



Multi-step-ahead predictor design for effective long-term forecast of hydrological signals using a novel wavelet neural network hybrid model

J.-S. Yang^{1,2}, S.-P. Yu^{1,2}, and G.-M. Liu¹

¹State Key Laboratory of Soil and Sustainable Agriculture, Nanjing Institute of Soil Science, Chinese Academy of Sciences, 71 East Beijing Road, Nanjing, 210008, China

²Dongtai Institute of Tidal Flat, Nanjing Branch of the Chinese Academy of Sciences, 8 Beihai Road, Dongtai, 224200, China

Correspondence to: S.-P. Yu (spyu@issas.ac.cn)

Received: 23 June 2013 – Published in Hydrol. Earth Syst. Sci. Discuss.: 15 July 2013

Revised: 13 October 2013 – Accepted: 12 November 2013 – Published: 10 December 2013

Abstract. In order to increase the accuracy of serial-propagated long-range multi-step-ahead (MSA) prediction, which has high practical value but also great implementary difficulty because of huge error accumulation, a novel wavelet neural network hybrid model – CDW-NN – combining continuous and discrete wavelet transforms (CWT and DWT) and neural networks (NNs), is designed as the MSA predictor for the effective long-term forecast of hydrological signals. By the application of 12 types of hybrid and pure models in estuarine 1096-day river stages forecasting, the different forecast performances and the superiorities of CDW-NN model with corresponding driving mechanisms are discussed. One type of CDW-NN model, CDW-NF, which uses neuro-fuzzy as the forecast submodel, has been proven to be the most effective MSA predictor for the prominent accuracy enhancement during the overall 1096-day long-term forecasts. The special superiority of CDW-NF model lies in the CWT-based methodology, which determines the 15-day and 28-day prior data series as model inputs by revealing the significant short-time periodicities involved in estuarine river stage signals. Comparing the conventional single-step-ahead-based long-term forecast models, the CWT-based hybrid models broaden the prediction range in each forecast step from 1 day to 15 days, and thus reduce the overall forecasting iteration steps from 1096 steps to 74 steps and finally create significant decrease of error accumulations. In addition, combination of the advantages of DWT method and

neuro-fuzzy system also benefits filtering the noisy dynamics in model inputs and enhancing the simulation and forecast ability for the complex hydro-system.

1 Introduction

The hydrological signal forecast, especially a long-term forecast, is important for the study and guidance of water resource management. Nevertheless, the hydrological signals are highly complex nonlinear systems and have severe variations in time and space, which make accurate forecasts difficult. Generally, the hydrological time series were predicted with models based on physical considerations or other numerical theories, such as the LR (linear regressive) analysis methods (Salas et al., 1980) based on the stochastic theory, the grey models (Deng, 1992) based on the grey information theory, the chaos models (Jayawardena and Lai, 1994; Islam and Sivakumar, 2002) based on the local similarity of signals, the fuzzy prediction models (Jang, 1993; Jang et al., 1997; Chen, 2005) based on the fuzzy theory, the TAR (threshold auto-regression), BL (bilinear time series), and SVM (support vector machine) models (Tong, 1990; Liang and Sivapragasm, 2002; Zou et al., 2010) based on the nonlinear time-series analysis, the ANN (artificial neural network) models (Raman and Sunlikumar, 1995; Yu et al., 2008; Yang et al., 2009) based on the black-box theory, and the NNB

(nearest neighbour bootstrapping) regressive models (Wang et al., 2001) based on the nonparametric prediction theory. However, these models were generally not successful enough in producing accurate predictions due to some inaccurate initial conditions, parameterisation schemes of sub-scale phenomena, and limited spatial resolution (Olson et al., 1995).

Many hybrid models have been proposed as predictors to improve the accuracy of hydrological time-series forecasts, such as the wavelet artificial neural network (ANN) model (Ancil and Tape, 2004; Krishna et al., 2011; Nayak et al., 2013), the periodic ANN (PANN) model (Wang et al., 2006), the chaotic ANN model (Karunasinghe and Liong, 2006), the hybrid fuzzy-ANN model (Nayak et al., 2007), the wavelet-based grey model (Chou, 2007), the wavelet-based NF (neuro-fuzzy) model (Partal and Kisi, 2007; Engin et al., 2007; El-Shafie et al., 2007), the non-supervised ANN-EA (evolutionary algorithms) model (Cao and Park, 2007; Chang et al., 2007), the fuzzy-SVM model (Hua et al., 2008), the wavelet-based multi-layer perceptron model (Kisi, 2008), the wavelet-regression (WR) model (Kisi, 2011), and the wavelet-based fuzzy logic model (Ozger et al., 2012). These hybrid models have shown different advantages for accurate predictions due to their capabilities of utilising present information effectively. Among these hybrid models, the neural network (NN) models, such as the NF (neuro-fuzzy) and ANN, are the most popularly utilised sub-models for signal forecast due to their capabilities of effectively learning complex and nonlinear relationships (Maier et al., 2010). The ANN model has been commonly used in hydrological signal forecasts by a number of researchers (French et al., 1992; Jain et al., 1999; ASCE, 2000; Cigizoglu, 2005; Marzano et al., 2006; Zou et al., 2010). The NF model has been introduced and successfully used in the hydrological sciences in recent years (Nayak et al., 2004, 2005; Kisi, 2005; Chang and Chang, 2006). In addition, the wavelet transform, as a strong mathematical tool in providing the good local representation of a signal in both the time and frequency domains, has become a useful method for analysing variations, periodicities, and trends in time series (Daubechies, 1994; Torrence and Compo, 1998; Coulibaly and Burn, 2004; Partal and Kucuk, 2006). Among the various types of wavelet transforms, the discrete wavelet transform (DWT) is popularly used as the data preprocessing method in a hybrid model to decompose the original signal inputs due to its capabilities of effectively classifying a hydro-meteorological time series into distinct time and frequency domains (Smith et al., 1998; Kim and Valdes, 2003; Labat, 2005).

Because of the common Markovian property (Bolch et al., 2006) embedded in the hydro-meteorological time series, most recent pure and hybrid models use data series at different previous time points as model inputs to forecast the original data series at the current time point. For daily time series, the data series from 1 day prior to a few days prior are usually used as model inputs, namely using data series S_{t-1} , S_{t-2} , ... as inputs, to forecast S_t . The data series at 1 day

prior is always selected as one of the inputs because of the usually high lag-1 autocorrelation (Kisi, 2008, 2011; Zhou et al., 2008). This selection principle denotes a type of commonly used single-step-ahead (SSA) prediction (Parlos et al., 2000), in which each single forecasting step of the Markovian property-based model can only predict the next 1-day datum (Fig. 1a). However, the SSA prediction may not provide enough information, especially in the situation in which it is desirable to understand the behaviour of multiple steps in the future, such as signal processing and time-series prediction. Given this issue, the serial-propagated multi-step-ahead (MSA) prediction (Fig. 1b), which attempts to make predictions several time steps into the future without the availability of output measurements, has attracted an increasing number of scientific studies (Su et al., 1992; Schenker and Agarwal, 1995; Coulibaly et al., 2000; Gao et al., 2002; Chang et al., 2007; Yong et al., 2010; Chang et al., 2012). However, the MSA predictors, especially the long-range MSA predictors, are difficult to develop, because the lack of measurements in the prediction horizon necessitates the recursive use of SSA predictors to reach the end point on the horizon. Even small SSA prediction errors at the beginning of the horizon accumulate and propagate, often resulting in a poor prediction accuracy. Over the last 20 yr, the MSA predictor design to increase the MSA prediction accuracy has received much attention, and different types of neural networks have been used successfully for some short-range MSA predictions (Su et al., 1992; Parlos et al., 2000; Chang et al., 2007).

As mentioned above, the crucial and most difficult point in a “true” long-term forecast of hydrological signal (Yu et al., 2013) is the development of effective models to reduce the error accumulation and increase the accuracy of the long-range serial-propagated MSA prediction. In view of this, the present study designed a novel hybrid model CDW-NN, combining continuous and discrete wavelet transforms and neural networks, as the MSA predictor for effective long-term forecast of hydrological signals by broadening the prediction range in each forecast step and reducing the total iteration steps in the long-term forecasting process. In the remainder of this paper, the long-term forecast methodologies of the MSA predictor CDW-NN are presented. In the next section, the details of daily river stage data series in different hydro-stations in Yangtze River estuary, China, are presented, and the CDW-NN hybrid models are applied to the long-term forecasts of different river stage signals. The results are discussed by comparing with the performances of other pure and hybrid models in the subsequent section, and finally conclusions are drawn.

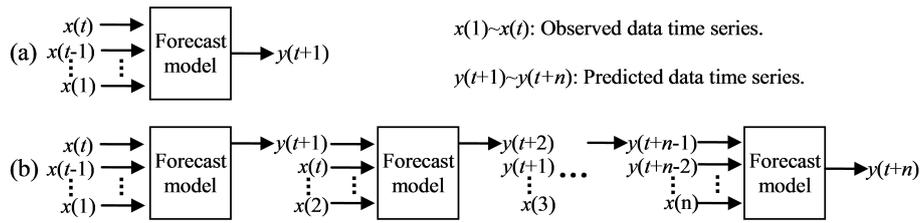


Fig. 1. Architectures of the conventional single-step-ahead (SSA) forecast (a) and the SSA-based long-range serial-propagated multi-step-ahead (MSA) forecast (b) (1000 dpi, 88 mm × 21 mm).

2 Methodologies

2.1 Continuous wavelet transform (CWT) and discrete wavelet transform (DWT)

Wavelet transform is a mathematical tool that allows the decomposition of the signal $f(t)$ in terms of elementary contributions called wavelets (Sadowskey, 1996; Labat et al., 2005). For the time series $f(t) \in L^2(R)$ or finite energy signal, the CWT of the signal $f(t)$ with the analysing wavelet ϕ is the convolution of $f(t)$ with a set of dilated and translated wavelets:

$$W_f(a, b) = \langle f(t), \varphi_{a,b}(t) \rangle = \sqrt{\frac{\delta t}{a}} \int_R f(t) \bar{\varphi}\left(\frac{t-b}{a}\right) dt, a, b \in R, a > 0, \tag{1}$$

where $\bar{\varphi}(t)$ is the complex conjugate function of $\phi(t)$, a the dilation (scale or frequency) parameter, b the translation (position or time) parameter, R the domain of real numbers, and δt the time interval of the data series. In this paper, the time interval of the data series equals 1.0 day, and the popularly used Morlet wavelet is selected as ϕ (Mallat, 1989; Daubechies, 1994; Torrence and Compo, 1998). The Morlet wavelet, which is a complex wavelet consisting of a plane wave modulated by a Gaussian function, is defined by

$$\phi(t) = \pi^{-1/4} e^{i\omega_0 t} e^{-(t^2/2)}, \tag{2}$$

where ω_0 is the non-dimensional frequency (usually taken to be 6 to satisfy the admissibility condition) (Farge, 1992).

The global wavelet power spectrum is defined as the power density at different timescale a , which is calculated by

$$E_a = \frac{1}{N} \sum_{b=1}^N |W_f(a, b)|^2, \tag{3}$$

where N is the length of the data. The signal’s periodicity can be indicated at the timescale at which the wave crest of wavelet power spectrum is observed. The significance of the global wavelet power spectrum is tested using a white or red noise model by comparing with the theoretical global

wavelet power spectrum (P). P is given as (Torrence and Compo, 1998)

$$P = \sigma^2 P_a \frac{x_v^2(p)}{v}, \tag{4}$$

where σ^2 is the variance of data series, $x_v^2(p)$ the inverse of chi-squared cumulative distribution with v degrees of freedom at the requested confidence level $1 - p$, and p the distribution fraction. For the lag-1 autocorrelation, $r(1) < 0.1$ P_a is the white noise spectrum, and for $r(1) > 0.1$, P_a is the red noise spectrum. For the Morlet wavelet, P_a is given as Eq. (5), and v is given as Eq. (6). In this study, a significance level of 0.005 was selected (e.g. $\chi_2^2(99.5\%) = 10.597$).

$$P_a = \frac{1 - r(1)^2}{1 + r(1)^2 - 2r(1) \cos\left(\frac{2\pi \delta t}{1.033a}\right)} \tag{5}$$

$$v = 2\sqrt{1 + \left(\frac{N \delta t}{2.32a}\right)^2} \tag{6}$$

The continuous wavelet (Eq. 1) is often discrete in real applications. When $a = a_0^j$, $b = k b_0 a_0^j$, $a_0 > 1$, $b_0 \in R$, and k and j are integer numbers, the DWT of $f(t)$ can be written as

$$W_f(j, k) = \frac{1}{\sqrt{a_0^j}} \int_R f(t) \bar{\varphi}\left(a_0^{-j} t - k b_0\right) dt. \tag{7}$$

Based on the commonly used Mallat algorithm for calculating discrete wavelet coefficients, the most common and simplest choice for the parameters a_0 and b_0 is 2 and 1 time steps, respectively, and the Daubechies wavelet, which has no explicit mathematical expressions and can be calculated only numerically, is commonly used in the DWT (Mallat, 1989; Daubechies, 1994; Partal and Kucuk, 2006; Kisi, 2011). For a discrete time series $f(t)$ occurring at different times t (e.g. integer time steps are used herein), the DWT can be defined as

$$W_f(j, k) = \frac{1}{\sqrt{2^j}} \sum_{t=0}^{N-1} f(t) \bar{\varphi}\left(2^{-j} t - k\right), \tag{8}$$

where N is the number of discrete time steps, and $W_f(j, k)$ is the wavelet coefficient for the discrete wavelet of scale $a = 2^j$ and time $b = 2^j k$.

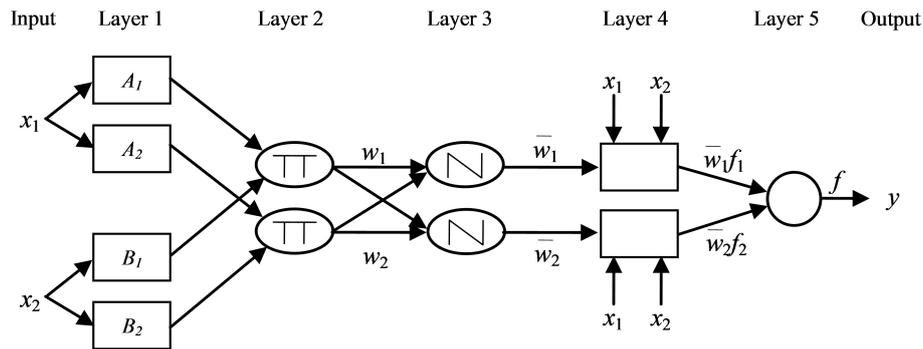


Fig. 2. Schematic diagram of a typical five-layer ANFIS structure (1000 dpi, 86 mm \times 32 mm).

2.2 Neuro-fuzzy (NF) and BP-ANN

The popular neural network (NN) model neuro-fuzzy, based on an adaptive neuro-fuzzy inference system (ANFIS) (Jang et al., 1997; Partal and Kisi, 2007), is utilised as the sub-model for the different hydro-meteorological signals forecasts in this paper. The ANFIS, first introduced by Jang (1993), is a universal approximator and, as such, is capable of approximating any real continuous function on a compact set to any degree of accuracy. The ANFIS is functionally equivalent to the Sugeno first-order fuzzy model (Jang et al., 1997; Drake, 2000), and its typical architecture with five learning layers is shown in Fig. 2. The crucial point in determining the optimal NF model structure is the selection of transfer function and rule numbers in Layer 1 of ANFIS architecture (Engin et al., 2007). Gauss and Bell functions are two commonly used functions, and the least two rule numbers of each node in Layer 1 are necessary (Jang et al., 1997). Since there is no theory yet about the suitable selection for any case, the optimal NF model structure is usually determined by many trials (i.e. trying different transfer functions and different rule numbers). Because more rule numbers may increase more complexity and calculation difficulty of the ANFIS, the rule numbers from 2 to 5 with a step size of 1 in each trial are used in testing the optimal NF model structure in the presented case study. The significant advantage of the NF model depends on the hybrid learning algorithm in ANFIS, which combines gradient descent, back-propagation, and the least-squares method and can rapidly train and adapt the ANFIS. Each learning epoch of the ANFIS is composed of a forward pass and a backward pass, and more information for neuro-fuzzy and ANFIS can be found in Jang's papers (Jang, 1993; Jang et al., 1997).

Another popular NN model, BP-ANN (back-propagation artificial neural networks), is utilised in our case to compare the forecast performance with the NF model. Based on the back-propagation algorithm, a common three-layer feed-forward type of BP-ANN is considered; the Levenberg–Marquardt methodology, which is more powerful than conventional gradient descent techniques (Hagan and Menhaj,

1994; Kisi, 2011), is used to adjust the weights of the ANN model, and the tangent sigmoid and linear activation functions are used for the hidden and output node(s), respectively. Same as the NF model structure determination, since there is no theory yet to determine how many hidden layer nodes in the BP-ANN are needed to approximate any given function, the hidden layer node number in BP-ANN is commonly determined by the trial-and-error approach (Nayak et al., 2013). Since each type of ANN model in the presented case study has two input layer nodes, the trial and error procedure starts with one hidden layer node initially, and the hidden layer node numbers are increased up to 10 with a step size of 1 in each trial. By many trials, the optimal BP-ANN structure is determined by selecting the hidden layer node number with the best model training efficiency and simulating performance.

2.3 Architecture of the long-term forecasting based on the MSA predictor CDW-NN

In order to reduce the error accumulation and increase the accuracy of the long-range serial-propagated MSA prediction, the present study designs a novel hybrid model CDW-NN, combining CWT, DWT, and NN, as the MSA predictor for effective long-term forecast of hydrological signals. The architecture of CDW-NN hybrid model is shown in Fig. 3. Firstly, for the original given daily data series $x(1) \sim x(t)$, the CWT method is utilised to reveal its short-term periodicities, i.e. the periods at $a1 \sim ai$ days in Fig. 3 ($a1 < a2 < \dots < ai$). Meanwhile, the decomposition of the original signal DWT is carried out to get new data series $TD(1) \sim TD(t)$, which is constructed by selecting and combining optimal DWT decomposition components. Then, by combining the CWT and DWT results, the new TD series at $a1 \sim ai$ days ahead ($TD(t - a1) \sim TD(t - ai)$) are selected as the NN model inputs for model training to forecast the datum $x(t)$. According to the serial-propagated prediction principle, using $TD(t - a1 + 1) \sim TD(t - ai + 1)$ as model inputs can predict the first future day datum $y(t + 1)$, and using $TD(t) \sim TD(t - ai + a1)$ as inputs can predict $y(t + a1)$.

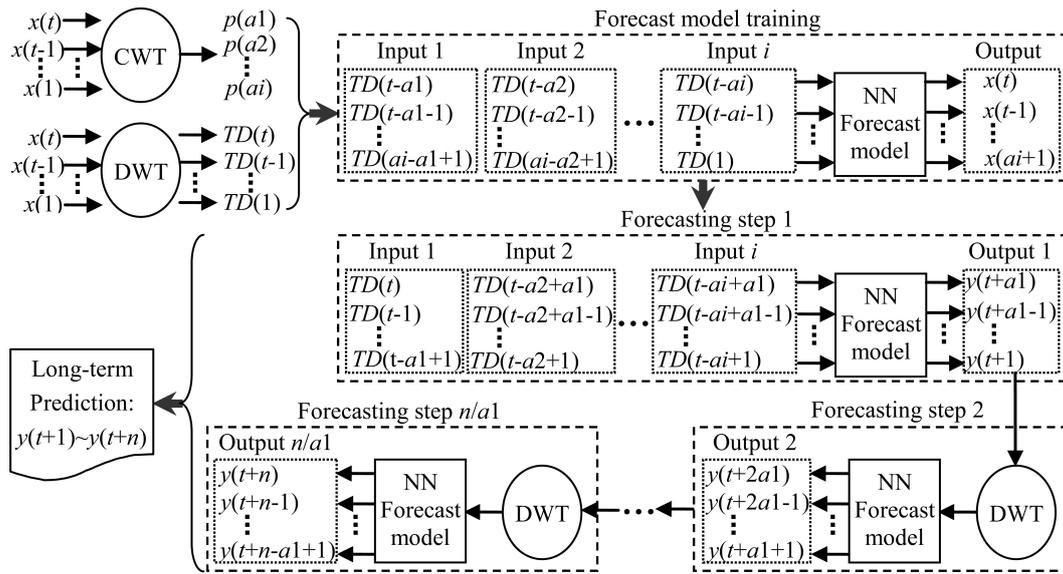


Fig. 3. Architecture of the novel CDW-NN hybrid model (1000 dpi, 132 mm × 71 mm).

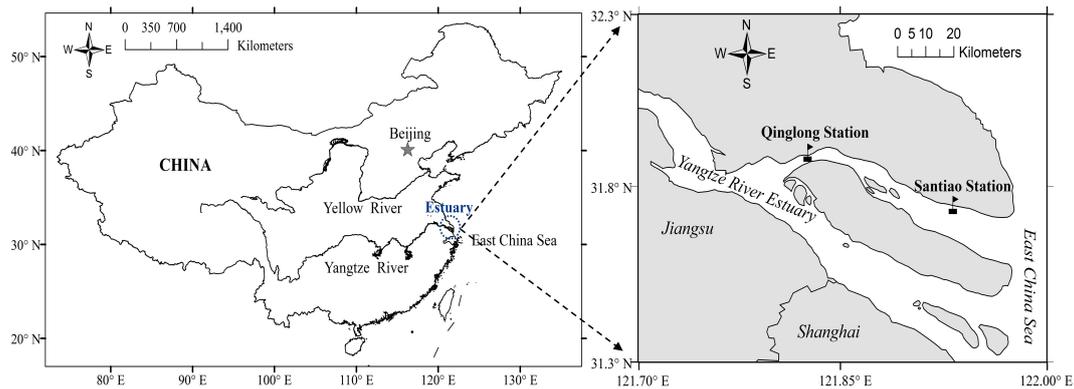


Fig. 4. Location map of the hydrological stations (1000 dpi, 148 mm × 54 mm).

Here, the first batch of outputs $y(t + 1) \sim y(t + a1)$ are predicted from the first forecasting step, and then can be used as new observations for the second step DWT decomposing and NN forecasting to predict the second batch of outputs $y(t + a1 + 1) \sim y(t + 2a1)$. Just as the 1-day prediction from each SSA forecast involved in conventional MSA forecast, the day number of predictions in the output of each CDW-NN forecasting step is $a1$. So, after about $n/a1$ steps of forecasting process, the final long-term prediction series $y(t + 1) \sim y(t + n)$ can be obtained.

3 Application

3.1 Studied area and data

The high-tide level data at two time points each day during 13 yr (4748 days) covering 1998–2010, supported by the Water Resources Department of Jiangsu Province, were

observed and collected from estuarine Santiao Port hydrologic station (31.721° N , 121.698° E) and Qinglong Port hydrologic station (31.862° N , 121.239° E), located about 19.0 and 70.0 km, respectively, upstream from Chinese Yangtze River entrance into the East China Sea (Fig. 4). The daily river stage data series at each station are obtained based on the average value of high-tide levels at two time points each day. The first 10 yr of river stage data (3652 days) are used for training and establishing hybrid models (i.e. $t = 3652$ in Fig. 3). The remaining 3 yr of river stage data (1096 days) are used for testing the long-term forecasting performance of hybrid models (i.e. $n = 1096$ in Fig. 3).

3.2 Short-term periodic features of estuarine daily river stage series by CWT

The Morlet wavelet transform coefficients of the training data series at relatively fine timescales (from 1-day to 50-day

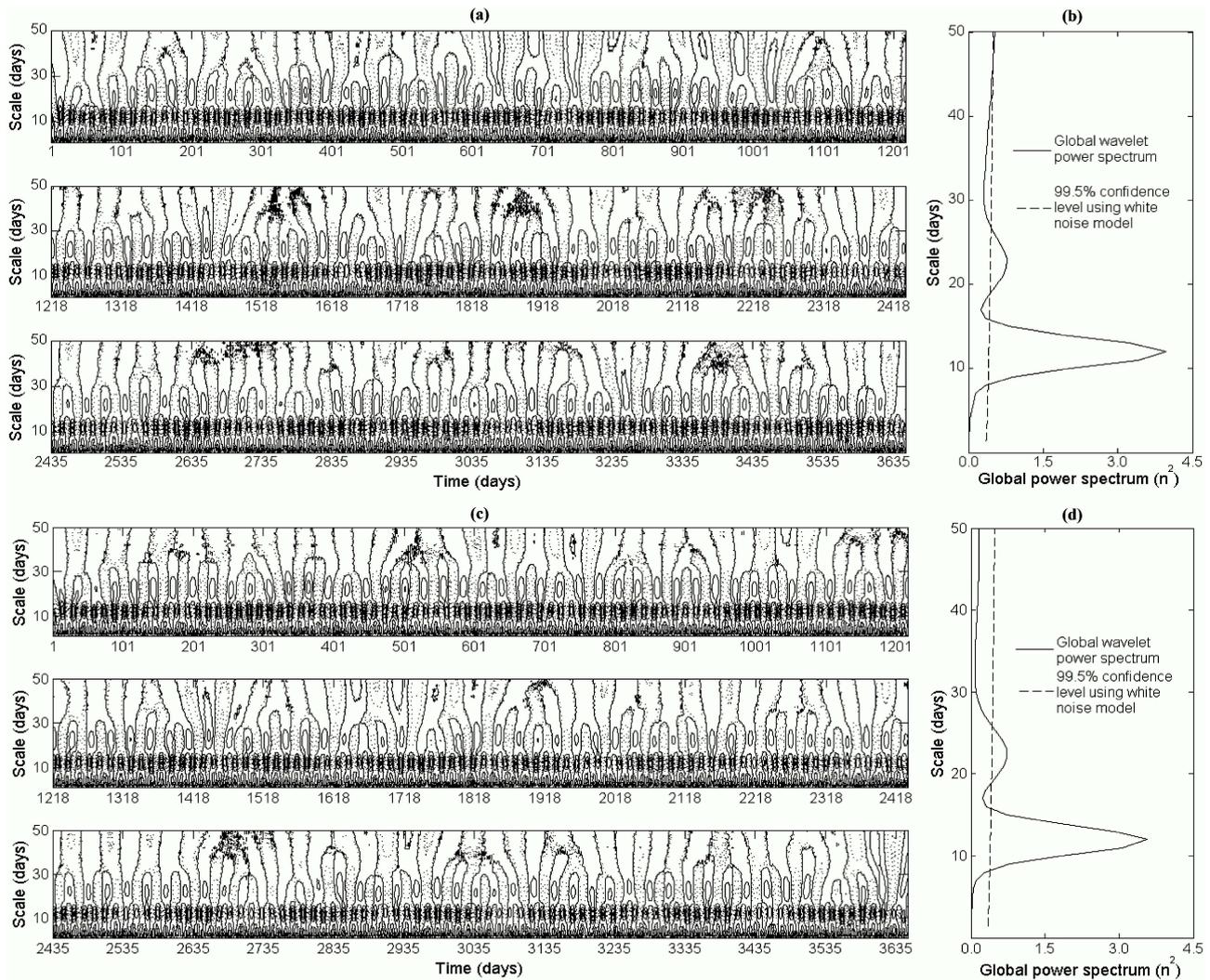


Fig. 5. Real parts of the Morlet wavelet transform coefficients at Santiao station (a) and Qinglong station (c), and their global wavelet power spectrums and corresponding confidence tests using white noise models (b, d), based on the 3652-day river stage data during 1998–2007 (1000 dpi, 139 mm \times 111 mm).

scales) are calculated by MATLAB language. The real parts of Morlet wavelet transform coefficients at Santiao station (Fig. 5a) and Qinglong station (Fig. 5c) clearly indicate the distribution conditions of the river stage signals at different timescales. The solid isograms indicate positive wavelet coefficients and a relatively high river stage period, and the dashed isograms indicate negative coefficients and a relatively low period. Further calculating the global wavelet power spectrums and the corresponding theoretical power spectrums using white noise model, the significance of wavelet power densities of daily river stage series at different timescales at Santiao and Qinglong stations is calculated and shown in Fig. 5b and d.

Results show that the Morlet wavelet transform coefficients of daily river stage series at Santiao station generate obvious two kinds of quasi-periodic oscillations (QPOs),

namely at 12-day and 23-day timescales, and both of their global wavelet power spectrums are prominent at the 99.5 % confidence level. The QPO of estuarine daily river stage at a fine timescale is often nested in a broad timescale. At the 12-day timescale, the average changing periodicity (T) of river stage time series (i.e. the average cycle days between each two time domains with positive wavelet coefficients) is 15 days obtained by calculating and averaging the day numbers of each two neighbouring high and low river stage periods. At the 23-day scale, the average T is 28 days. At Qinglong station, the Morlet wavelet transform coefficients of daily river stage series generate obvious two QPOs at 12-day and 22-day timescales, of which the corresponding T s are 15 days and 28 days, the same as that at Santiao station, and both of their global wavelet power spectrums are prominent at the 99.5 % confidence level. Based on the prominent

Table 1. Squared correlation coefficients (R^2) between each discrete wavelet component series at 15 days and 28 days prior ($t - 15$ and $t - 28$) and original river stage data series on the current day (S_t).

Discrete wavelet components in 10 resolution levels	R^2 with S_t at Santiaio station			R^2 with S_t at Qinglong station		
	Single-factor analysis at $t - 15$	Single-factor analysis at $t - 28$	Double-factor analysis at $t - 15$ and $t - 28$	Single-factor analysis at $t - 15$	Single-factor analysis at $t - 28$	Double-factor analysis at $t - 15$ and $t - 28$
A_{10}	0.005	0.005	0.005	0.012	0.012	0.012
D_1	0.000	0.000	0.000	0.000	0.000	0.000
D_2	0.003	0.002	0.006	0.000	0.000	0.001
D_3	0.430	0.309	0.501	0.399	0.299	0.468
D_4	0.067	0.144	0.228	0.052	0.147	0.221
D_5	0.009	0.000	0.009	0.008	0.000	0.009
D_6	0.003	0.000	0.007	0.003	0.000	0.006
D_7	0.016	0.008	0.019	0.016	0.008	0.019
D_8	0.102	0.085	0.111	0.135	0.112	0.146
D_9	0.006	0.005	0.007	0.008	0.006	0.013
D_{10}	0.007	0.006	0.007	0.000	0.000	0.001
$S = A_{10} + \sum D_i$	0.549	0.510	0.672	0.555	0.558	0.674
$TD = D_3 + D_4 + D_8$	0.576	0.539	0.681	0.557	0.564	0.699

short-term periodic features of estuarine daily river stage time series, estuarine daily river stages at 15 days prior and 28 days prior are determined to simulate and forecast the river stage at the current day (i.e. a_1 and a_2 in Fig. 3 equal 15 and 28, respectively).

3.3 Decomposition of daily river stage time series and optimal DWT components combination

The decomposition process of DWT consists of a number of filtering steps following the Mallat algorithm. The original signal of training data series is first decomposed into an approximation (A_1) and details (D_1), and the A_1 is then broken down into many lower resolution components (A_i and D_i). The details are the low-scale high-frequency components of the signal, while the approximations are the high-scale low-frequency components. The higher scales consist of the extended version of a wavelet, and the corresponding coefficients refer to the slowly changing coarse features of low-frequency components. The lower scales present the condensed wavelet and follow the rapidly changing details (high-frequency components) of the signal (Mallat, 1989). In our case, 10 decomposition components (A_{10} , D_1 – D_{10}) of the original daily river stage signals (1998–2007) in 10 resolution levels are calculated by MATLAB language, and then are used to analyse the optimum input factors in the hybrid models.

The squared correlation coefficients (R^2) between each discrete wavelet component (A_{10} , D_1 – D_{10}) at 15 days and 28 days prior ($t - 15$ and $t - 28$) and the original time series at time t (S_t) are computed and presented in Table 1. By the single-factor and double-factor analysis, at both Santiaio and Qinglong stations, the D_3 , D_4 and D_8 components at 15 days and 28 days prior show the prominently higher R^2 with S_t than the other DWT components, especially significant in

the double-factor analysis. Instead of using each DWT component individually as the model input, employment of the added suitable DWT components is more useful and can highly increase the forecast performance. Based on the revealed dominant DWT components of different hydrological series, the new series (TD) obtained by adding D_3 , D_4 and D_8 at 15 days and 28 days prior are selected as two NN model inputs for the daily river stage forecast at both Santiaio and Qinglong stations. Comparing the lagged TD series with the lagged original series (S), the lagged new series (TD) show slightly higher correlations at both 15-day and 28-day delay time nodes with S_t , which indicate that the new TD series keeps the main information of the original signal dynamics in spite of the filtering of many other weakly correlated information by DWT.

3.4 CDW-NN model training and long-term forecasting of daily river stage signals

According to the above CWT and DWT results, two new daily series TD(1)~TD($t - 28$) and TD(14)~TD($t - 15$) extracted from the training data series are used as the NN model inputs to simulate and forecast the original series $x(29) \sim x(t)$. Program codes were written in MATLAB language for training the neuro-fuzzy and BP-ANN submodels and determining their optimal model structures. At Santiaio station, by many trials the optimal CDW-NF hybrid model structure is determined as CDW-NF (5-Bell), which denotes the Bell-type transfer function and five rules for each input in the layer one of ANFIS, and the optimal CDW-ANN hybrid model structure is determined as CDW-ANN(2-3-1), which denotes two input layer nodes, three hidden layer nodes and one output layer node in the BP-ANN submodel. At Qinglong station the optimal CDW-NF(4-Gauss) model and the optimal CDW-ANN(2-4-1) model are obtained. After

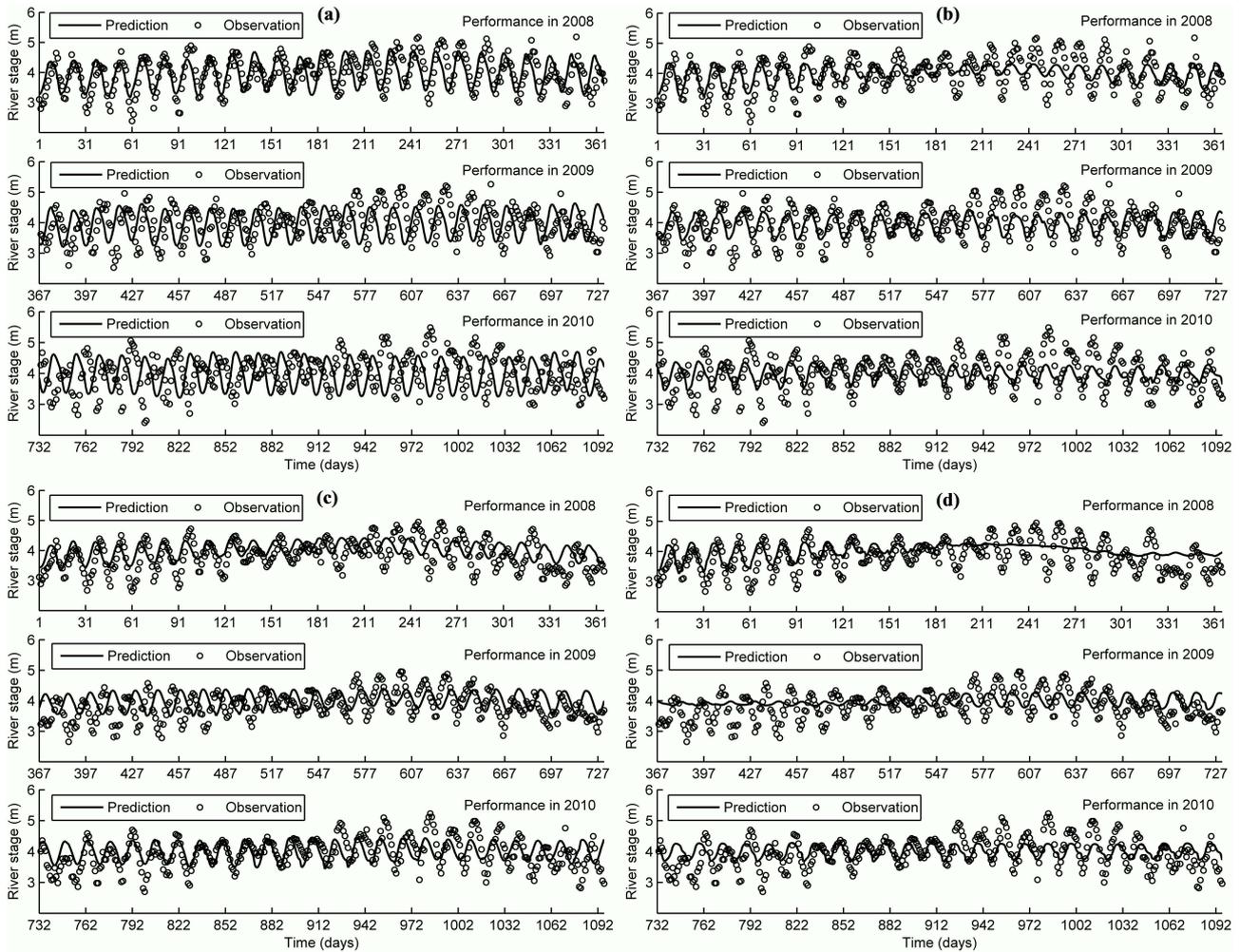


Fig. 6. Observed and predicted values of the daily river stages from 2008 to 2010 using the CDW-ANN hybrid model and CDW-NF hybrid model at Santiaog station (a, b) and Qinglong station (c, d) (1000 dpi, 136 mm × 105 mm).

the training processes of NN submodels, two TD series $TD(t-27) \sim TD(t-13)$ and $TD(t-14) \sim TD(t)$ are used as model inputs to forecast the future 15-day river stage series $y(t+1) \sim y(t+15)$ (i.e. the first forecasting step in Fig. 3). Then the predicted 15-day river stage series are treated by DWT and used as the new inputs to forecast the next 15-day river stage in the next forecasting step. Because each single forecasting step using CDW-NN model obtains 15-day predictions, the total 1096-day future river stage values from 2008 to 2010 are predicted after 74 steps of forecasting.

The root mean square errors (RMSEs), mean absolute errors (MAEs) and correlation coefficient (R) statistics are used to evaluate the model performance of simulation and prediction. The RMSE and MAE are defined as

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_{i\text{observed}} - Y_{i\text{estimate}})^2}, \quad (9)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |Y_{i\text{observed}} - Y_{i\text{estimate}}|, \quad (10)$$

where n is the number of data sets, and Y_i is the daily river stage.

The forecast performances of CDW-NF and CDW-ANN models at Santiaog and Qinglong stations are shown in Fig. 6. Results show that the CDW-NF models perform the significantly better correlations between the observed and predicted river stage data during 2008–2010 with the higher R^2 of 0.284 and 0.173 at Santiaog and Qinglong stations, respectively, while the CDW-ANN with the lower R^2 of 0.020 and 0.030 at Santiaog and Qinglong stations, respectively. The CDW-ANN hybrid model shows better forecast performances during the first year than that in the last two years.

Table 2. Comparison among the performances of 12 types of river stage forecasting models in respect to root mean square errors (RMSEs), mean absolute errors (MAEs) and R^2 in the training and test periods.

Model	Santiao station						Qinglong station					
	Training period			Test period			Training period			Test period		
	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2
CDW-NF	0.314	0.250	0.702	0.476	0.381	0.284	0.310	0.245	0.698	0.457	0.372	0.173
CDW-ANN	0.318	0.254	0.696	0.774	0.635	0.020	0.313	0.247	0.692	0.527	0.430	0.030
CDW-LR	0.325	0.261	0.681	0.562	0.463	0.012	0.322	0.255	0.674	0.499	0.410	0.014
CW-NF	0.316	0.250	0.697	0.623	0.494	0.033	0.302	0.283	0.713	0.570	0.459	0.009
CW-ANN	0.322	0.255	0.687	0.702	0.568	0.000	0.309	0.244	0.701	0.505	0.409	0.002
CW-LR	0.329	0.262	0.672	0.565	0.463	0.005	0.309	0.244	0.699	0.487	0.397	0.011
DW-NF	0.211	0.166	0.865	0.574	0.475	0.039	0.211	0.169	0.860	0.471	0.384	0.081
DW-ANN	0.214	0.168	0.861	0.568	0.468	0.037	0.214	0.171	0.857	0.474	0.386	0.066
DW-LR	0.225	0.177	0.848	0.554	0.454	0.036	0.226	0.181	0.839	0.492	0.400	0.053
Neuro-fuzzy	0.263	0.198	0.791	0.558	0.454	0.009	0.227	0.168	0.838	0.499	0.404	0.012
BP-ANN	0.271	0.203	0.780	0.557	0.451	0.016	0.226	0.168	0.839	0.497	0.402	0.009
LR	0.284	0.216	0.757	0.559	0.456	0.013	0.242	0.180	0.815	0.498	0.403	0.012

4 Discussion

4.1 Forecast performance comparison among CDW-NN models and the other 10 types of hybrid and pure models

Similar to the establishing process of CDW-NN models, the hybrid CDW-LR model is established by combining CWT, DWT and a linear regression (LR) model. Using the training data series, the simulating equations of the CDW-LR models at Santiao station and Qinglong station are obtained and shown as Eqs. (11) and (12), respectively. Without the DWT treatment, three other CWT-based (CW-) models are established by using the original daily river stage series at 15 days and 28 days prior (S_{t-15} and S_{t-28}) as model inputs to simulate S_t . By the model training processes, the optimal CW-NF structures for Santiao and Qinglong stations are determined as CW-NF(5-Bell) and CW-NF(2-Gauss), respectively; the optimal CW-ANN structures for Santiao and Qinglong stations are CW-ANN(2-3-1) and CW-ANN(2-1-1), respectively; and the CW-LR models for Santiao and Qinglong stations are shown as Eqs. (13) and (14), respectively.

$$S_t = 0.454TD_{t-28} + 0.533TD_{t-15} + 3.966 \quad (11)$$

$$S_t = 0.496TD_{t-28} + 0.487TD_{t-15} + 4.004 \quad (12)$$

$$S_t = 0.428S_{t-28} + 0.493S_{t-15} + 0.315 \quad (13)$$

$$S_t = 0.473S_{t-28} + 0.460S_{t-15} + 0.266 \quad (14)$$

As mentioned in Fig. 1, many conventional studies on MSA predictor design focused on the methodology of recursive use of SSA predictors (i.e. generally using S_{t-1} and S_{t-2} as model inputs to forecast S_t). In view of this, six types of conventional SSA-based long-term forecast hybrid and pure models are established for comparing with the CWT-based hybrid models. Among the six types of models, the

DW-R, DW-ANN and DW-NF hybrid models utilise the optimum decomposition components combinations (TD_{t-1} and TD_{t-2}), determined by DWT, as model inputs to forecast S_t . With respect to the pure LR, BP-ANN and neuro-fuzzy models, the original daily river stage series at 1 day prior and 2 days prior (S_{t-1} and S_{t-2}) are used as model inputs to forecast S_t . By the model training processes, the optimal DW-NF structures for Santiao and Qinglong stations are determined as DW-NF(5-Gauss) and DW-NF(5-Bell), respectively; the optimal DW-ANN structures for Santiao and Qinglong stations are DW-ANN(2-5-1) and DW-ANN(2-7-1), respectively; the optimal pure NF model structures for Santiao and Qinglong stations are determined as NF(5-Bell) and NF(2-Gauss), respectively; and the optimal pure BP-ANN model structures for Santiao and Qinglong stations are BP-ANN(2-3-1) and BP-ANN(2-5-1), respectively. The DW-LR models for Santiao and Qinglong stations are obtained and shown as Eqs. (15) and (16), respectively, and the pure LR models for Santiao and Qinglong stations are shown as Eqs. (17) and (18), respectively.

$$S_t = -0.973TD_{t-2} + 1.818TD_{t-1} + 3.966 \quad (15)$$

$$S_t = -0.971TD_{t-2} + 1.826TD_{t-1} + 4.005 \quad (16)$$

$$S_t = -0.316S_{t-2} + 1.124S_{t-1} + 0.761 \quad (17)$$

$$S_t = -0.397S_{t-2} + 1.234S_{t-1} + 0.649 \quad (18)$$

As shown in Table 2, the training and forecasting performances of 12 types of hybrid and pure models are compared with each other in respect to the RMSE, MAE and R^2 statistics. Due to the high lag-1 and lag-2 autocorrelations in the hydrological time series, in the training periods the six types of conventional SSA-based long-term forecast models perform better than the six types of CWT-based models with the higher R^2 and smaller RMSE and MAE. Nevertheless, in the test periods, without the observation data as model inputs

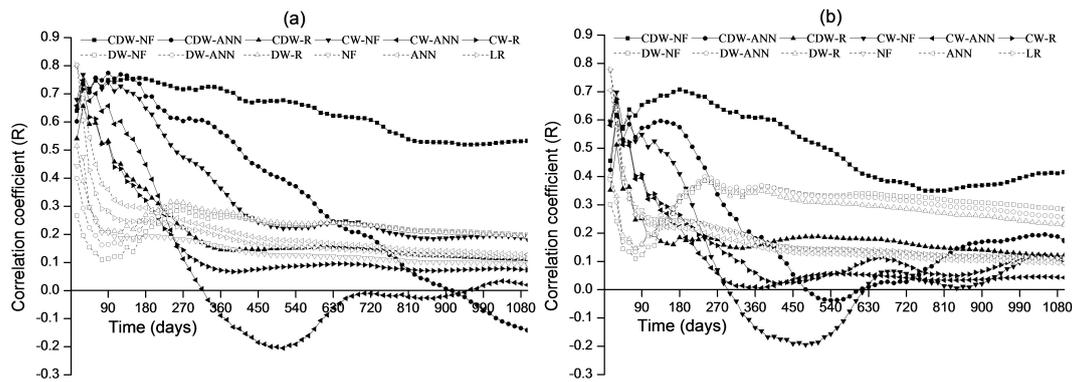


Fig. 7. Forecast performances in respect to the correlation coefficients (R) of 12 types of hybrid models during the 1096-day river stages forecasting at Santiaog station (a) and Qinglong station (b) (1000 dpi, 132 mm \times 46 mm).

in each forecasting step, the CDW-NF hybrid model shows the most significant performances among all the 12 models, especially at Santiaog station. In addition, due to the prominent ability of the decomposition approach DWT in filtering weakly correlated details from original signal, each DWT-based hybrid model performs better than the corresponding model without DWT both in the training and test periods.

4.2 Driving mechanism of the advantages of CDW-NN models in long-range MSA predictions

Prediction performance details in respect to R for the 12 types of hybrid and pure models at different forecasting steps during the overall 1096-day river stage forecasting are calculated and shown in Fig. 7. According to the serial-propagated MSA prediction theory, the error accumulation increases with the iteration step increase in a long-term forecast. Therefore, the prediction performances of all 12 types of models have overall decreasing trends with the increasing of predicted data length from 1 day to 1096 days. Nevertheless, during approximately the first 200–600 days of river stage forecasting, all six types of CWT-based models performed better than the other six types of SSA-based models. In particular, both at Santiaog and Qinglong stations the CDW-NF models show the significantly better performances than the other models during the overall 3 yr (1096 days) river stage forecasting steps. The main explanations are that the CWT-based models reduce the overall forecasting iteration steps to 74 steps by using the 15-day prior data series as the first model input, while the conventional SSA-based models needs 1096 steps by using 1-day prior data series as the first model input. The prominent decrease of forecasting steps consequently brings significant reduction of error accumulation in the long-range MSA prediction. In addition, the combination of the advantages of the DWT method and neuro-fuzzy system also benefits weakening the noisy dynamics of the model inputs and enhancing the simulation and forecast ability for the complex hydro-system.

In view of the above discussion, the CDW-NF hybrid model is proven to be the most effective MSA predictor for the long-term forecasts of estuarine daily river stage signals. In addition, the CDW-ANN hybrid model can be taken as the second selection for the short-term and mid-term forecasts of estuarine daily river stage signals because of its high performance during the first year forecast process. It should be noted that the methodology of accurate MSA predictor design by reducing iteration steps and error accumulation is innovated in this study by revealing the short-time periodic features of estuarine river stage dynamics, which is mainly caused by the half-month periodicity involved in the astronomical tidal fluctuation in river estuary. With respect to other kinds of hydro-meteorological signals, which have non-significant short-term periodic features, other types of algorithms and models for reducing error accumulation in long-term forecasting steps need to be further studied in future research. In addition, although the novel CDW-NF hybrid model succeeds in prominently increasing the long-range MSA prediction accuracy in R^2 , RMSE and MAE, the overall relatively poor performances of most of models in the “true” long-term forecast lead to the difficulties in the accurate peak value estimation. Thus, the improvement of modelling capabilities in hydro-signal’s peak value estimation and physical process revealed by developing more advanced algorithms still needs further research.

5 Conclusions

Studies on the long-term forecast of hydrological signals have high practical value. However, the accurate long-range MSA predictor design is very difficult, especially in conventional SSA-based MSA predictions because of the huge error accumulation in the serial-propagated long-term forecast. In this study, we design a novel hybrid model CDW-NN, combining continuous and discrete wavelet transforms and neural networks, as the MSA predictor for effective long-term forecasts of hydrological signals. By the application of CDW-NN

hybrid models and the other 10 types of hybrid and pure models in estuarine daily river stage series long-term forecasts, the 1096-day estuarine river stage data are forecasted, and the superiorities of CDW-NN models with corresponding driving mechanisms are proven as follows.

1. Comparing the conventional SSA-based models, the CWT-based hybrid models broaden the prediction range in each forecast step from conventional 1 day to now 15 days and reduce the overall forecasting iteration steps from conventional 1096 steps to now 74 steps by using the 15-day and 28-day prior data series as model inputs, which is determined by revealing the hydro-signal's significant short-time periodicities using CWT. The prominent reduction of forecast steps has created significant decrease of error accumulations and increase of long-term forecast performances in the CWT-based hybrid models.
2. Among the CWT-based models, one type of CDW-NN model, CDW-NF, has been proven to be the most effective MSA predictor for the prominent accuracy enhancement during the overall 1096-day long-term forecasts of estuarine hydro-signals. The other type of CDW-NN model, CDW-ANN, has been proven to be the second selection for short-term and mid-term forecasts of estuarine hydro-signals. The main explanation is the combination of the advantages of the CWT and DWT methods and neuro-fuzzy system in reducing the error accumulation, filtering weakly correlated details from original signal and weakening the noisy dynamics of the model inputs, and enhancing the simulation and forecast ability for the complex hydro-system.
3. It should be noted that because the successful application of the novel CDW-NF model in hydro-signals long-term forecasts largely depends on the significant short-term periodicities involved in estuarine hydro-signals, some other innovative algorithms and models still need to be further studied in future research for other kinds of hydro-meteorological signals without significant short-term periodic features.

Acknowledgements. This study is financially supported by the National Natural Science Foundation of China (no. 41101518), the National Basic Research Program of China (no. 2010CB429001), the Natural Science Foundation of Jiangsu Province of China (no. BK2010610, no. BK2011425 and no. BY2013062), and the National Non-Profit Research Program of China (no. 200903001 and no. 2011418020). Cordial thanks should be extended to the editors and reviewers for their invaluable comments, which greatly improved the quality of this paper.

Edited by: E. Toth

References

- Ancil, F. and Tape, D. G.: An exploration of artificial neural network rainfall-runoff forecasting combined with wavelet decomposition, *J. Environ. Eng. Sci.*, 3, S121–S128, 2004.
- ASCE Task Committee: Artificial neural network in hydrology, *J. Hydrol. Eng.-ASCE*, 5, 124–144, 2000.
- Bolch, G., Greiner, S., de Meer, H., and Trivedi, K. S.: *Queueing Networks and Markov Chains*, John Wiley, New York, 2006.
- Cao, H. Q. and Park, H. D.: Elucidation and short-term forecasting of microcystin concentrations in Lake Suwa (Japan) by means of artificial neural networks and evolutionary algorithms, *Water Res.*, 41, 2247–2255, 2007.
- Chang, F. J. and Chang, Y. T.: Adaptive neuro-fuzzy inference system for prediction of water level in reservoir, *Adv. Water Res.*, 29, 1–10, 2006.
- Chang, F. J., Chiang, Y. M., and Chang, L. C.: Multi-step-ahead neural networks for flood forecasting, *Hydrolog. Sci. J.*, 52, 114–130, 2007.
- Chang, L. C., Chen, P. A., and Chang, F. J.: A Reinforced Two-Step-Ahead Weight Adjustment Technique for On-Line Training of Recurrent Neural Networks, *IEEE T. Neural Networ. Learn. Syst.*, 23, 1269–1278, 2012.
- Chen, S. Y.: *Theories and methods of variable fuzzy sets in water resources and flood control system*, Dalian University of Technology Press, Dalian, 2005.
- Chou, C. M.: Applying multi-resolution analysis to differential hydrological grey models with dual series, *J. Hydrol.*, 332, 174–186, 2007.
- Cigizoglu, H. K.: Application of the generalized regression neural networks to intermittent flow forecasting and estimation, *J. Hydrol. Eng.-ASCE*, 10, 336–341, 2005.
- Coulibaly, P. and Burn, H. D.: Wavelet analysis of variability in annual Canadian streamflows, *Water Resour. Res.* 40, W03105, doi:10.1029/2003WR002667, 2004.
- Coulibaly, P., Ancil, F., and Bobee, B.: Neural network-based long-term hydropower forecasting system, *Comput.-Aided Civ. Inf.*, 15, 355–364, 2000.
- Daubechies, I.: *Ten lectures on wavelets*, CBMS, SIAM, 61, 194–202, 1994.
- Deng, J. L.: *Grey prediction and grey decision*, Hua Zhong Li Gong University Press, Wuhan, 1992.
- Drake, J. T.: *Communications Phase Synchronization Using the Adaptive Network Fuzzy Inference System*, PhD Thesis, New Mexico State University, Las Cruces, New Mexico, 2000.
- El-Shafie, A., Taha, M. R., and Noureldin, A.: A neuro-fuzzy model for inflow forecasting of the Nile river at Aswan high dam, *Water Resour. Manage.*, 21, 533–556, 2007.
- Engin, A., Davut, H., and Asaf, V.: An expert Discrete Wavelet Adaptive Network Based Fuzzy Inference System for digital modulation recognition, *Expert Syst. Appl.*, 33, 582–289, 2007.
- Farge, M.: Wavelet transforms and their applications to turbulence, *Annu. Rev. Fluid Mech.*, 24, 395–457, 1992.
- French, M. N., Krajewski, W. F., and Cuykendal, R. R.: Rainfall forecasting in space and time using a neural network, *J. Hydrol.*, 137, 1–37, 1992.
- Gao, X. Z., Ovaska, S. J., and Vasilakos, A. V.: Temporal difference method-based multi-step-ahead prediction of long term deep fading in mobile networks. *Comput. Commun.* 25, 1477–1486, 2002.

- Hagan, M. T. and Menhaj, M. B.: Training feed forward networks with the Marquardt algorithm, *IEEE Trans. Neural Netw.*, 6, 861–867, 1994.
- Hua, B., Xiong, W., and Cheng, H.: Application of fuzzy support vector machine to runoff forecast, *Eng. J. Wuhan Univ.*, Wuhan, 41, 5–8, 2008.
- Islam, M. N. and Sivakumar, B.: Characterization and prediction of runoff dynamics: a nonlinear dynamical view, *Adv. Water Resour.*, 25, 179–190, 2002.
- Jain, S. K., Das, D., and Srivastava, D.: Application of ANN for reservoir inflow prediction and operation, *J. Water Resour. Pl.-ASCE*, 125, 263–271, 1999.
- Jang, J. S. R.: ANFIS: adaptive-network-based fuzzy inference system, *IEEE T. Syst. Man. Cy.*, 23, 665–685, 1993.
- Jang, J. S. R., Sun, C. T., and Mizutani, E.: *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*, Prentice Hall, Upper Saddle River, New Jersey, 1997.
- Jayawardena, A. W. and Lai, F.: Analysis and prediction of chaos in rainfall and streamflow time series, *J. Hydrol.*, 153, 23–52, 1994.
- Karunasinghe, D. S. K. and Liong, S. Y.: Chaotic time series prediction with a global model: artificial neural network, *J. Hydrol.*, 153, 23–52, 2006.
- Kim, T. W. and Valdes, J. B.: Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks, *J. Hydrol. Eng.-ASCE*, 6, 319–328, 2003.
- Kisi, O.: Suspended sediment estimation using neuro-fuzzy and neural network approaches, *Hydrolog. Sci. J.*, 50, 683–696, 2005.
- Kisi, O.: Stream flow forecasting using neuro-wavelet technique, *Hydrol. Process.*, 22, 4142–4152, 2008.
- Kisi, O.: Wavelet Regression model as an alternative to neural networks for river stage forecasting, *Water Resour. Manage.*, 25, 579–600, 2011.
- Krishna, B., Rao, Y. R. S., and Nayak, P. C.: Time series modeling of river flow using wavelet neural networks, *J. Water Resour. Prot.*, 3, 50–59, 2011.
- Labat, D.: Recent advances in wavelet analyses: Part 1. A review of concepts, *J. Hydrol.*, 314, 275–288, 2005.
- Labat, D., Ronchail, J., and Guyot, J. L.: Recent advances in wavelet analyses: Part 2. Amazon, Parana, Orinoco and Congo discharges time scale variability, *J. Hydrol.*, 314, 289–311, 2005.
- Liong, S. Y. and Sivapragasm, C.: Flood stage forecasting with SVM, *J. Am. Water Resour. Assoc.*, 38, 173–186, 2002.
- Maier, H. R., Jain, A., Dandy, G. C., and Sudheer, K. P.: Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions, *Environ. Model. Softw.*, 25, 891–909, 2010.
- Mallat, S. G.: A theory for multi resolution signal decomposition: the wavelet representation, *IEEE T. Pattern Anal. Mach. Intell.*, 11, 674–693, 1989.
- Marzano, F. S., Fionda, E., and Ciotti, P.: Neural-network approach to ground-based passive microwave estimation of precipitation intensity and extinction, *J. Hydrol.*, 328, 121–131, 2006.
- Nayak, P. C., Sudheer, K. P., Rangan, D. M., and Ramasastri, K. S.: A neuro-fuzzy computing technique for modeling hydrological time series, *J. Hydrol.*, 291, 52–66, 2004.
- Nayak, P. C., Sudheer, K. P., Rangan, D. M., and Ramasastri, K. S.: Short-term flood forecasting with a neurofuzzy model, *Water Resour. Res.*, 41, W04004, doi:10.1029/2004WR003562, 2005.
- Nayak, P. C., Sudheer, K. P., and Jain, S. K.: Rainfall-runoff modeling through hybrid intelligent system, *Water Resour. Res.*, 43, W07415, doi:10.1029/2006WR004930, 2007.
- Nayak, P. C., Venkatesh, B., Krishna, B., and Jain, S. K.: Rainfall runoff modelling using conceptual, data driven and wavelet based computing approach, *J. Hydrol.*, 493, 57–67, 2013.
- Olson, D. A., Junker, N. W., and Korty, B.: Evaluation 33 years of quantitative precipitation forecasting at NMC, *Weather Forecast.*, 10, 498–511, 1995.
- Ozger, M., Mishra, A. K., and Singh, V. P.: Long lead time drought forecasting using a wavelet and fuzzy logic combination model: a case study in Texas, *J. Hydrometeorol.*, 13, 284–297, 2012.
- Parlos, A. G., Rais, O. T., and Atiya, A. F.: Multi-step-ahead prediction using dynamic recurrent neural networks, *Neural Networks*, 13, 765–786, 2000.
- Partal, T. and Kisi, O.: Wavelet and neuro-fuzzy conjunction model for precipitation forecasting, *J. Hydrol.*, 342, 199–212, 2007.
- Partal, T. and Kucuk, M.: Long-term trend analysis using discrete wavelet components of annual precipitations measurements in Marmara region (Turkey), *Phys. Chem. Earth*, 31, 1189–1200, 2006.
- Raman, H. and Sunilkumar, N.: Multivariate modeling of water resource time series using artificial neural network, *Hydrolog. Sci. J.*, 40, 145–163, 1995.
- Sadowsky, J.: Investigation of signal characteristics using the continuous wavelet transform, *J. Hopkins APL Tech. D.*, 17, 258–269, 1996.
- Salas, J. D., Delleur, J. W., Yevjevich, V., and Lane, W. L.: Applied modeling of hydrologic time series, *Water Resouces Pub.*, Michigan, 105–108, 1980.
- Schenker, B. and Agarwal, M.: Long-range prediction for poorly-known systems, *Int. J. Control.*, 62, 227–238, 1995.
- Smith, L. C., Turcotte, D. L., and Isacks, B.: Stream flow characterization and feature detection using a discrete wavelet transform, *Hydrol. Process.*, 12, 233–249, 1998.
- Su, H. T., McAvoy, T. J., and Werbos, P.: Long-term prediction of chemical processes using recurrent neural networks: a parallel training approach. *Indust. Appl. Chem. Eng. Res.*, 31, 1338–1352, 1992.
- Tong, H.: *Non-linear time series: a dynamical system approach*, Oxford University Press, Oxford, 1990.
- Torrence, C. and Compo, G. P.: A practical guide to wavelet analysis, *B. Am. Meteorol. Soc.*, 79, 61–78, 1998.
- Wang, W., Van Gelder, P. H. A. J. M., Vrijling, J. K., and Ma, J.: Forecasting daily streamflow using hybrid ANN models, *J. Hydrol.*, 324, 383–399, 2006.
- Wang, W. S., Xiang, H. L., and Ding, J.: Prediction of hydrology and water resources with nearest neighbor bootstrapping regressive model, *Water Resour. Pow.*, 19, 8–14, 2001.
- Yang, L. N., Peng, L., Zhang, L. M., Zhang, L. L., and Yang, S. S.: A prediction model for population occurrence of paddy stem borer (*Scirpophaga incertulas*), based on Back Propagation Artificial Neural Network and Principal Components Analysis, *Comput. Electron. Agr.*, 68, 200–206, 2009.

- Yong, S., Li, Y. B., Wang, Q., and Li, C. H.: Multi-steps prediction of chaotic time series based on echo state network, in: Proceedings of the 5th International Conference on Bio-Inspired Computing: Theories and Applications, BIC-TA 2010, 8–10 September and 23–26 September 2010, Liverpool, UK/Changsha, China, 2, 669–672, 2010.
- Yu, S. P., Yang, J. S., Liu, G. M., and Zou, P.: Simulation and prediction of soil salt dynamics in the Yangtze River Estuary with BP artificial neural network, *Soil*, 40, 976–979, 2008.
- Yu, S. P., Yang, J. S., and Liu, G. M.: A novel discussion on two long-term forecast mechanisms for hydro-meteorological signals using hybrid wavelet-NN model, *J. Hydrol.*, 497, 189–197, 2013.
- Zhou, H. C., Peng, Y., and Liang, G. H.: The research of monthly discharge predictor-corrector model based on wavelet decomposition, *Water Resour. Manage.*, 22, 217–227, 2008.
- Zou, P., Yang, J. S., Fu, J. R., Liu, G. M., and Li, D. S.: Artificial neural network and time series models for predicting soil salt and water content, *Agr. Water Manage.*, 97, 2009–2019, 2010.