A hybrid real-time tidal prediction mechanism based on harmonic method and variable structure neural network

Jian-Chuan Yin a,*, Ni-Ni Wang b,c, Jiang-Qiang Hu a

a Navigation College, Dalian Maritime University, Dalian 116026, Liaoning, PR China
b Institute of Geographic Science and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, PR China
c Department of Mathematics, Dalian Maritime University, Dalian 116026, Liaoning, PR China

ARTICLE INFO

Article history:
Received 18 December 2013
Received in revised form
31 July 2014
Accepted 2 March 2015
Available online 21 March 2015

Keywords:
Tidal prediction
Hybrid model
Variable neural network
Harmonic method
Sliding data window

ABSTRACT

Accurate real-time tidal prediction is essential for human activities in coastal and marine fields. Tidal changes are influenced not only by periodic revolutions of celestial bodies but also by time-varying meteorological factors. For accurate real-time tidal prediction, a hybrid prediction mechanism is constructed by taking both advantages of harmonic analysis and neural network. In the proposed mechanism, conventional harmonic analysis is employed for representing the influences of celestial factors; and neural network is used for representing the nonlinear influences of meteorological factors. Furthermore, to represent time-varying tidal dynamics influenced by meteorological factors, a variable neural network is real-time constructed with the neurons and the connecting parameters are adaptively adjusted based on a sliding data window (SDW). The hybrid prediction method uses only the latest short-period data to generate predictions sequentially. Hourly tidal data measured at four American tidal stations are used to validate the effectiveness of the hybrid sequential tidal prediction model. Simulation results of tidal prediction demonstrate that the proposed model can generate accurate short-term prediction of tidal levels at very low computational cost.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Tidal prediction is an important issue in areas of coastal construction design, tidal energy utilization, ocean natural calamities prevention and military affairs (Fang et al., 1986). Accurate real-time tidal prediction is also vital for marine safety, such as stipulating ship operation schedule of navigating over shallow waters or through an overhead bridge. Harmonic analysis, where the tide is expressed as superposition of several sinusoidal constituents (Pugh, 1987), is the most commonly used tidal prediction method and still the basis for long-term tidal prediction (Lee, 2004). Apart from relative motion of celestial bodies and costal topography, tidal level is also influenced by meteorological factors such as atmospheric pressure, wind, rainfall, ice etc. However, the harmonic analysis is made for average meteorological conditions by neglecting the real-time hydro-meteorological effects. Variation in sea level is a complex outcome of many environmental forces, and the neglect of influences of meteorological factors may cause the prediction values diverge from the actual ones (Ghorbani et al., 2010). These differences can be very large under extreme weather conditions such as storm surges or strong winds. Furthermore, meteorological factors’ influences on tide are complex and are hard to be accurately represented by strictly founded mathematical model, which presents challenges for accurate tidal prediction.

In recent years, artificial neural network (ANN) became increasingly popular in areas of prediction attributes to its merits such as nonlinearity, adaptivity, arbitrary approximation capability and parallel information processing mechanism (Haykin, 1999). As a kind of data-driven technique, neural network can represent complex nonlinear and non-periodic relationships based on the learning of data samples, which is superior to the harmonic method and suitable for tidal predictions.

A fundamental principle in data modeling is to incorporate useful a priori information regarding the underlying data generating mechanism into the modeling process (Benhammadi et al., 2011; Chen et al., 2011). The topology of neural prediction model is of great importance for the predictive efficiency (Günaydın, 2008). Inclusion of more useful information which makes contribution to the system mapping is helpful to improve the model’s prediction accuracy. Therefore, a reliable prediction of the sea levels can be performed using more sophisticated models by incorporating relevant information. Various research works have been carried out to make best use of useful information. Rajasekaran et al. (2006) construct functional networks and sequential learning neural networks based on history tidal observations; Pashova and Popova (2011) utilized statistical parameters of tidal level for
daily mean sea level prediction. Huang et al. (2003) proposed regional neural network for water level (RNN-WL) prediction method by taking use of tidal level data of stations distributed in regional scale; Lee (2004) employed tidal constituents for generating long-term predictions; Chang and Lin (2006) incorporated the factors related to tide-generating forces in the neural network; Lee and Jeng (2002) considered the residual fluctuation in the network's input layer for prediction purposes; Altunkaynak (2012) employed current wind speed and the previous significant wave height as inputs to give predictions of current significant wave height; Liang et al. (2008) incorporated wind information in the neural tidal prediction model, such as wind speed, wind direction, and day-averaged wind speed. Herman et al. (2007) utilized the principle components of tide, together with the wind components and historical tidal data, for predictions of tidal level and currents. Sufficient information is more crucial for predictions of abnormal tidal changes under violent meteorological conditions, such as strong winds and sudden changes in atmospheric pressure (Lee, 2008). Tseng et al. (2007) used typhoon's characteristics, local meteorological conditions and typhoon surges at a considered tidal station as inputs for typhoon surges forecast; Lee (2006, 2008) incorporated wind information and atmospheric pressure in the model for storm surge prediction.

All above-mentioned factors related to tide are time-varying and many of them (especially meteorological factors) vary rapidly with time. However, most of the neural networks employed for application are static with fixed structure and parameters, which cannot represent the fast variation of tide dynamics. Therefore, there is a practical need to establish a variable neural network whose structure and parameters can both be adaptively tuned in real-time to represent the time-varying changes of tide dynamics caused by meteorological factors. Furthermore, the theory of neural network also indicates that excessive learning usually leads to the phenomenon of over-fitting which will damage the generalization capability of network. The expansion of network dimension will also cause the “curse of dimension” and slower the computational speed consequently (Haykin, 1999). Therefore, there should be a mechanism to adjust the dimension of network.

In this paper, a hybrid model is established to for real-time tidal prediction. The harmonic analysis method is employed to give harmonic tidal level predictions; the residual influenced by meteorological factors is predicted by a variable neural network. The final prediction is the summation of prediction results of the two approaches. The efficiency of hybrid model is verified by tidal prediction simulations which are conducted based on the actual measurement of tidal stations at the port of Tampa and other three tidal stations in America.

2. Hybrid tidal prediction mechanism

2.1. Harmonic analysis method

The most conventionally used method for tidal level prediction is harmonic analysis. The principal long period cyclicities of tide are related to the astronomical factors like relative positions of the sun, moon, and earth. In harmonic analysis, the astronomical component is expressed as superposition of many sinusoidal constituents with amplitudes and frequencies determined by a local analysis of the measured tide waves. Thus, daily sea level $y(t)$ for certain time $t$ can be represented by a time-dependent function:

$$y(t) = a_0 + \sum_{i=1}^{n} h_i \cos (w_i t - \phi_i) + r(t)$$  (1)

where $a_0$ is a long-term mean sea level; $n$ is the number of constituents; $h_i$, $a_0$, and $\phi_i$ are the amplitude, temporal frequency (speed) and phase offset of the $i$th constituent, respectively. The residual $r(t)$ is the unmodelled random fluctuations representing the effects of meteorological factors, random noises and other unmodelled effects.

2.2. Variable RBF neural network

As the meteorological factors (such as atmospheric pressure $P$, wind force $F$, wind direction $\theta$, water temperature $T$) vary fast with time, their influences should be represented by variable model whose structure and parameters can be both adjustable. As meteorological and other unmodelled factors' effects on tidal level, the residual term $r(t)$ in Eq. (1) can be expressed as the functions of such factors:

$$r(t) = f(P, F, \theta, T, r) + \epsilon(t)$$  (2)

where $f(\cdot)$ is the mapping from meteorological factors to the residual, and $\epsilon(t)$ is the noises. Static neural network with fixed structure cannot represent time-varying dynamics efficiently, so there is a practical need of variable network whose structure and parameters are both adjusted in real time.

Construction of variable-structure neural network by sequential learning is a research focus in recent years (Suresh et al., 2010). Sequential learning algorithm is a popular method for representing time-varying system dynamics by variable neural network (Platt, 1991; Liang et al., 2006). It processes samples sequentially, and tunes the network's structure and parameters adaptively. By adjusting the network dimension, it can also avoid the phenomenon of over-fitting and under-fitting, thus guarantee generalization capability as well as accelerate processing speed. Among various types of network, radial basis function network (RBFN) is the most conventionally used network in sequential learning attributing its merits such as locally response character, simple topology, fast convergence speed and no phenomenon of local minima (Haykin, 1999).

In this study, to represent the time-varying dynamics of tide caused by various meteorological factors, the $f(\cdot)$ is modeled by a variable neural network constructed by a kind of sequential learning algorithm, which is described as follows:

In the algorithm, the adjustment of network structure includes network growing and network pruning. In the growing process, the newest sample is added to a new hidden neuron directly; in the pruning process, each neuron's contribution is evaluated individually and neurons which consecutively contribute less to the output will be pruned from the network. Once the neurons are determined, the connecting weight between the hidden layer and output layer is adjusted accordingly.

In a sequential learning scheme, the samples are presented sequentially but, if algorithm only learns the latest received single sample at one step, the resulted network would rely too much on the particular sample and may result in instability of network; while if the algorithm processes all the received samples at one time, the resulted network could not reflect the current changes in system dynamics, and the computational burden would also be increased. To make a compromise, a sliding data window (SDW) is employed and algorithm learns samples in the window (Yin et al., 2012). The SDW is a first-in-first-out (FIFO) sequence:

$$W_{SD} = [x_t, y_t, x_{t-1}, y_{t-1}, \ldots, x_{t-N+1}, y_{t-N+1}]$$  (3)

where $N$ denotes the width of sliding window, $x$ and $y$ denote the input vector and output vector at the corresponding time ($t, t-1, \ldots, t-N+1$), respectively. That is, the window is a combination of input matrix $X = [x_0, x_1, \ldots, x_{N-1}] \in R^{m \times N}$ and the output matrix $Y = [y_0, y_1, \ldots, y_{N-1}]^T \in R^{m \times N}$, with $n$ and $m$ being the dimensions of input and output, respectively. Therefore, the window is updated in real-time and can be employed to represent current dynamics.
of time-varying system. In this study, single hidden layer is adopted in RBFN and its jth output can be expressed as:

\[ y_i = \sum_{k=1}^{M} \theta_k \phi_k(||x - c_k||) = \sum_{k=1}^{M} \theta_k \exp \left( \frac{(||x - c_k||)^2}{2\sigma_k^2} \right), \]  

(4)

where \( \phi_k \) is the active function of the kth hidden neuron which is selected as Gaussian function in this study; \( c_k \) is the center of the kth hidden neuron; \( || \cdot || \) denotes the Euclidean distance operator; \( \sigma \) is the scale parameter of kth hidden neuron; \( j \) is the index of outputs; \( M \) is the current number of hidden neurons; \( \theta \) is the connection weight between the kth hidden neuron and the jth output. The total input–output mapping of RBFN can be represented in matrix form:

\[ Y = \Phi \Theta. \]  

(5)

To evaluate the contribution of each hidden neuron to the output separately, the response matrix is orthogonalized using Gram–Schmidt method (Chen et al., 1991):

\[ \Phi = WA \]  

(6)

where \( W \) is the orthogonalized response matrix whose individuals are orthogonal mutually; \( A \) is an upper triangle matrix. Then, the contribution of the kth hidden neuron to the output is evaluated according to its corresponding error reduction ratio (err):

\[ \text{err}_k = \frac{(w_k^T y)^2}{(w_k^T w_k) \text{trace}(Y^T Y)}. \]  

(7)

where \( w_k \) is the kth vector of the orthogonalized response matrix \( W \). The value of err is the projection of each hidden neuron in the space spanned by \( W \), which is used to evaluate the contribution of each hidden neurons to the output. The larger value of err implies larger contribution.

For multi-in multi-output (MIMO) process (Chen et al., 1992), the values of err are calculated by:

\[ \text{err}_{kj} = \frac{\sum_{h=1}^{m} (w_{kj}^T y)^2}{(w_{kj}^T w_{kj}) \text{trace}(Y^T Y)}. \]  

(8)

where \( j \) denotes the index of outputs.

However, it is noticed that the number of hidden neurons \( M \) usually is not the same as the width of sliding window \( N \), thus the matrices of \( \Phi \in \mathbb{R}^{N \times M} \) and \( W \in \mathbb{R}^{N \times N} \) are not square. Therefore, the sum of \( \text{err} \) \( \sum_{k=1}^{M} \text{err}_k \neq 1 \) and it is hard to evaluate the contribution of each hidden neuron by means of err. To apply the index of err in the SDW-based sequential learning scheme, the index of normalized error reduction ratio (nerr) is employed in this study to measure the contribution of each hidden neuron to the output (Yin et al., 2012):

\[ \text{nerr}_k = \frac{\text{err}_k}{\sum_{k=1}^{M} \text{err}_k}. \]  

(9)

It is known that \( \sum_{k=1}^{M} \text{nerr}_k = 1 \) thus the index of nerr can be used to evaluate the contribution of each hidden neuron directly. At each step, the nerr of each hidden neuron is calculated and the neurons with small nerr are selected by ranking them in ascending order. The selection is terminated once the summation of nerr reaches the preset accuracy threshold \( \rho \):

\[ \sum_{i=1}^{p+1} \text{nerr} \geq \rho. \]  

(10)

where \( p \) is the number of selected hidden neurons. It is concluded that the summed contribution of the selected \( p \) neurons is less than \( \rho \times 100\% \) to overall output. If certain neurons are selected for \( l \) consecutive steps, they will be regarded as obsolete and pruned from the network. This procedure is taken to assure that only the consecutively inactive neurons are removed from network, thus the stability of RBFN can be maintained. At each step, after the hidden neurons are adjusted by adding and pruning, connecting weights \( \Theta \) can be updated using the pseudo-inverse method:

\[ \Theta = \Phi^+ Y = (\Phi^T \Phi)^{-1} \Phi^T Y. \]  

(11)

As the elements in the input vector may have different features, it is reasonable to set different scale parameters in basis functions for the related elements. The generalized multi-scale RBFN provides more flexible representation capability of system dynamics and exhibits better generalization properties (Billings et al., 2007). In this application, variable RBFN is improved by adopting multiscale RBFs in the network.

Assume that there are \( h \) different categories of inputs and each contains \( d_k \) \( (k=1, 2, \ldots, h) \) terms in the prediction model, so we have \( \sum_{k=1}^{h} d_k = n \), and

\[ \phi_i = \exp \left\{ -\frac{1}{2} \left[ \frac{(||x_{i} - c_{i1}1)||^2}{\sigma_{i1}} + \cdots + \left( \frac{(||x_{i} - c_{ih})||^2}{\sigma_{ih}} \right)^2 \right] \right\}. \]  

(12)

where \( \phi_i \) is the response of ith center, and \( \sigma_{i1}, \sigma_{i2}, \ldots, \sigma_{ih} \) are center widths for each category of input, respectively.

Therefore, the vectors of centers and input of sliding window are divided into \( h \) categories. The implementation of multi-scale RBFN improves its fitness of input–output mapping hyperplane, thus reduces the dimension of hidden neurons in applications.

The structure and parameters are both determined sequentially by learning the data in real-time updated SDW. The resulted network can be consequently implemented for identification and prediction purposes. In our study, identification and prediction are accomplished in each step: structure and parameters are tuned upon the learning of SDW in the learning process, and prediction is conducted once the network is determined.

It is current system characteristic rather than holistic system characteristics needed to be modeled, so the algorithm needs only three parameters, which facilitates its practical applications.

2.3. Network structure of variable RBFN

Harmonic method gives predictions based on the long-term observations where the effects of meteorological factors have been averaged. Meteorological factors’ diverge from the average condition may consequently cause differences between the predicted and actual sea level. The fluctuations resulted from meteorological factors are usually small but, fluctuations may become obvious when they were influenced by extreme weather conditions such as strong winds or severe atmospheric pressure fluctuations. The water level responds to the changing atmospheric pressure and leads to the “inverse barometer effect”. That is, low pressure tends to raise sea level and high pressure tends to depress it. The effect of wind on tide is considerably variable and depends largely on the topography of the area in question. In general, it can be said that the wind will raise sea level in the direction towards which it is blowing. A strong wind blowing straight onshore will pile up the water and cause high waters to be higher than predicted, while winds blowing off the land will have the reverse effect (Admiralty Tide Tables, 2013). Ocean tidal change is also associated with water temperature; as water warms up, ocean volume increases and leads to the thermospheric sea level rise.

Therefore, all above-mentioned meteorological factors (atmospheric pressure, water temperature, wind speed and wind direction) are considered in the prediction model. Time series of residual \( r(t) \) contains much useful information concerning meteorological factors
and other unmodeled factors, thus implemented in the prediction model realized by variable RBFN.

The history residual time series \( (r(t-1), r(t-2), \ldots, r(t-p)) \) is employed to support the prediction to \( r(t) \), with \( p \) is the time lag. To decide the optimal value of \( p \), correlation values of time series with different time lag \( r(t-1), r(t-2), \ldots, r(t-p) \) is calculated with respect to \( r \) and the results are depicted in Fig. 1.

It can be seen in Fig. 1 that the correlation value decreases with the increase of enlargement of time lag \( p \). It can also be noted that there is a sharp decline in the curve with regard to time lag of 25. Therefore, the value of \( p \) is selected as 25, thus 25 h of residual time series are adopted as inputs of neural network. It is also known that a lunar day is 24 h and 50 min, which means a whole period of tidal changes is included in the network input layer. The correlation value of meteorological time series (atmospheric pressure \( P \), wind force \( F \), wind direction \( \varphi \), water temperature \( T \)) with different time lags is also calculated to \( r(t) \) and the results show that the values vary slightly with the enlargement of the time lag. Hence the samples of meteorological time series at time of \( t-1 \) \( (P(t-1), F(t-1), \varphi(t-1), T(t-1)) \) are selected to represent their mapping to \( r(t) \). This can be represented by the nonlinear autoregressive with exogenous input (NARX) prediction model (Pisoni et al., 2009). For 1-step-ahead prediction, the identification model is:

\[
r(t) = f(r(t-1), r(t-2), \ldots, r(t-p), P(t-1), F(t-1), \varphi(t-1), T(t-1)) + \epsilon(t)
\]

Therefore, prediction of \( r(t+1) \) can be achieved in prediction process by using the samples which is available currently:

\[
r(t+1) = f(r(t), r(t-1), \ldots, r(t-p+1), P(t), F(t), \varphi(t), T(t))
\]

For \( q \)-step ahead prediction, the NARX model during identification is:

\[
r(t) = f(r(t-q), r(t-q-1), \ldots, r(t-q-p+1), P(t-q), F(t-q), \varphi(t-q), T(t-q)) + \epsilon(t)
\]

NARX model for \( q \)-step ahead prediction becomes:

\[
r(t+q) = f(r(t), r(t-1), \ldots, r(t-p+1), P(t), F(t), \varphi(t), T(t))
\]

As a result, the topology structure of the variable RBF neural network model is \( I_29H_2O_1 \), where \( I_29 \) denotes 29 input neurons, \( O_1 \) denotes single output neuron of actual tidal level, and \( H_2 \) denotes the \( x \) hidden neurons, with the value of \( x \) being on-line determined by the sequential learning algorithm.

2.4. Hybrid tidal prediction method

The hybrid modeling method in this study is the combination of harmonic analysis and ANN which aims at bringing different advantages of first principle knowledge and data driven techniques into one model (Oliveira, 2004). The harmonic analysis takes consideration of astronomical factors thus can give stable predictions, but it is based on long-term average historical measurement and cannot represent the complex influence of time-varying meteorological factors. On the contrary, the nonlinearity and adaptivity natures of neural network enable its online tuning ability in representing time-varying nonlinear system dynamics. However, the representing capability of neural network depends on the learning of data, and cannot include any physical representation of object process directly (von Stosch et al., 2012), thus will result in limitation in extrapolation performance.

It has also been discovered that the combination of first principles knowledge and ANNs can better structure the process operation space (Fiedler and Schuppert, 2008), thus generate better calibration properties, better extrapolation capabilities and enhanced statistical confidence of the model estimates comparing to pure ANNs or other data-driven techniques (von Stosch et al., 2011; Omer Faruk, 2010).

Therefore, in this study, a hybrid model is employed for real-time tidal predictions to enhance the prediction performance in respect of robustness and efficiency. The influences of astronomical and meteorological factors are represented by harmonic method and variable RBFN, respectively.

The learning process is processed in two steps. In the identification process, the harmonic method is firstly implemented to represent the astronomical effects to tide, and the tidal level acquired by harmonic analysis is denoted as \( y_i(t) \). Identification residual \( r(t) \), the differences between actual level \( y(t) \) and \( y_i(t) \), is then identified by the variable RBFN which is employed to represent the meteorological factors’ effects to tidal levels. That is, the variable RBFN is tuned to represent the input–output mapping in Eq. (15). The configuration of the learning process in the hybrid method is illustrated in Fig. 2.

In the prediction process, the predictions of the harmonic method and the achieved RBFN, \( y_i(t+q) \) and \( r(t+q) \) are generated according to Eqs. (1) and (15), respectively. The overall output is then achieved by summing the prediction results of the two modules, \( y_f(t) \) and \( r(t) \). The configuration of the hybrid method during prediction process is illustrated in Fig. 3.

In the sequential processing mode, the prediction is conducted once the identification is completed and both are conducted at each step consecutively.

3. Tidal level prediction using the hybrid prediction method

3.1. Prediction model

Measurement data of the tidal level at Old Port Tampa is employed to validate the efficiency of the hybrid prediction...
method. The Old Port Tampa tidal station is situated on the western end of the Interbay peninsula in southern America. Geographic position of the tidal gauge is 27°51.4' N, 82°33.1' W and its location is depicted in Fig. 4.

The experimental measurements belong to hourly value of the Tampa tidal station over the duration of one year (from 1 July 00:00 GMT, 2012 to 30 June 23:00 GMT, 2013) with time resolution of 1 h, totally 8760 pairs of samples with respect to the tidal level, atmospheric pressure, water temperature, wind force and wind direction, respectively. All the measurement data employed in this study is available from the web site of the American National Oceanic and Atmospheric Administration (NOAA): http://co-ops.nos.noaa.gov/.

The harmonic predictions of tidal level at Old Port Tampa, which are also presented on the web site of NOAA, are shown in Fig. 5 along with the actual measured ones.

It can be seen from Fig. 5 that the predicted tidal levels can reflect the holistic changing law, which shows that the harmonic method is effective in representing the tidal changes caused by astronomical factors. However, it can also be noted that the predicted tidal levels differs from the measurements most of the time, and there exists a relatively large sustained deviation between the measured and predicted tidal levels with correlation coefficient (CC) of 0.8643, as shown in Fig. 6.

The difference between estimated and actual measured sea levels is shown in Fig. 7. It can be noted in Fig. 7 that sustained deviation exists for certain period of time which is caused by environmental changes. It also can be noted that the range of prediction error may reach −0.7390 m at 1376-th hour (0700 Aug. 27, 2012) and 0.8610 m at 8183-th hour (2200 June 6, 2013), which...
may cause danger in practical marine operations based on these prediction results.

The harmonic prediction result's deviation from actual measurement is mainly caused by time-varying environmental changes such as the atmospheric pressure and strong wind. The harmonic method can reflect the first principle of influence of celestial bodies, and the variable RBFN can represent time-varying system dynamics. To take both considerations of astronomical and meteorological factors, the proposed hybrid prediction method is implemented for real-time prediction of tidal level of Old Port Tampa, with the influences of astronomical and meteorological factors are represented by approaches of harmonic method and variable structure RBFN, respectively. The samples of the elements concerned in the prediction model are illustrated in Fig. 8.

3.2. Performance of tidal level prediction

The effectiveness of the proposed hybrid prediction method is validated using the measured tidal level measurement of Old Port Tampa. All the simulations in this study were carried out in MATLAB 7.4 environment running at 2.80 GHz (CPU) and 3.25 GB memory (RAM).

The identification and prediction performances of model are evaluated by indices of root mean square error (RMSE), with identification RMSE (RMSE_{iden}) and prediction RMSE (RMSE_{pre}), respectively:

\[
RMSE_{iden} = \sqrt{\frac{\sum_{t=1}^{n} (y_{iden}(t) - y(t))^2}{n}}
\]

\[
RMSE_{pre} = \sqrt{\frac{\sum_{t=1}^{n} (y_{pre}(t+q) - y(t+q))^2}{n}}
\]

where \(y_{iden}(t)\) and \(y(t)\) are the identified and actual value of tidal level at time \(t\); \(y_{pre}(t+q)\) and \(y(t+q)\) are the predicted and actual values of tidal level at time \(t+q\), respectively. These two indices are used for validating the model identification and prediction accuracies (Ömer Faruk, 2010).

The parameters for constructing variable RBFN are selected as: \(N=72, \rho=0.001, l=3\). The width for the elements of residual, atmospheric pressure, wind force, wind direction, water temperature are selected as 1.44, 1028.5, 15.4, 109 and 39.1, respectively. For harmonic prediction method, 37 constituents are selected and corresponding parameters (amplitude, phase, speed) for each constituent are also available at http://co-ops.nos.noaa.gov/.

At each step, the model learns the samples in SDW and adjusts the structure and parameters of variable RBFN. Based on currently available tidal and metrological measurements, tidal prediction is made by combining the predictions generated by Harmonic method and variable RBFN, respectively. The 1-h-ahead tidal level prediction results are depicted in Fig. 9 together with the actual measurements.

It is shown in Fig. 9 that the predicted results track the tendencies of tidal rise and fall well with RMSE_{pre} of 0.0262688 m, comparing to 0.1528718 m using conventional harmonic method. The abrupt changes of tidal levels at time of 1376-th and 8183-th hours are also detected and predicted accordingly with small errors, which is shown in Fig. 10(a) and (b), respectively.

This shows that the hybrid model can incorporate the influences of meteorological factors to tidal changes and possesses better prediction accuracy than custom harmonic method. The prediction error is depicted in Fig. 11.

It can be noticed in Fig. 11 that there are no large sustained deviations between the predicted values and actual ones, and the predicted values locate around the measured ones uniformly, which can also be observed in Fig. 12. The closeness between the measured tidal levels and the predicted ones can be further verified by the corresponding high CC of 0.9949, as shown in Fig. 12.

The predictions are made based on the identification of data in the SDW. The RMSE_{iden} is 0.0146015 m, which is close to the RMSE_{pre} of 0.0262688 m.

The RBFN structure is variable with hidden neurons are added or pruned adaptively according their corresponding index nerr. The number of hidden neurons is a variable value ranged from 5 to 32, and the average hidden neurons number (HNN) during simulation is 16.6591. The evolution history of number of HNN is depicted in Fig. 13. In this study, sequential learning RBFN in this study employs small number of hidden neurons, which also enable its fast processing speed of the hybrid prediction model. The overall consumed time for 8700 steps of processing is 144.260448 s, the average time for each step (including identification and prediction process) is 0.016582 s.

Different time span of prediction influences the magnitude of prediction error. To validate the prediction accuracy for longer time scope, simulations are conducted based on the same measurement data. For 1-h-ahead and 2-h-ahead prediction, the parameters of model are selected as: \(N=72, \rho=0.001, l=3\). For 3 and more hours ahead prediction, the parameters are adjusted as \(N=24, \rho=0.01, l=3\). Same harmonic parameters are adopted as the harmonic prediction method. Simulation results are shown in Table 1.

It is noticed in Table 1 that the RMSEs of prediction are 0.026269 and 0.0377226 for 1-step-ahead and 2-steps-ahead predictions, respectively; values of RMSE remain among ranges of 0.044520 and 0.046571 for 3 and more steps ahead predictions. This shows that the proposed hybrid prediction method can give stable predictions for short-term tidal predictions with satisfactory prediction accuracy. It can also infer from the results that the variable RBFN employs only a small number of hidden neurons for representing the time-varying influences of meteorological factors, and the processing speed is rather fast consequently.

For the comparison of tidal prediction performance using other approach, conventional NARX method using least square algorithm is
implemented under the same experimental conditions. The models (15) and (16) are used for identification and prediction, respectively. The simulation results are shown in Table 2.

It is noticed by comparing prediction results illustrated in Tables 1 and 2 that, for 1-h-ahead prediction, the prediction accuracy using conventional NARX model is slightly higher than that with proposed modular prediction method. However, with the increase of prediction time span, the proposed method shows higher accuracy both in identification and prediction. This shows the nonlinear neural network is superior to linear least squares algorithm in approximating nonlinear mapping. The proposed method also possesses higher processing speed attributing to that it only needs to process the samples in SDW.

Measurements at other three American tidal stations, which differ in geographic conditions such as shelter and wave reflection, are employed to evaluate the universality of the proposed prediction method. The stations are Monterey (36°36.3’N, 121°53.2’W), Honolulu (21°18.4’N, 157°52.0’W) and Chesapeake Bay Bridge Tunnel (36°58’N, 76°6.8’W).

The time span of hourly tidal measurement for the three stations is 1 Jan. 00:00 GMT, 2013 to 31 Nov. 23:00 GMT, 2013.

![Fig. 8. Samples of measured data for identification and prediction.](image)

![Fig. 9. Measured data and predicted results by hybrid method.](image)

![Fig. 10. Measured data and predicted results by hybrid method ((a) 1000–1500 h; (b) 8000–8500 h).](image)

![Fig. 11. Prediction error by hybrid method.](image)

![Fig. 12. Scatter diagram of the measured and predicted results by hybrid method.](image)
Same parameters are employed as above experiment. That is, \( N = 72, \rho = 0.001 \) and \( l = 3 \) for 1-h-ahead prediction; \( N = 24, \rho = 0.01 \) and \( l = 3 \) for 6- and 24-h-ahead prediction. Simulation results of tidal prediction are shown in Table 3.

It is shown in Table 3 that the proposed method give stable and accurate tidal predictions for measurement data of three different tidal stations. And the tidal stations of Old Port Tampa, Monterey, Chesapeake Bay Bridge Tunnel and Honolulu are located on the southern, western, eastern coast of America and the middle of northern Pacific, respectively, which shows the generality of the method. It is also noted that same parameters are employed for different tidal stations in this study. Further fine tuning of parameters should improve the performance of the hybrid method.

### 4. Conclusion

A hybrid prediction model is proposed by taking both advantages of harmonic method and variable RBF neural network. The harmonic analysis and variable neural network represent the influences of celestial bodies and meteorological factors, respectively. Simulation results of real-time tidal level prediction demonstrate that the proposed method generates accurate predictions with high processing speed, which demonstrate the feasibility and effectiveness of the hybrid tidal prediction method. To adapt to the diversities of the tidal changes and meteorological changes, the self-adaptation of system parameters will be one of our future research efforts.

### Acknowledgements

This work is supported by the National Natural Science Foundation of China (Grant no. 51061130548), the Applied Basic Research Fund of

---

**Table 1**

| Prediction time span | RMSE_{prop} (m) | CC     | RMSE_{iden} (m) | Average HNN | Steps | Time (s) |
|----------------------|------------------|--------|-----------------|-------------|-------|----------|
| 1-h-ahead            | 0.026269         | 0.994875 | 0.014602        | 16.659080   | 8700  | 144.260448 |
| 2-h-ahead            | 0.037226         | 0.989708 | 0.018275        | 19.189310   | 8700  | 150.072742 |
| 3-h-ahead            | 0.043920         | 0.983752 | 0.014413        | 8.010804    | 8700  | 76.118526  |
| 6-h-ahead            | 0.044520         | 0.985469 | 0.014091        | 7.986896    | 8700  | 75.269322  |
| 12-h-ahead           | 0.046571         | 0.983907 | 0.014275        | 7.947931    | 8700  | 76.468794  |
| 24-h-ahead           | 0.045237         | 0.984855 | 0.014439        | 7.960920    | 8700  | 75.604252  |

**Table 2**

| Prediction | RMSE_{prop} (m) | CC     | RMSE_{iden} (m) | Steps | Time (s) |
|------------|------------------|--------|-----------------|-------|----------|
| 1-h-ahead  | 0.026217         | 0.981230 | 0.025043        | 8700  | 213.583181 |
| 2-h-ahead  | 0.038237         | 0.967752 | 0.032962        | 8700  | 214.914666 |
| 3-h-ahead  | 0.045140         | 0.941232 | 0.043585        | 8700  | 211.382728 |
| 6-h-ahead  | 0.072568         | 0.845749 | 0.070174        | 8700  | 216.315268 |
| 12-h-ahead | 0.095987         | 0.708399 | 0.093277        | 8700  | 212.482581 |
| 24-h-ahead | 0.114384         | 0.541636 | 0.112212        | 8700  | 212.179456 |

**Table 3**

| Tidal station       | Prediction Span (h) | RMSE_{prop} (m) | CC     | RMSE_{iden} (m) | Average HNN | Steps | Time (s) |
|---------------------|---------------------|------------------|--------|-----------------|-------------|-------|----------|
| Monterey            | 1                    | 0.023826         | 0.998815 | 0.014910        | 20.268161   | 7900  | 140.473108 |
|                     | 6                    | 0.030585         | 0.998049 | 0.012452        | 9.407241    | 7900  | 66.134935  |
|                     | 24                   | 0.032806         | 0.997760 | 0.012423        | 9.219310    | 7900  | 67.756684  |
| Honolulu            | 1                    | 0.015604         | 0.996309 | 0.009866        | 20.343670   | 7900  | 128.515855 |
|                     | 6                    | 0.019586         | 0.994225 | 0.009447        | 7.765949    | 7900  | 71.065877  |
|                     | 24                   | 0.021896         | 0.993271 | 0.008109        | 9.785316    | 7900  | 73.365363  |
| Chesapeake Bay Tunnel | 1                  | 0.035915         | 0.994283 | 0.019704        | 15.492025   | 7900  | 124.627586 |
|                     | 6                    | 0.048092         | 0.989714 | 0.019737        | 6.867848    | 7900  | 67.024692  |
|                     | 24                   | 0.050218         | 0.988893 | 0.019037        | 6.876582    | 7900  | 68.403307  |
References

Admiralty Tide Tables. 2013. Pacific Ocean (Including Tidal Stream Tables), vol. 4. United Kingdom Hydrographic Office, Taunton, UK.

Altunkaynak, A., 2012. Prediction of significant wave height using geno-multilayer perceptron. Ocean Eng. 58, 144–153.

Benhammadi, F., Gessoum, Z., Molftari, A., 2011. CPU load prediction using neuro-fuzzy and Bayesian inferences. Neurocomputing 74, 1606–1616.

Billings, S., Wei, H., Balikhin, M., 2007. Generalized multiscale radial basis function networks. Neural Netw. 20, 1081–1094.

Chang, H.K., Lin, L.C., 2006. Multi-point tidal prediction using artificial neural network with tide-generating forces. Coastal Eng. 53, 857–864.

Chen, S., Cowan, C., Grant, P., 1991. Orthogonal least squares learning algorithm for radial basis function networks. IEEE Trans. Neural Netw. 2, 302–309.

Chen, S., Grant, P., Cowan, C., 1992. Orthogonal least-squares algorithm for training multioutput radial basis function networks. IEE Proc.: F 139, 378–384.

Chen, S., Hong, X., Harris, C.J., 2011. Grey-box radial basis function modeling. Neurocomputing 74, 1564–1571.

Fang, G.H., Zheng, W.Z., Chen, Z.Y., 1986. Analysis and Prediction of Tide and Tidal Current. Ocean Press, Beijing, China.

Fiedler, B., Schuppert, A., 2008. Local identification of scalar hybrid models with tree structure. IMA J. Appl. Math. 73, 449–476.

Ghorbani, M.A., Khatibi, R., Aytek, A., Makarynskyy, O., Shiri, J., 2010. Sea water level forecast using genetic programming and comparing the performance with artificial neural networks. Comput. Geosci. 36, 620–627.

Guaydun, K., 2008. The estimation of monthly mean significant wave heights by using artificial neural network and regression methods. Ocean Eng. 35, 1406–1415.

Haykin, S., 1999. Neural Networks: A Comprehensive Foundation. Prentice Hall, New Jersey.

Herman, A., Kaiser, R., Niemeyer, H.D., 2007. Modelling of a medium-term dynamics in a shallow tidal sea, based on combined physical and neural network methods. Ocean Modell. 17, 277–299.

Huang, W., Murray, C., Kraus, N., Rosati, J., 2003. Development of a regional neural network for coastal water level predictions. Ocean Eng. 30, 2275–2295.

Liang, N.Y., Huang, G.B., Saratchandran, P., Sundararajan, N., 2006. A fast and accurate online sequential learning algorithm for feedforward networks. IEEE Trans. Neural Netw. 17, 1411–1422.

Liang, S.X., Li, M.C., Sun, Z.C., 2008. Prediction models for tidal level including strong meteorologic effects using a neural network. Ocean Eng. 35, 666–675.

Lee, T.L., Jeng, D.S., 2002. Application of artificial neural networks in tide forecasting. Ocean Eng. 29, 1003–1022.

Lee, T.L., 2004. Back-propagation neural network for long-term tidal predictions. Ocean Eng. 31, 225–238.

Lee, T.L., 2006. Neural network prediction of a storm surge. Ocean Eng. 33, 483–494.

Lee, T.L., 2008. Back-propagation neural network for the prediction of the short-term storm surge in Taichung harbor, Taiwan. Eng. Appl. Artif. Intell. 21, 63–72.

Oliveira, R., 2004. Combining first principles modeling and artificial neural networks: a general framework. Comput. Chem. Eng. 28, 755–766.

Oner Faruk, D., 2010. A hybrid neural network and ARTIMA model for water quality time series prediction. Eng. Appl. Artif. Intell. 23, 586–594.

Pashova, L., Popova, S., 2011. Daily sea level forecast at tide gauge Burgas, Bulgaria using artificial neural networks. J. Sea Res. 66, 154–161.

Pisoni, E., Farina, M., Carnevale, C., Piriddoli, L., 2009. Forecasting peak air pollution levels using NARX models. Eng. Appl. Artif. Intell. 22, 593–602.

Platt, J., 1991. A resource allocating network for function interpolation. Neural Comput. 3, 213–225.

Pugh, D., 1987. Tides, Surges and Mean Sea-Level John Wiley & Sons, Swindon, UK.

Rajasekaran, S., Thiruvenkatasamy, K., Lee, T., 2006. Tidal level forecasting using functional and sequential learning neural networks. Appl. Math. Modell. 30, 85–103.

Suresh, S., Dong, K., Kim, H.J., 2010. A sequential learning algorithm for self-adaptive resource allocation network classifier. Neurocomputing 7, 3012–3019.

Tseng, C.M., Jan, C.D., Wang, J.S., Wang, C.M., 2007. Application of artificial neural networks in typhoon surge forecasting. Ocean Eng. 34, 1757–1768.

von Stosch, M., Oliveira, R., Peres, J., Feyo de Azevedo, S., 2011. A hybrid modeling framework for process analytical technology: application to Bordetella pertussis cultures. J. Biotechnol. Prog. 28, 284–291.

von Stosch, M., Oliveira, R., Peres, J., Feyo de Azevedo, S., 2012. A general hybrid semi-parametric process control framework. J. Process Control 22, 1171–1181.

Yin, J.C., Wang, L.D., Wang, N.N., 2012. A variable-structure gradient RBF network with its application to predictive ship motion control. Asian J. Control 14, 716–725.