1. Structure of WFRON1 and WFRON2](image-url)
The ChangYuan River is a first order tributary of the Fen River and has the river length of 87 km with the drainage area of about 1029.7 km². The basin climate is a temperate and semi-arid monsoon climate with an average annual precipitation of 442mm, more than 60% of the precipitation falls between June and September with occasional heavy storms. The hydrological station at Pantuo was selected as the target gauge in this study. These runoff data covering 1935-2004 were provided to establish the runoff simulation model. Previous 60 years of data were used to determine the model, After 10 years of data were used to verify the model.

The original sequence is decomposed by Mallat discrete wavelet transform algorithm. According to the trend change, the original sequence is decomposed into four layers, and five different sub-sequence (see Fig. 2) are obtained, that is, the low frequency sequence (approximate part cA4) with time scale 4 and the high frequency sequence of each time scale detail section (cD1, cD2, cD3, cD4), the approximate part shows the trend of the original sequence. The Pearson correlation coefficients (see Table 1) and the significance level of the original sequence Qt were calculated for each sub-sequence (the first month, the first 2 months, the first 3 months, the first 4 months). It can be seen that the correlation between the cD1 sequence and the original sequence is low, and it can be considered that the independent random component st, cD2, cD3 and cD4 are highly correlated with the original sequence, and cA4 are considered to be the definite component qt. The prediction of the original sequence can be transformed into a predictive of the deterministic and stochastic components. The monthly runoff deterministic and stochastic component changes with time as shown in Fig.3.

![Figure 2. Wavelet-decomposed sub-series components (Qi) of monthly runoff data at Pantuo station](image-url)
Figure 3. Variation in stochastic component and Variation in deterministic component of the monthly runoff at Pantuo station

Table 1. Comparison of deterministic and stochastic components with monthly runoff predicted by WFRON1 and WFRON2 models using different inputs

| Model inputs | Deterministic ingredients | Stochastic ingredients | WFRON1 | WFRON2 |
|--------------|--------------------------|------------------------|--------|--------|
|              | MAE          | DC           | R   | MAE          | DC           | R   | MAE          | DC           | R   |
| qt-1         | 19.25        | 0.678        | 0.826 | 21.75        | 0.689        | 0.823 | 22.23        | 0.682        | 0.823 |
| qt-1, qt-2   | 8.56         | 0.945        | 0.965 | 10.21        | 0.923        | 0.965 | 9.54         | 0.936        | 0.963 |
| qt-1, qt-2, qt-3 | 8.67     | 0.935        | 0.959 | 10.43        | 0.915        | 0.968 | 9.25         | 0.926        | 0.962 |
| qt-1, qt-2, qt-3, qt-4 | 8.62  | 0.942        | 0.961 | 10.98        | 0.915        | 0.956 | 9.56         | 0.926        | 0.952 |

The FRON model is used to predict the deterministic composition. The deterministic components of the current month's flow are predicted by the deterministic components of the runoff of the previous months. The inputs are: (a) the first month (qt-1), (b) the first 2 months (qt-1, qt-2), (c) the first 3 months (qt-1, qt-2, qt-3), (d) the first 4 months (qt-1, qt-2, qt-3, qt-4). The simulation results of stochastic components are the same as above, and the prediction results are superimposed to obtain the prediction results of WFRON1 model. Table 2 gives the deterministic composition qt, randomness components st and the evaluation index of WFRON1 monthly runoff simulation forecasting model validation period. From the table, we can see that the prediction accuracy of the input b combination model is the highest when the deterministic component is predicted. When the stochastic component is predicted, the prediction accuracy of the input d model is the highest. The deterministic coefficient DC and the correlation coefficient R of the random component are very different from the predicted results of the deterministic component. The results show that the GRNN1 monthly runoff prediction results are ideal. The accuracy of the model is the highest when the input b is combined, the average absolute error MAE, the deterministic coefficient DC and the correlation coefficient R are 10.21 m$^3$/s, 0.923, 0.965.

However, it can be clearly seen from Table 1 that the simulated effect of FRON on the random component is much lower than that of the deterministic component, which is due to the poor self-dependence of the random component, which affects the improvement of the simulation effect. In this paper, the WFRON2 model is established to eliminate the invalid sequence cD1 which is not related to the original sequence. The new sequence is obtained by superimposing the effective sequence cD2, cD3, cD4 and the approximate part cA4 as the input of FRON.

Comparing the evaluation indexes of FRON1 model and WFRON2 model in Table 1, we can see that the WFRON2 model has the best prediction accuracy when the input b is the same. And the accuracy of the model is higher than that of the WFRON1 model after removing the invalid sequence. The MAE, DC and R of the WFRON2 model are 9.54 m$^3$/s, 0.936 and 0.963 respectively.
Figure 4 show the comparison between the measured values of monthly runoff and the predicted values of WFRON1 and WFRON2. It can be seen that the predicted values of the two models are in good agreement with the measured values, and the model has higher prediction accuracy, the overall trend of the simulation results are better. The effect of the WFRON2 model on the peak is better than that of the WFRON1 model, and the experimental results and the correlation coefficients are better than those of the WFRON1 model, but the WFRON1 and WFRON2 models have large errors in the simulation of the minimum.

The original sequence is simulated by using the FRON model without discrete wavelet transform, and the input group with the highest precision is predicted. That is compared with the WFRON1 model and WFRON2 model with the highest accuracy of the input combination forecast (see Table 2). As can be seen from Table 2, the evaluation index of WFRON2 model is better than that of WFRON1 model and GRNN model, the deterministic coefficient is 0.926, WFRON2 model is the best for predicting the runoff of the Pantuo Gorge in three models. The prediction accuracy of the FRON model is the lowest, indicating that the coupling model of the discrete wavelet transform and the generalized regression neural network can better predict the monthly runoff than the single generalized regression neural network model because the characteristics of the sub-sequence of the wavelet transform are more obvious than the original sequence, Wavelet transform has improved effect on runoff simulation accuracy.

| Model  | MAE/(m³/s) | DC   | R    |
|--------|------------|------|------|
| WFRON1 | 9.75       | 0.912| 0.953|
| WFRON2 | 9.54       | 0.926| 0.963|
| FRON   | 10.236     | 0.856| 0.924|

3. Conclusion
The discrete wavelet transform can effectively reveal the variation characteristics of the runoff series at different resolution levels, and can effectively improve the simulation accuracy of runoff prediction. By predicting the deterministic and stochastic components separately, which can further reflect the physical mechanism of the runoff sequence that is affected by the deterministic and uncertain factors, and the prediction accuracy of the model is further improved. However, the poor self-dependence of the random sequence causes the low prediction accuracy of the model, which affects the effect of the monthly runoff simulation.
Coupled discrete wavelet transform and fast resource optimization network Algorithm, the monthly flow forecasting model has a good effect on the overall trend of runoff, but the effect of peak and minimum simulation is not ideal.

After the discrete wavelet transform, eliminating the ineffective sequences with poor correlation with the original sequence as the input of the generalized regression neural network to predict the runoff prediction model is more accurate, which is feasible in the actual production application.

4. Acknowledgments

Science and Technology Key Project on Social Development of Shanxi Province (20140313023-4), International Science and Technology Cooperation Program of Shanxi Province (201603D421011)

5. References

[1] Pagano, T.C., Wood, A.W., Ramos, M.H., Cloke, H.L., Pappenberger, F., Clark, M.P., Cranston, M., Kavetski, D., Mathevet, T., Sorooshian, S., Verkade, J.S., 2014. Challenges of operational river forecasting. J. Hydrometeorol. 15(4), 1692-1707.

[2] Zhao, X., Chen, X., 2015. Auto regressive and ensemble empirical mode decomposition hybrid model for annual runoff forecasting. Water Resour Manage. 29(8),2913-2926.

[3] Sanikhani, H., Kisi, O,2012. River flow estimation and forecasting by using two different adaptive Neuro-fuzzy approaches. Water Resour Manage 26(6), 1715-1729

[4] Zhou Xiuping, Li Tianxiang, Wang Wensheng. 2013. Annual runoff prediction based on least square support vector machines-Markov chain combined model. Journal of Hydroelectric Engineering. 32(4),16-19

[5] Kisi, O.,2011. A combined generalized regression neural network wavelet model for monthly streamflow prediction. Journal of Civil Engineering. 15(8), 1469-1479

[6] Valipour M, Banihabib M E, Behbahani S M R. 2013. Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. Journal of Hydrology., 476(2),433-441.

[7] Kisi O, Shiri J, Tombul M. 2013. Modeling rainfall-runoff process using soft computing techniques. Computers & Geosciences, 51(2),108-117.

[8] Abrahart R J, See L. 2000. Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments. Hydrological Processes, 14(11/12), 2157-2172.

[9] Kisi, O., Cimen, M., 2011. A wavelet-support vector machine conjunction model for monthly streamflow forecasting. 399(1/2),132-140

[10] ZHANG L B, JIN JL, CHENG JL. 2007. The forecasting method of selecting-best model based on artificial neural networks and its application. Journal of Hydroelectric Engineering, 26(6), 12-16

[11] Wang W C, Chau K W, Qiu L, Chen Y B. 2015. Improving forecasting accuracy of medium and long-term runoff using artificial neural network based on EEMD decomposition. Environmental Research, 139, 46-54

[12] TANG Q, WANG HR, XU XY, Wang C. 2014. Hydrological time series model based on SVM with mixed kernel function and its application. Systems Engineering—Theory & Practice, 34(2), 521-529