Acta Oceanol. Sin., 2017, Vol. 36, No. 11, P. 94–105
DOI: 10.1007/s13131-017-1140-x
http://www.hyxb.org.cn
E-mail: hyxbe@263.net

A precise tidal prediction mechanism based on the combination of harmonic analysis and adaptive network-based fuzzy inference system model

ZHANG Zeguo[1], YIN Jianchuan[1]*, WANG Nini[2], HU Jiangqiang[1], WANG Ning[3]
[1] Navigation College, Dalian Maritime University, Dalian 116026, China
[2] Department of Mathematics, Dalian Maritime University, Dalian 116026, China
[3] Marine Engineering College, Dalian Maritime University, Dalian 116026, China

Received 25 February 2017; accepted 21 March 2017

Abstract

An efficient and accurate prediction of a precise tidal level in estuaries and coastal areas is indispensable for the management and decision-making of human activity in the field wok of marine engineering. The variation of the tidal level is a time-varying process. The time-varying factors including interference from the external environment that cause the change of tides are fairly complicated. Furthermore, tidal variations are affected not only by periodic movement of celestial bodies but also by time-varying interference from the external environment. Consequently, for the efficient and precise tidal level prediction, a neuro-fuzzy hybrid technology based on the combination of harmonic analysis and adaptive network-based fuzzy inference system (ANFIS) model is utilized to construct a precise tidal level prediction system, which takes both advantages of the harmonic analysis method and the ANFIS network. The proposed prediction model is composed of two modules: the astronomical tide module caused by celestial bodies' movement and the non-astronomical tide module caused by various meteorological and other environmental factors. To generate a fuzzy inference system (FIS) structure, three approaches which include grid partition (GP), fuzzy c-means (FCM) and sub-clustering (SC) are used in the ANFIS network constructing process. Furthermore, to obtain the optimal ANFIS based prediction model, large numbers of simulation experiments are implemented for each FIS generating approach. In this tidal prediction study, the optimal ANFIS model is used to predict the non-astronomical tide module, while the conventional harmonic analysis model is used to predict the astronomical tide module. The final prediction result is performed by combining the estimation outputs of the harmonious analysis model and the optimal ANFIS model. To demonstrate the applicability and capability of the proposed novel prediction model, measured tidal level samples of Fort Pulaski tidal station are selected as the testing database. Simulation and experimental results confirm that the proposed prediction approach can achieve precise predictions for the tidal level with high accuracy, satisfactory convergence and stability.

Key words: tidal level prediction, harmonious analysis method, adaptive network-based fuzzy inference system, correlation analysis

Citation: Zhang Zeguo, Yin Jianchuan, Wang Nini, Hu Jiangqiang, Wang Ning. 2017. A precise tidal prediction mechanism based on the combination of harmonic analysis and adaptive network-based fuzzy inference system model. Acta Oceanologica Sinica, 36(11): 94–105, doi: 10.1007/s13131-017-1140-x

1 Introduction

The precise tidal level prediction is a significant activity for the design of coastal constructions such as wharves and harbors. A precise tidal prediction also has a great influence on decision-making procedures of vessels or drilling platforms in offshore areas (Fang et al., 1986). An accurate tidal prediction is particularly crucial in the fields of operation scheduling, such as making navigation plans of ships through shallow water or bridge. Under these conditions, decision consideration of the under-keel clearance or the water depth under bridge can be adopted based on the hourly tidal level forecast. Furthermore, tidal information is a significant affecting factor for navigation scheduling of platform's operation planning. Consequently, the accurate forecasting of tidal level in the fields of coastal and ocean engineering is an issue of concern (Xiao et al., 2016).

Nowadays, the most popular mechanism in tidal research is the conventional harmonic analysis method, in which the components of tidal can be expressed as the superposition of several sinusoidal constituents determined by the harmonic analysis approach. In addition, the conventional harmonic analysis method remains the basis for long-term tidal prediction (Lee and Jeng, 2002; Lee, 2006). Nevertheless, there are some deficiencies in practical applications of the harmonic analysis method. First, the essential constituents of the harmonic analysis method need to be determined by a considerably long period of tidal level records. Nevertheless, the long-term tidal level records may not be

---

Foundation item: The National Natural Science Foundation of China under contract No. 51379002; the Fundamental Research Funds for the Central Universities of China under contract Nos 3132016322 and 3132016314; the Applied Basic Research Project Fund of the Chinese Ministry of Transport of China under contract No. 2014329225010.

*Corresponding author, E-mail: yinjianchuan@dlmu.edu.cn obtained owing to the high cost of the *in situ* monitoring data; in addition, the only consideration of this forecasting model is the influence of celestial bodies and coastal topography like the framework of the coastline and the profile of the seabed. Consequently, it cannot convey the intricate time-varying meteorological impacts on tide, and those impacts are generated by meteorological elements like sea rainfall, sea wind, sea ice, atmospheric temperature and pressure and so on. Furthermore, there are still some other time-varying factors such as river discharge and fluctuation which will also have great influence on the tidal level in an estuarine area. Consequently, the forecasting precision of the conventional harmonic analysis method is quite low under some circumstances.

In recent years, with the rapid development of artificial intelligence technology such as fuzzy logic and neural network, the artificial intelligence technology has been widely applied into the field of engineering computation and simulation (Yin et al., 2014; Yin and Wang, 2013) due to its strong nonlinear mapping and self-learning ability. Artificial neural networks (ANNs) have been proved to be of versatile utility in engineering optimization and other scientific computation fields (Haykin, 1999), owning to their excellent abilities of nonlinear mapping, generalization and self-learning. As an approach of data analysis and processing, the artificial neural network can express complicated non-periodic and nonlinear relationships based on the study of the training data set, which possesses superior predictive performance to the conventional harmonic analysis approach. As a consequence, it could be suitable for the precise tidal level predictions. The neural network based on error back propagation (BP) is the most popular and practical neural network, and it is widely utilized in tide level forecasting (Lee, 2004, 2008). An online sequential extreme learning machine based on an improved Gath-Geva fuzzy segmentation and a variable-structure radial basis function neural network constructed by sequential learning are proposed for tidal prediction (Yin and Wang, 2015; Yin et al., 2013; He et al., 2012). One day ahead tide prediction of the west coast of India New Mangalore Tide Station was carried out by using the neural networks (Jain and Deo, 2007), and the artificial neural network together with regression methods was utilized for the estimation of monthly mean significant wave heights (Günaydın, 2008), and a development of a regional neural network were proposed for coastal water level predictions (Huang et al., 2003).

Recently, due to the excellent learning and approximation capability of the adaptive network-based fuzzy inference system(ANFIS), some novel models based on the ANFIS have been proved to be a prevalent universal approximator which can express highly nonlinear functions (Yetilmezsoy et al., 2011). Actually, in the system of fuzzy inference, the experiment or knowledge of humanity together with the process of inference could be analyzed and characterized qualitatively by means of fuzzy if-then rules. However, it cannot achieve an accurate quantitative analysis. Furthermore, the artificial neural network possesses the excellent capability of self-learning, self-organizing and self-adapting. Nevertheless, they are not able to tackle qualitative knowledge and the procedure of inference. As a consequence, ANFIS model which combines the advantages of the artificial neural networks and the fuzzy inference systems was proposed by Jang (1993), Jang and Sun (1995). The ANFIS structure has been successfully utilized to approximate highly nonlinear functions, determine nonlinear constitutes in a control and prediction system (Stefanakos, 2016; Mekanik et al., 2016; Chang and Lai, 2014), and forecast a chaotic time series (Ruano, 2005). Accordingly, this proposed novel mechanism is capable of dealing with quite complicated and highly nonlinear issues (Cakmakci, 2007). Because of the influence of many uncertain factors, especially the influence of meteorological factors, the non-astronomical tide section has the characteristics of strong nonlinearity, time varying and fuzzy uncertainty. As a result, the ANFIS model which combines the advantages of the artificial neural network and the fuzzy inference systems can represent the complicated highly nonlinearity meteorological influences on tide which is generated by the meteorological factors like atmosphere, wind, pressure, rainfall, ice, etc. Based on the above analysis, the ANFIS possesses superior predictive performance to the conventional harmonic analysis approach and the conventional artificial neural network. Consequently, it could be more suitable for the accurate tidal level predictions.

The ANFIS combined with the harmonic analysis model is employed in this study with the purpose of constructing a forecasting model of the precise tidal level using the measured tidal level data base. A correlation analysis (Zhao et al., 2014; Nezlin and Li, 2003) is used to analyze the time series of the precise tidal level data for the determination of the input dimensions of the ANFIS prediction model and the hybrid prediction model which is the combination of the ANFIS prediction model and the harmonic analysis prediction model. Furthermore, grid partition (GP), fuzzy c-means (FCM) and sub-clustering (SC) approaches are exploited for the generation of fuzzy inference system (FIS). For each approach, a number of different model structures are established by adjusting the corresponding parameters. The precise tidal level forecasting performances of the established models are investigated by numerous simulation experiments in order to obtain the optimal ANFIS based forecasting model. In the hybrid prediction model, the harmonic analysis method (Clue, 2004) is employed to predict the astronomical tidal parts, while the optimal ANFIS based forecasting model is utilized to predict the non-astronomical tidal parts with strong nonlinearity. The proposed prediction method combines the advantages of the two methods: the harmonic analysis method can achieve a long-term, stable astronomical tide forecast, and the ANFIS method can complete the nonlinear fitting and prediction of tides with high accuracy and fast convergence rate. And the final prediction result is accomplished by combining the estimation outputs of the optimal ANFIS- based forecasting model and the harmonic analysis model. Additionally, the main purpose of this research is to solve the issue of precise tidal level forecasting by means of constructing a stable and reliable prediction model, and thus a useful and valuable reference could be provided to marine engineering managers for the future early prediction and management of the tidal water level. The proposed model is employed to study its suitability and applicability for the hourly prediction of the tidal water level. The measured tidal data of Fort Pulaski Tidal Station is selected as the testing data base. In this research, conventional harmonious analysis and the BP method together with the traditional AR (auto regression) approach are also employed to compare the performance with the proposed hybrid forecasting model. The results of the simulation have confirmed that the proposed novel prediction model can give predictions for the precise tidal level with high accuracy, excellent convergence and satisfactory stability. Furthermore, the experiment results acquired could be a valuable reference to the corresponding researchers in the study area of tidal prediction.

The rest of the paper is organized as follows: a brief description about the ANFIS is given in Section 2; the ANFIS modeling process is represented in Section 3 together with the comparison of three approaches which are employed to generate the FIS for obtaining the optimal ANFIS- based forecasting model; harmonic analysis method together with the hybrid prediction model which is the combination of the harmonic analysis method and the optimal ANFIS-based forecasting model is provided in Section 4; simulation results and detailed discussions are displayed in Section 5; and conclusions are shown in the last section.

2 Adaptive neuro-fuzzy inference system

ANFIS belongs to a kind of multilayer feedforward network which utilizes training and learning algorithms of the neural network together with the fuzzy inference approach to map the input space to the output space. It is composed of input-output information and a fuzzy rule data base of the Takagi-Sugeno type (Jang, 1993). For simplification, it is assumed that the structure of the ANFIS possesses two inputs variables x, y and one output variable f. Figure 1 shows the framework of the ANFIS model, which consists of five layers with two input variables (x and y). Actually, the fuzzy if-then rule consists of two primary modules: IF module and THEN module which are named as antecedent part and consequent part respectively. Then, for a first-order Sugeno fuzzy model (Takagi and Sugeno, 1985), the corresponding fuzzy rule data set with two fuzzy if-then rules can be represented as:

for Rule 1, if x is A_1 and y is B_1, then $f_1=p_1x+q_1y+r_1$;
for Rule 2, if x is A_2 and y is B_2 then $f_2=p_2x+q_2y+r_2$;
where p_i, q_i and r_i (i=1 or 2) denote linear parameters in the consequent part of the Sugeno fuzzy model; and A_i and B_i represent the linguistic labels which are expressed by fuzzy rule sets.

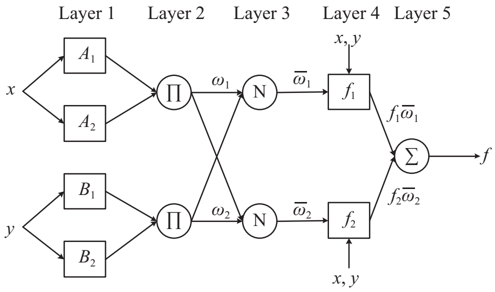

Fig. 1. The ANFIS architecture.

The basic principle of the neuro-adaptive learning mechanisms is very simple. It provides an effective approach for the fuzzy modeling process to learn knowledge from a data base, to calculate the corresponding parameters of a membership function. For the sake of calculating the parameters of consequent part (linear parameters or consequent parameters) by employing a hybrid learning algorithm, the ANFIS model maps the FIS into neural network architecture. In the hybrid algorithm, the membership function parameters (nonlinear parameters or antecedent parameters) are confirmed by utilizing the back-propagation learning algorithm. Meanwhile, the consequent parameters (linear parameters) are calculated by utilizing the least square approach.

Figure 1 expresses the framework of the ANFIS model consisting of two inputs variables x, y, and one output variable f with two fuzzy rules which are represented in the antecedent section. The first layer of the ANFIS modeling is the fuzzification layer where A_i and B_i denote linguistic labels. The output value of the fuzzification layer is membership grades of the corresponding fuzzy sets. Namely, in the first layer, the antecedent parameters are computed. The firing strength for the corresponding fuzzy rule is obtained by using the AND operator in the second layer. The third layer is a normalization layer, and the main target of which is to obtain the ratio of the ith fuzzy rule's firing strength to the sum of entire fuzzy rules' firing strength. In the fourth layer, the results from Layer 3 are multiplied with the function of Sugeno inference system. Only one node exists in Layer 5, and this single node calculates the sum of entire outputs of each fuzzy rule from Layer 4. After that, the weighted averaged approach is applied to achieve the procedure of defuzzification, which converts the fuzzy outputs into actual results.

In the forward propagation process of the ANFIS model, the outputs of premise section will be delivered to Layer 4 and the least squares estimate approach is employed for the adjustment of the consequent parameters. In the backward propagation process, the error information will be conveyed backward to Layer 1 and the gradient descent algorithm-based back-propagation approach is applied to adjust the antecedent parameters.

3 Prediction model based on the adaptive neuro fuzzy inference system

3.1 *Tidal data analysis*

The conception of correlation analysis (Young and Shellswell, 1972) originates from the signal analysis and processing, and correlation analysis reflects that how the correlation between any two values in a time series is changed over time. The autocorrelation analysis characterizes the correlation of each time series and it can also represent the correlation between neighboring variables of the time series in detail. Autocorrelation function and partial autocorrelation function are effective methodologies to analyze and tackle the issues of complicated time series: the autocorrelation function depicts the relationship of each time series, as a consequence, it can also represent the correlation between the adjacent variables of time series; while the partial correlation function eliminates the effect of other intermediate variables in time series. In the tidal data analysis, the time series of tidal level data is influenced by numerous variable factors that are hard to be determined by nautical equipment, which makes it difficult to calculate the contribution to the tidal level respectively. As a consequence, the correlation analysis approach is applied to analyze the correlation between the time series of tidal level, and then to obtain the input dimension of the ANFIS forecasting model. In this study, the correlation value of 0.4 is chosen as a reference standard to determine the input dimension and the analysis results is illustrated in Fig. 2 and Fig. 3. Figure 2 shows that the autocorrelation coefficient of tidal level time series data is tailing. Furthermore, Fig. 3 implies that the partial autocorrelation coefficient is a four-order truncation. To summarize, the correlation analysis confirms that the time series value from t-1 to t-4 moment has a noticeable relevance with the time series value of moment t, which can be chosen as the input of the ANFIS forecasting model.

Generally, fuzzy rules of ANFIS model are produced automatically. For the sake of generating FIS architecture from sampling data in the ANFIS model, three approaches including GP, FCM and SC are investigated in this research. Additionally, these approaches are employed to construct and determine the best ANFIS model for tidal level prediction model. In the simulation process, parameters of ANFIS forecasting model are determined within the stage of training. The checking data set is utilized to verify the effectiveness and the accuracy of the forecasting model.

As shown in Fig. 4, the symbol of the white star represents

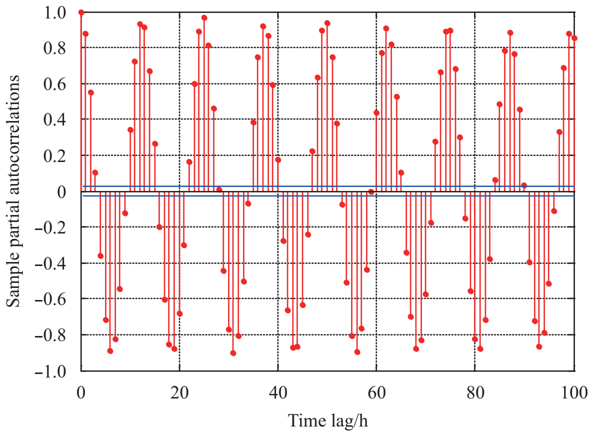

Fig. 2. Autocorrelation analysis of tidal data.

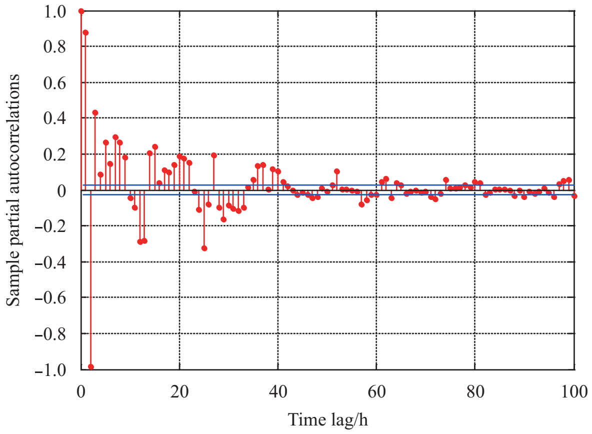

Fig. 3. Partial autocorrelation analysis of tidal data.

Fort Pulaski Tidal Station which is selected as the testing point, and the other five white diamond symbols represent the other five tidal stations chosen for the validation of the proposed tidal level forecasting model. The experimental measurements of Fort Pulaski Tidal Station over the duration from 1 January 00:00 GMT 2015 to 7 September 23:00 GMT 2015 with time resolution of 1 h, totally 6 000 pairs of observed tidal level data which includes 70% of the training data, 30% of the checking data, are chosen as the checking and training data base for the selection of the optimal ANFIS prediction model. Furthermore, these measured tidal data of Fort Pulaski Tidal Station will also be employed to verify the validity and practicability of the proposed hybrid tidal prediction model. All the measured tidal level data employed in this study can be obtained from the web site http://co-ops.nos.noaa.gov/. In this study, The ANFIS model is constructed by totally five layers, which includes input layer with four input nodes (tidal level data at *t*-1, *t*-2, *t*-3 and *t*-4 moment), input membership function layer with several input membership function nodes, fuzzy rule layer with several fuzzy rules, and output membership function layer with several output membership function nodes which equals to the number of fuzzy rules, and the output layer with one output node (tidal level data at moment *t*). The detailed ANFIS model for the tidal level prediction is illustrated in Fig. 5.

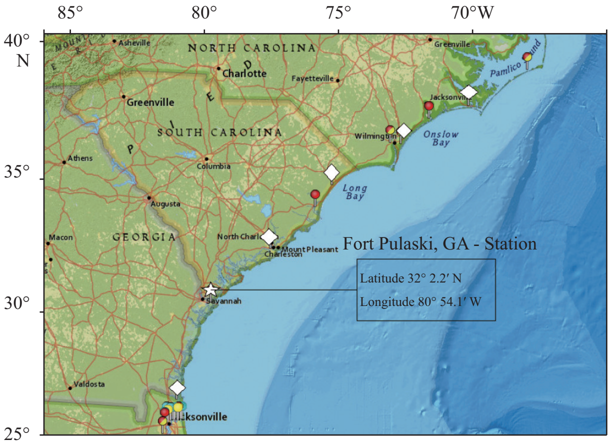

Fig. 4. Location map of the tidal level station on the Georgia and South Carolina coast of the United States (http://co-ops.nos.noaa.gov/).

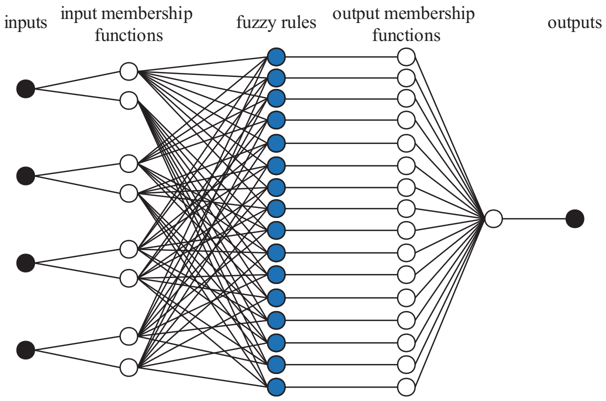

Fig. 5. The ANFIS structure of the tidal level prediction system.

3.2 *Grid partition mechanism*

The GP approach is employed for the generation of Sugeno FIS by utilizing the learning data base. The sample set is divided into numerous local fuzzy regions by utilizing an axis-paralleled partition based on the predefined membership functions in the process of the GP algorithm. The GP approach consists of eight different types of the membership function (gauss MF, dsig MF, gauss2 MF, gbell MF, pi MF, psig MF, trap MF and tri MF). The number of the input membership functions can be designated in combination with each input variable. Since the ANFIS-GP model is a kind of Takagi-Sugeno system, only one output can be obtained, and the linear output function is chosen in this study. Furthermore, the number of the fuzzy rules produced by the GP approach is equal to that of output membership functions. In this subsection, the alternative prediction models include numerous FIS structures are created by employing diverse membership function characteristics (types, numbers). The optimal forecasting model is investigated through a large number of simulation experiments. Eight sub-models (eight different types of input membership function together with the linear output function) are created for each alternative forecasting model in the simulation. Nevertheless, the experiment results of these alternative models are not provided in detail. Only, the forecasting model which possessed the minimum checking MAE and RMSE value, and its characteristics are represented as an optimal forecasting model. Through numerous experimental explorations, the characteristics and experimental results of the entire alternative forecasting models are shown in Table 1. The training parameters of

Table 1. Characteristics and experimental results of the alternative ANFIS-GP models

Model	Number of input MFs	Type of input MF	Type of output MFs	Number of output MFs	Number of rule	Time/s	Epoch	RMSE/m		MAE/m	
								Training set	Checking set	Training set	Checking set
ANFIS-GP1	[2 2 3 2]	gauss MF	linear	24	24	112.561 969	100	0.050 90	0.058 59	0.038 25	0.044 57
ANFIS-GP2	[2 3 2 2]	dsig MF	linear	24	24	105.177 371	100	0.051 05	0.058 65	0.038 37	0.044 67
ANFIS-GP3	[2 2 2 2]	gauss2 MF	linear	16	16	74.530 401	100	0.051 44	0.058 84	0.038 62	0.044 60
ANFIS-GP4	[2 2 2 2]	gbell MF	linear	16	16	86.102 379	100	0.051 33	0.058 56	0.038 52	0.044 52
ANFIS-GP5	[2 3 2 2]	pi MF	linear	24	24	100.846 349	100	0.051 29	0.059 16	0.038 54	0.045 02
ANFIS-GP6	[2 3 2 2]	psig MF	linear	24	24	100.406 400	100	0.051 05	0.058 65	0.038 37	0.044 67
ANFIS-GP7	[2 2 2 2]	trap MF	linear	16	16	55.677 716	100	0.052 01	0.059 25	0.039 02	0.044 82
ANFIS-GP8	[2 3 2 2]	tri MF	linear	24	24	104.189 945	100	0.051 72	0.058 96	0.038 78	0.044 85
ANFIS-GP9	[3 3 3 3]	gbell MF	linear	81	81	1051.324 340	100	0.049 03	0.060 93	0.037 15	0.046 27

the GP fuzzy system consists of maximum number of epochs, anticipation error, initial step size, step size decreasing rate and increasing rate, and the corresponding value of these parameters are 100, 0, 0.01, 0.9 and 1.1, respectively.

Table 1 shows that the ANFIS-GP4 which possesses the minimum checking error (MAE is 0.044 52 and RMSE is 0.058 56) is the optimal prediction model. The input membership function type of the ANFIS-GP4 model is the bell-shaped membership function named as gbellmf. For each input variable, there are two input membership functions, respectively. In addition, the optimal fuzzy rule number is 16. The fuzzy rule number has increased substantially with the growth of the membership function number for each input variable, which makes the training speed of the simulation model reduce significantly. Furthermore, it can be inferred from the ANFIS-GP9 model that with the growth of the fuzzy rule number, the checking error will be increased, while the training error will be reduced, and the corresponding training speed is also descended significantly.

3.3 *Fuzzy-c means mechanism*

FCM clustering approach is employed for the sake of producing a FIS structure in this subsection. The FIS structure will generate the fuzzy rules which are extracted from the sampling data characteristic. In this approach, the numbers of the membership functions and the fuzzy rules for the input and output variables will be obtained by the FCM mechanism. The number of fuzzy cluster centers which are created by the FCM method can be predefined in the modeling process. Furthermore, when the type of the FIS architecture is assigned as Takagi-Sugeno, the types of the input and output membership functions will be designated as gauss and linear, respectively. Nineteen alternative prediction models are created by predefining the number of the fuzzy cluster centers with a variation of 2–20. Additionally, the numbers of the membership functions and the fuzzy rules are determined by the number of cluster centers. Simulation experiments for these alternative prediction models are conducted for the selection of the optimal FIS architecture, and the model characteristics together with experimental results are shown in Table 2. It should not be concluded that better experimental results can be acquired with more cluster centers, which can be inferred from Table 2. Besides, as shown in Table 2, the ANFIS-FCM12 prediction model with 12 fuzzy cluster centers shows the smallest checking error (MAE is 0. 0449 7, and RMSE is 0. 059 28) and at this moment the RMSE value of the training error changes into 0.051 93. Consequently, the ANFIS-FCM12 model with 12 fuzzy rules is considered as the optimal prediction model. Furthermore, it can be concluded from Table 2 that the prediction model becomes over-fitting with the increase or decrease of the num-

Table 2. Characteristics and experimental results of the alternative ANFIS-FCM models

Model	Number of cluster center	Number of input MFs	Number of output MFs	Number of rule	Time/s	Epoch	RMSE/m		MAE/m	
							Training set	Checking set	Training set	Checking set
ANFIS- FCM2	2	[2 2 2 2]	2	2	26.266 174	100	0.056 37	0.061 18	0.042 49	0.047 10
ANFIS- FCM3	3	[3 3 3 3]	3	3	15.762 437	100	0.054 81	0.060 16	0.041 21	0.046 37
ANFIS- FCM4	4	[4 4 4 4]	4	4	20.328 44	100	0.054 86	0.060 48	0.041 34	0.046 42
ANFIS- FCM5	5	[5 5 5 5]	5	5	21.921 39	100	0.053 59	0.059 65	0.040 28	0.045 67
ANFIS- FCM6	6	[6 6 6 6]	6	6	23.433 090	100	0.053 81	0.059 99	0.040 51	0.046 04
ANFIS- FCM7	7	[7 7 7 7]	7	7	26.553 552	100	0.052 63	0.059 46	0.039 58	0.045 24
ANFIS- FCM8	8	[8 8 8 8]	8	8	29.771 149	100	0.052 51	0.059 30	0.039 44	0.045 11
ANFIS- FCM9	9	[9 9 9 9]	9	9	33.856 435	100	0.052 34	0.059 43	0.039 22	0.045 22
ANFIS- FCM10	10	[10 10 10 10]	10	10	38.464 749	100	0.052 22	0.059 62	0.039 13	0.045 24
ANFIS- FCM11	11	[11 11 11 11]	11	11	42.347 99	100	0.052 30	0.059 70	0.039 29	0.045 22
ANFIS- FCM12	12	[12 12 12 12]	12	12	47.102 289	100	0.051 93	0.059 28	0.038 94	0.044 97
ANFIS- FCM13	13	[13 13 13 13]	13	13	51.665 819	100	0.051 72	0.059 67	0.038 71	0.045 31
ANFIS- FCM14	14	[14 14 14 14]	14	14	56.653 430	100	0.051 52	0.059 97	0.038 60	0.045 43
ANFIS- FCM15	15	[15 15 15 15]	15	15	63.366 390	100	0.051 44	0.059 78	0.038 54	0.045 24
ANFIS- FCM16	16	[16 16 16 16]	16	16	66.637 524	100	0.051 40	0.059 69	0.038 51	0.045 24
ANFIS- FCM17	17	[17 17 17 17]	17	17	75.753 891	100	0.051 48	0.059 69	0.038 68	0.045 32
ANFIS- FCM18	18	[18 18 18 18]	18	18	79.162 627	100	0.051 11	0.060 06	0.038 26	0.045 57
ANFIS- FCM19	19	[19 19 19 19]	19	19	85.838 959	100	0.050 94	0.060 09	0.038 26	0.045 69
ANFIS- FCM20	20	[20 20 20 20]	20	20	166.260 874	100	0.051 04	0.060 48	0.038 30	0.045 93

ber of cluster centers, and the training speed of the model is also gradually descended with the increase of the number of cluster centers.

The training parameters of the FCM method in this study are composed of exponent of the partition matrix with defaulted value of 2, maximum number of iterations with defaulted value of 100 and error stopping criterion with defaulted value of 0.000 05. The number of the clustering centers will determine the number of the membership functions and the fuzzy rules of the FCM fuzzy system, and the training parameters of the FCM fuzzy system are the same as the SC fuzzy system.

3.4 *Sub clustering mechanism*

In the SC approach, it is presumed that each datum point is a potential clustering center. The likelihood that each datum point would become the clustering center is calculated based on the density index of neighboring data points, and those data points which are selected as the clustering centers possess the highest density index. To determine the new clustering center and its location, the SC method eliminates the entire data points around the first cluster center as confirmed by the radii, and the iteration loop continues until the entire data points are included in the radius of the clustering center (Cakmakci, 2007). Subtractive clustering method is composed of four algorithm parameters which include squash factor, range of influence, rejected ratio and accepted ratio (Chiu, 1994).

The optimal ANFIS model based on the SC approach is obtained by adjusting the clustering parameters manually. Since the variation of the range of influence (RI) value has a great influence on the model's fuzzy rules produced by the SC algorithm, and the fuzzy rules will affect the performance of the whole network model. Consequently, the RI value is varied from 0.1 to 0.8 with 0.05 increment. Furthermore, the remaining three parameters of the SC algorithm are set to defaulted value which includes squash factor with defaulted value of 1.25, accepted ratio with defaulted value of 0.5 and rejected ratio with defaulted value of 0.15, and because the variation of these SC parameters will not have a significant impact on the entire model. Consequently, 13 alternative prediction models are established for the selection of the optimal prediction model with the variation of the RI value. Types of the input and output membership functions are specified as gauss and linear for the entire alternative prediction models. The number of the membership function is determined by the SC algorithm. Furthermore, the number of the fuzzy rules is the same as the number of membership function of each input variable. The training parameters of the SC fuzzy system are the same as the GP fuzzy system.

The experimental results of these alternative prediction models are illustrated in Table 3. It can be inferred from Table 3 that the optimal SC forecasting model is ANFIS-SC5 which possesses the minimum checking error (corresponding MAE is 0. 044 93 and minimum RMSE is 0. 059 16). The simulation results of the ANFIS-SC5 model are obtained when the RI value is set as 0.40, and this model has ten fuzzy rules. Furthermore, it can be concluded from Table 3 that the lower RI values will generate the higher rule numbers. In addition, with the increase of the RI values, the convergence speed of the simulation model is gradually accelerated, while the prediction model becomes over-fitting with the increase or decrease of the value of RI.

3.5 *Comparison of three ANFIS prediction model*

Based on the simulation results of the ANFIS model of three different fuzzy system generating approaches, the optimal simulation models are ANFIS-GP4, ANFIS-SC5 and ANFIS-FCM12. In order to determine the best ANFIS model for tidal level forecasting from the three optimal models, the prediction performance of these produced models is evaluated by utilizing the testing data base obtained from Fort Pulaski Tidal Station with the duration from 2015-09-08 00:00 to 2015-12-31 23:00, totally 2 760 pairs of testing samples. The experimental results based on the same simulation environment and the same training parameters are shown in Table 4. The simulation results confirm that the ANFIS-GP4 with the GP approach is the best tidal level prediction model which possesses the highest *r* (0.997 90) and the lowest MAE and RMSE value (0.037 97 and 0.050 24).

Table 3. Characteristics and experimental results of the alternative ANFIS-SC models

Model	RI	Number of input MFs	Number of output MFs	Number of rule	Time/s	Epoch	RMSE/m		MAE/m	
							Training set	Checking set	Training set	Checking set
ANFIS-SC1	0.20	[30 30 30 30]	30	30	179.748 311	100	0.050 26	0.060 86	0.037 85	0.046 16
ANFIS-SC2	0.25	[18 18 18 18]	18	18	80.960 301	100	0.051 23	0.059 87	0.038 37	0.045 46
ANFIS-SC3	0.30	[13 13 13 13]	13	13	51.961 271	100	0.051 76	0.059 30	0.038 74	0.044 94
ANFIS-SC4	0.35	[12 12 12 12]	12	12	48.074 167	100	0.051 61	0.059 37	0.038 58	0.044 91
ANFIS-SC5	0.40	[10 10 10 10]	10	10	40.382 104	100	0.052 05	0.059 16	0.038 94	0.044 93
ANFIS-SC6	0.45	[8 8 8 8]	8	8	41.668 747	100	0.052 24	0.059 19	0.039 14	0.044 92
ANFIS-SC7	0.50	[7 7 7 7]	7	7	31.642 998	100	0.052 31	0.059 25	0.039 14	0.044 98
ANFIS-SC8	0.55	[6 6 6 6]	6	6	24.628 198	100	0.052 45	0.059 36	0.039 22	0.045 10
ANFIS-SC9	0.60	[6 6 6 6]	6	6	35.701 686	100	0.052 57	0.059 42	0.039 35	0.045 12
ANFIS-SC10	0.65	[6 6 6 6]	6	6	24.940 734	100	0.052 61	0.059 30	0.039 40	0.045 09
ANFIS-SC11	0.7	[4 4 4 4]	4	4	20.486 733	100	0.053 89	0.059 54	0.040 56	0.045 52
ANFIS-SC12	0.75	[4 4 4 4]	4	4	19.848 708	100	0.054 00	0.059 85	0.040 68	0.045 94
ANFIS-SC13	0.80	[4 4 4 4]	4	4	21.346 840	100	0.054 08	0.059 89	0.040 75	0.045 99

Table 4. Comparison of three ANFIS prediction models

Model	Time/s	Epoch	*r*	RMSE/m	MAE/m
ANFIS-GP4	646.043 193	100	0.997 90	0.050 24	0.037 97
ANFIS-SC5	46.093 081	100	0.997 85	0.050 81	0.038 21
ANFIS- FCM12	27.241 647	100	0.997 84	0.050 92	0.038 28

4 Hybrid tidal level prediction system

4.1 *Harmonic analysis*

Theoretically explaining, tides are the periodic vertical movements of the sea level, which are caused by the gravitational attraction between celestial bodies (especially the earth, the moon and the sun). The origin of tide is the tide-generating force of the celestial bodies in space which is a combination of the centrifugal and gravitational forces between the earth and the other celestial bodies.

In practice, a most commonly and widely utilized tidal level prediction approach is harmonic analysis method which approach decomposes the complicated tides into a couple of periodic constituents and each of which is generated by a hypothetical celestial body. At moment t, for a definite tide station with the height $h(t)$ of tide, it can be expressed as

$$h(t)=H_0+\sum_{k=1}^{\infty}f_kH_k\cos[\sigma_kt+(v_0+u)_k-g_k], \tag{1}$$

where H_0 is the mean sea level, f_k is the node factor, σ_k is the angular velocity of tidal components, $(v_0+u)_k$ is the initial phase of the tidal components, H_k is the amplitude of tidal components, and g_k is the time interval from the moments of the celestial body passing the upper transit meridian to the moments of high water occurance. H_k and g_k are the harmonic analysis constants.

In practice, the quantity of tidal components may be quite huge; fortunately, most of the tidal components can be ignored because of their smaller amplitude (H_k) and the longer period (g_k). In actual calculation, Eq. (1) can be rewritten as

$$h(t)=H_0+\sum_{k=1}^{n_c}h_k\cos(\omega_kt-\phi_k), \tag{2}$$

where n_c is the number of components, h_k is the amplitudes of the tidal components, ω_k is the angular velocity of the tidal components, and ϕ_k is the initial phase of the tidal components.

The harmonic analysis method is widely employed in tide prediction because of its stable prediction performance and simple calculation process. Though the harmonic analysis method is extensively used for tidal level prediction, the drawback of the conventional harmonic analysis method is quite apparent: the harmonic constant of this approach requires a considerable number of the measured tidal data which is a laborious and time-consuming job.

4.2 *Hybrid tide prediction model based on the combination of the ANFIS and harmonic model*

According to the origin of the tide, tidal level forecasting can be divided into two parts: the astronomical tide parts caused by the celestial bodies' movement and the non-astronomical tide parts caused by various meteorological factors. In addition, the trends of these meteorological factors are shown in Fig. 6.

The astronomical tide which is mainly caused by the tidal force of the celestial bodies has a notable time-varying characteristic, while the non-astronomical tide shows strong nonlinearity. In this study, the conventional harmonic analysis method is implemented to predict the astronomical tide section, and the residual tide section is predicted by the ANFIS prediction model. The residual section is a complicated association caused by the topography, and hydrological and meteorological factors. The ANFIS prediction model fits the residual tide section which owns obvious nonlinear performances accurately. The final prediction model is composed of the harmonic prediction model and the optimal ANFIS (ANFIS-GP4) prediction model. The composition and training process of the final modular prediction model are illustrated in Fig. 7.

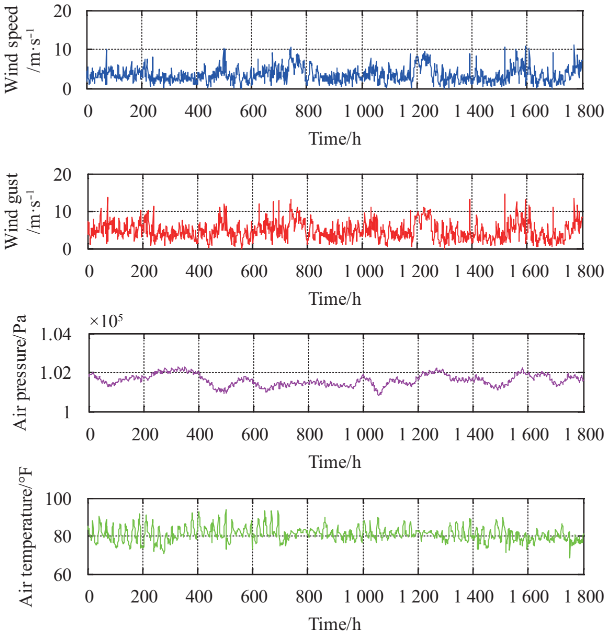

Fig. 6. Changes of meteorological elements.

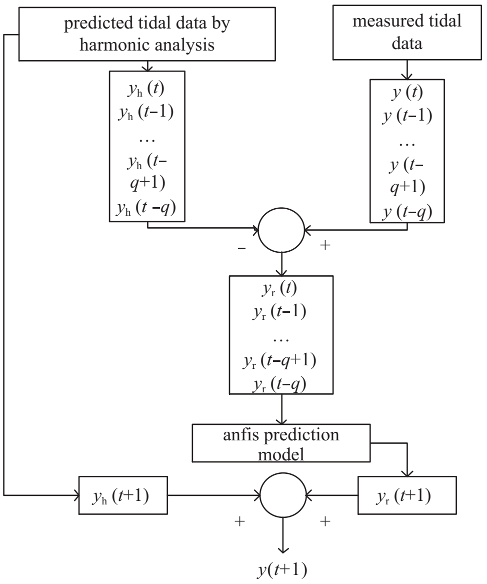

Fig. 7. Flow chart of the hybrid tidal prediction model.

In the hybrid prediction model, the observed tidal data $y(t)$, $y(t\text{-}1)$, ..., $y(t\text{-}q)$ are set as the input of the model. Where y_h, $y_h(t\text{-}1)$, ..., $y_h(t\text{-}q)$ are the tidal prediction values obtained by the harmonic prediction method, y_r, $y_r(t\text{-}1)$, ..., $y_r(t\text{-}q)$ are the difference between y and y_h. Finally, the ultimate modular prediction model is the combination of the harmonic prediction model and the ANFIS-GP4 prediction model. Eventually, the predictive tidal level $y(t+1)$ is the association of $y_h(t+1)$ and $y_r(t+1)$.

5 Tidal level prediction based on different forecasting approaches

The measured tidal level data at six east coast tidal stations in America are applied to validating the effectiveness of the proposed novel method. The locations of these tide gauge stations are shown in Fig. 4. The sampling period of each tide gauge station is identical to the sampling period of the Fort Pulaski Tidal Station. The simulation experiment of this research is based on Matlab2012b.

For a further comprehensive comparison of the simulation results, root mean square error (RMSE), mean square error (MSE), mean absolute error (MAE), standard deviation (SD) and mean error (ME) are introduced as an evaluation indicators to evaluate the performance of the proposed prediction model. The concepts of the $e_{\rm rms}$, the $e_{\rm ms}$, the $e_{\rm ma}$, the $d_{\rm s}$ and the $e_{\rm m}$ are shown as follows:

$$e_{\rm rms}=\sqrt{\frac{\sum_{k=1}^{n_{\rm sample}}(y_k-\hat{y}_k)^2}{N}}, \tag{3}$$

$$e_{\rm ms}=\frac{\sum_{k=1}^{n_{\rm sample}}(y_k-\hat{y}_k)^2}{N}, \tag{4}$$

$$e_{\rm ma}=\frac{1}{N}\sum_{k=1}^{n_{\rm sample}}{\rm abs}(y_k-\hat{y}_k), \tag{5}$$

$$d_{\rm s}=\sqrt{\frac{1}{N}\sum_{k=1}^{n_{\rm sample}}(y_k-\mu)^2}, \tag{6}$$

$$e_{\rm m}=\frac{1}{N}\sum_{k=1}^{n_{\rm sample}}(y_k-\hat{y}_k), \tag{7}$$

where $n_{\rm sample}$ is the number of the sample data, μ represents the arithmetic mean value of the observed data, y_k represents the value of the observation, and $\hat{y}_k$ denotes the value of prediction.

5.1 *Tidal level prediction by using the harmonic analysis model*

The tidal level prediction of Fort Pulaski Tidal Station by using harmonic analysis forecasting method is illustrated in Fig. 8 together with the observed ones, which is shown at the website of the NOAA, with the time resolution of 1 h. It can be obviously concluded from Fig. 8 that the predicted value of the tidal level falls below the observation value from the beginning to the end, which illuminates that the tidal level can be influenced by numerous unexpected factors of the environmental changes, such as the heavy sea surface wind and rough sea. Furthermore, there exists a considerable sustained deviation with the coefficient of regression r of 0.984 67 which is depicted in Fig. 8. The sustained offset between the predicted and observed tidal data reaches up to 0.5 m and the forecasting error of the harmonic analysis model is in a range of [−0.5, 0.6], which can be revealed from the error diagram in Fig. 8. Furthermore, the harmonic prediction model also has a larger prediction error which can be seen from the error histogram analysis in Fig. 8

Figure 9 is the linear regression diagram of the measured and predicted results by using the harmonic analysis method.

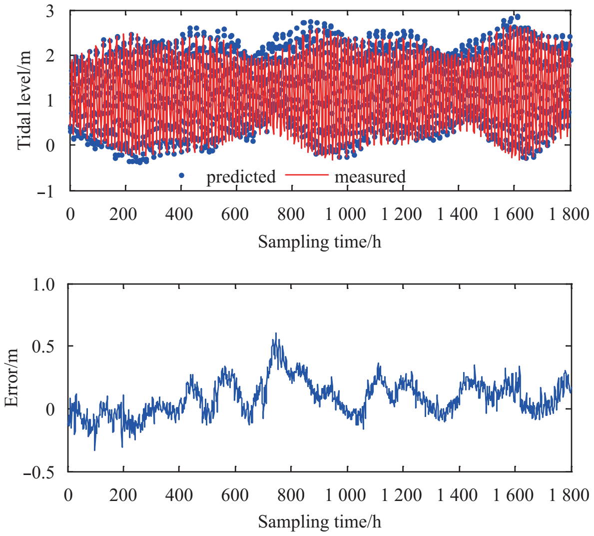

Fig. 8. Simulation results by using the harmonic analysis method (Fort Pulaski Tidal Station). The red line denotes the measured value of the tidal level, and the dot symbols represent the predicted values of the tidal level by using the harmonic analysis method.

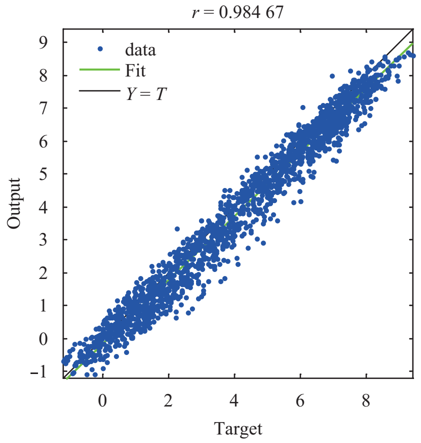

Fig. 9. Linear regression diagram by using the harmonic analysis method (Fort Pulaski Tidal Station). Y=T is a standard linear function. This figure and subsequent graphics represent the degree of correlation between the output variable (predicted value) and the target variable (measured value).

5.2 *Tidal level prediction by using the conventional auto regression (AR) and back propagation (BP) model*

The BP neural network is a multilayer feed-forward networks which is proposed by Rumelhant and McClelland in 1986. The learning and training procedure of the BP network is composed by the back propagation of error information and the forward propagation of learning information. The BP neural network has been employed by Lee (2004, 2006, 2008) and Lee and Jeng (2012) for the prediction of the short-term storm surge and tidal predictions, and it also acquired favorable effect in practice.

The AR model is a traditional linear prediction mechanism, and it is also known as N-order autoregressive model. The core arithmetic of this model is relatively simple, but its application is quite reliable. The AR model is a kind of classic researching model in the time series analysis and is widely employed in the field of prediction (Mabrouk et al., 2008; Lin et al., 2004). The significance of the AR model is that it can characterize the influence and the effect of the factors on the prediction target only through the historical observation of the time series, and it is independent of the assumption that the model variables are independent of each other. In addition, the constructed model can eliminate the adverse effect due to the choice of independent variables and multicol-linearity problem in traditional regression forecasting models. A novel sparse AR model is also proposed for the prediction of tidal current (Lu et al., 2015).

The results of tidal level predictions at Fort Pulaski Tidal Station by using the AR and BP methods are shown in Figs 10 and 11, respectively. The deviation between the predicted and measured values is quite apparent and the predicted results are in poor agreement with the measured results, which can be further confirmed by linear regression diagram in Figs 12 and 13. In addition, the prediction error of the conventional AR model is significantly larger than that of the conventional harmonic analysis prediction error and the BP prediction error, which can also be concluded from the error diagram in Figs 8, 10 and 11. The prediction error of AR and BP forecasting model is in the range of [–0.7, 0.7] and [–0.3, 0.3], respectively.

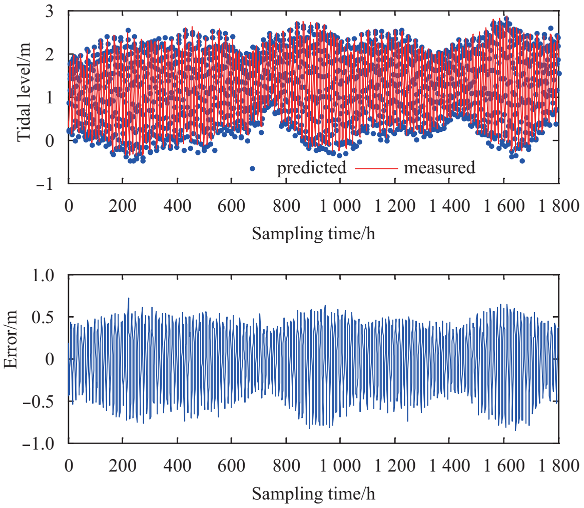

Fig. 10. Simulation results by using the AR method (Fort Pulaski Tidal Station). The red line denotes the measured value of the tidal level, and the dot symbols represent the predicted values of the tidal level by using the AR method.

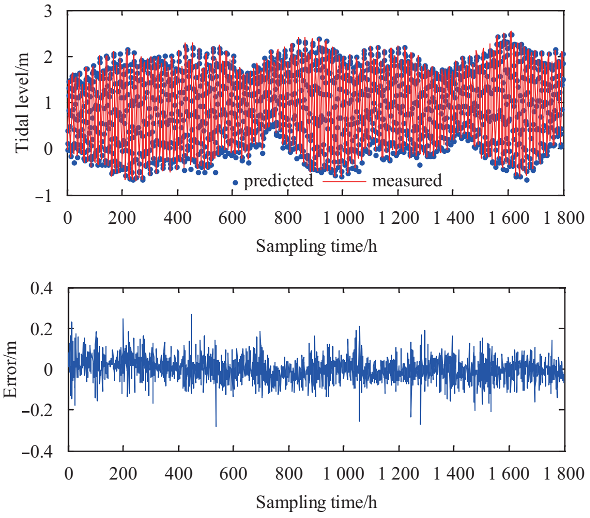

Fig. 11. Simulation results by using the BP method (Fort Pulaski Tidal Station). The red line denotes the measured value of the tidal level, and the dot symbols represent the predicted values of the tidal level by using the BP method.

Figures 12 and 13 are the linear regression diagrams of the measured and predicted results by using the conventional AR and BP forecasting methods. The regression coefficients of the conventional AR and BP models are 0.889 00 and 0.997 15, respectively.

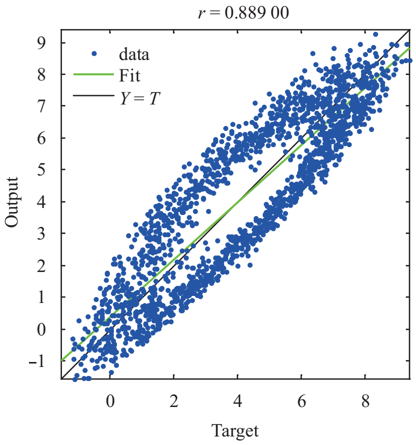

Fig. 12. Linear regression diagram by using the AR method (Fort Pulaski Tidal Station).

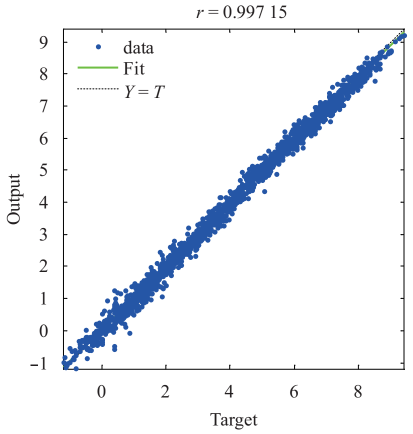

Fig. 13. Linear regression diagram by using the BP method (Fort Pulaski Tidal Station).

5.3 *Tidal level prediction by using the conventional ANFIS-GP4 prediction model*

It can be apparently recognized from Fig. 14 that the predicted values generated by the ANFIS-GP4 forecasting model coincide well with the observed tidal data than the conventional harmonic method. The prediction error of the ANFIS forecasting model is in the range of [–0.25, 0.25], as shown in Fig. 14. Figure 14 also illustrates that the prediction accuracy is significantly improved compared with the harmonic prediction model. Nevertheless, the error fluctuation range is relatively large as shown in Fig. 14.

Figure 15 is the linear regression diagram of the measured and predicted results by using the conventional ANFIS-GP4 fore-

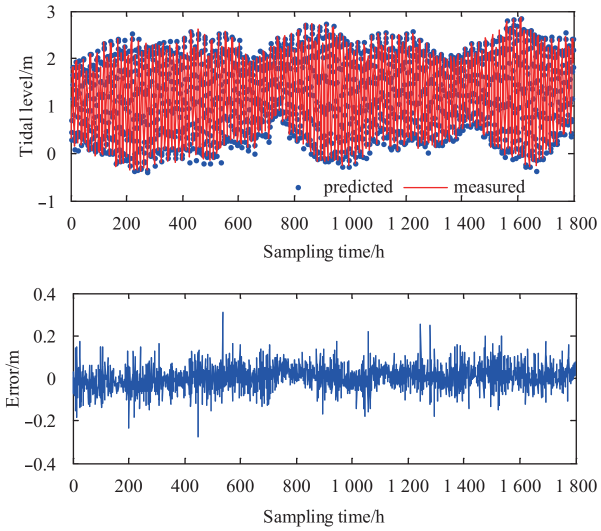

Fig. 14. Simulation results by using the conventional ANFIS-GP4 method (Fort Pulaski Tidal Station). The red line denotes the measured value of the tidal level, and the dot symbols represent the predicted values of the tidal level by using the ANFIS-GP4 model.

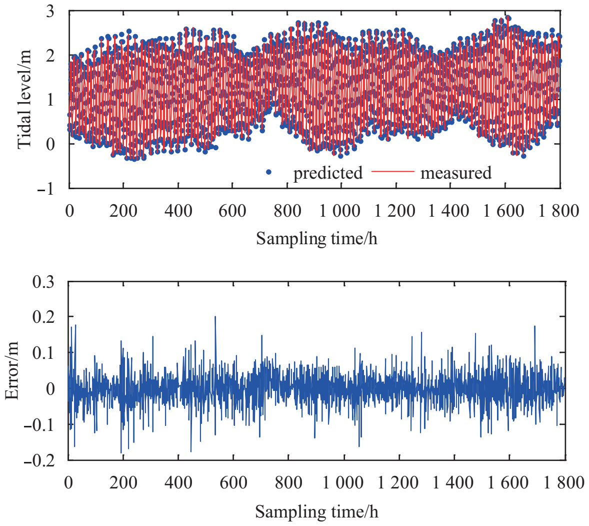

Fig. 16. Simulation results by using the hybrid ANFIS-GP4 method (Fort Pulaski Tidal Station). The red line denotes the measured value of the tidal level, and the dot symbols represent the predicted values of the tidal level by using hybrid ANFIS-GP4 method.

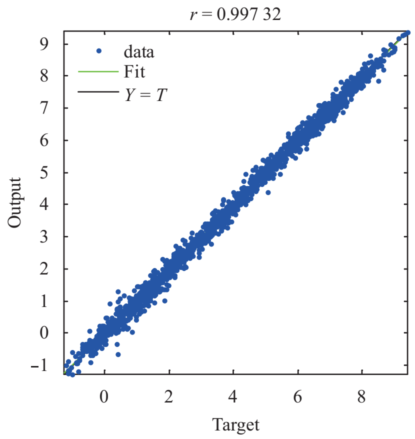

Fig. 15. Linear regression diagram by using the conventional ANFIS-GP4 method (Fort Pulaski Tidal Station).

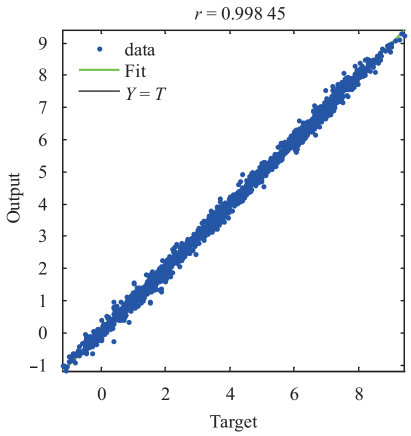

Fig. 17. Linear regression diagram by using hybrid ANFIS-GP4 (Fort Pulaski Tidal Station).

casting method. The closeness between the observed tidal values and the forecasted ones can be further validated by the corresponding higher regression coefficient of 0.99732, as shown in Fig. 15.

5.4 *Tidal level prediction by using the hybrid ANFIS prediction model*

The tidal level forecasting results by using the hybrid ANFIS prediction method are represented in Fig. 16 together with the real observations. It can be noticed from the Error diagram in Fig. 16 that there is no large continuous offset between the predicted values and the actual ones. Furthermore, the forecasting results are located around the real ones conformably, which can also be obtained from Fig. 17. The closeness between the observed tidal values and the forecasted ones can be further validated by the corresponding highest regression coefficient of 0.998 45, as shown in Fig. 17. Besides, the forecasting error of the hybrid ANFIS-GP4 model is in the range of [−0.2, 0.2], which can be revealed from the error diagram in Fig. 16. In addition, the error fluctuation range is significantly reduced compared with the previously two models, which is shown in Fig. 16. In conclusion, the proposed hybrid prediction model significantly improves the prediction accuracy of the tidal level.

Figure 17 expresses the linear regression diagram of the measured and predicted results by using the hybrid ANFIS-GP4 approach.

In order to further quantitatively validate the predictive performance of the hybrid ANFIS-GP4 prediction model proposed in this study, detailed simulation performance results based on the same simulation environment, the same simulation parameters and the same measured tidal data of Fort Pulaski Tidal Station are shown in Table 5.

It is obviously seen from Table 5 that the corresponding error value of the hybrid ANFIS-GP4 prediction model is the smallest of the three prediction models. Furthermore, the corresponding training speed is the fastest, too. Consequently, the experimental results demonstrate that the proposed prediction model has a higher forecasting accuracy for the tidal level prediction.

Different time spans of time series prediction will influence the magnitude of the prediction error. In order to further compare the accuracy of the conventional ANFIS model and the hybrid ANFIS model in the prediction of the tidal level, simulation experiments of different time spans are preformed based on the same observed tidal level data, and the same simulation environment. The simulation results are shown in Tables 6 and 7. From Table 6 we can see that the RMSEs of the tidal level prediction by using the conventional ANFIS-GP4 method are 0.107 70 and 0.191 66 for 1 h ahead and 3 h ahead forecasting, respectively. In addition, values of RMSE are in a range of 0.107 70 and 0.207 99 for 3 h or more hours-ahead predictions. Besides, the RMSE values of tidal level prediction by using the hybrid ANFIS-GP4 method for 1 h ahead and 3 h ahead predictions are 0.061 37 and 0.074 63 which can be obtained from Table 7, and values of RMSE are in the range of 0.061 37 and 0.106 29 for 3 h or more hours-ahead forecasting. The simulation results confirm that the proposed novel hybrid prediction method is able to provide stable and reliable prediction results for the tidal level forecast with satisfactory prediction accuracy. In addition, with the increase of the forecasting time interval, the proposed novel method also exhibits higher forecasting accuracy compared with the conventional ANFIS-GP4 method, which can also be obtained from the corresponding error value illustrated in Tables 6 and 7. Moreover, the time consumption of the whole simulation process has reduced obviously, which validates the high-efficiency of the proposed approach.

Table 5. Tidal level simulation results of Fort Pulaski Tidal Station

Model	Time/s	Training epoch	Step	r	RMSE/m	MAE/m	MSE/m	Mean error/m	Standard deviation/m
Harmonic analysis	–	–	1 800	0.984 67	0.163 36	0.126 69	0.087 55	0.085 60	0.139 17
AR	11.137 496	–	1 800	0.889 00	0.377 62	0.334 58	0.467 81	0.009 76	0.377 59
BP	8.672 851	100	1 800	0.997 15	0.061 44	0.047 25	0.012 39	−0.010 92	0.060 48
ANFIS-GP4	60.649 794	100	1 800	0.997 32	0.058 56	0.044 52	0.011 25	0.004 98	0.058 36
Hybrid-ANFIS-GP4	55.331 230	100	1 800	0.998 45	0.044 46	0.033 32	0.006 48	−0.001 52	0.044 44

Table 6. Fort Pulaski Tidal Station p-hour-ahead tidal prediction by the conventional ANFIS-GP4

Prediction time span	RMSE/m	MAE/m	MSE/m	Mean error/m	Standard deviation/m	r	Time/s
1 h ahead	0.107 70	0.083 25	0.038 06	0.012 68	0.106 98	0.990 97	91.045 2
3 h ahead	0.191 66	0.150 92	0.120 51	0.026 51	0.189 87	0.971 24	70.794 0
6 h ahead	0.207 99	0.165 84	0.141 93	0.031 15	0.205 70	0.966 23	48.711 9
12 h ahead	0.197 71	0.157 92	0.128 25	0.025 38	0.196 14	0.969 38	530.104 9
24 h ahead	0.128 19	0.101 30	0.053 91	0.014 64	0.127 38	0.987 24	36.081 3

Table 7. Fort Pulaski Tidal Station p-hour-ahead tidal prediction by the hybrid ANFIS-GP4

Prediction time span	RMSE/m	MAE/m	MSE/m	Mean error/m	Standard deviation/m	r	Time/s
1 h ahead	0.061 37	0.046 75	0.012 36	−0.002 07	0.061 36	0.997 03	38.926 2
3 h ahead	0.074 63	0.057 10	0.018 27	−0.003 86	0.074 55	0.995 60	38.736 3
6 h ahead	0.074 43	0.057 40	0.018 17	−0.005 41	0.074 25	0.995 67	36.604 2
12 h ahead	0.082 30	0.063 64	0.022 22	−0.006 07	0.082 09	0.994 71	32.704 2
24 h ahead	0.106 29	0.083 94	0.037 07	−0.013 78	0.105 43	0.991 27	35.417 5

Table 8. Simulation of tidal level obtained from five tidal stations by using the hybrid ANFIS-GP4

Tidal station	Prediction span/h	RMSE/m	MAE/m	MSE/m	Mean error/m	Standard deviation/m	r	Time/s
Beaufort	1	0.040 12	0.029 61	0.005 28	0.000 30	0.040 12	0.994 19	88.970 13
	12	0.063 13	0.046 20	0.013 08	0.000 53	0.063 14	0.985 53	143.761 875
	24	0.074 44	0.056 69	0.018 18	0.000 49	0.074 44	0.979 99	145.898 412
	48	0.088 07	0.068 09	0.025 45	0.000 56	0.088 08	0.972 09	185.274 982
Wrightsville Beach	1	0.054 85	0.040 50	0.009 87	−0.000 48	0.054 85	0.993 44	242.249 947
	12	0.084 45	0.062 54	0.023 40	0.000 29	0.084 46	0.984 31	161.770 060
	24	0.098 74	0.076 18	0.031 99	0.000 24	0.098 75	0.978 60	151.974 041
	48	0.113 78	0.089 79	0.042 47	−0.000 44	0.113 79	0.971 73	82.406 640
Springmaid Pier	1	0.065 22	0.047 72	0.013 96	−0.000 55	0.065 22	0.993 81	127.460 160
	12	0.101 27	0.073 95	0.033 64	0.000 49	0.101 27	0.984 92	145.312 979
	24	0.115 54	0.087 23	0.043 80	−0.000 06	0.115 55	0.980 37	139.445 828
	48	0.126 20	0.098 59	0.052 25	0.000 15	0.126 21	0.976 86	140.372 480
Charleston	1	0.055 80	0.040 47	0.010 22	−0.001 06	0.055 80	0.995 85	128.853 659
	12	0.093 67	0.068 63	0.028 79	−0.000 68	0.093 68	0.988 16	122.790 396
	24	0.111 40	0.085 23	0.040 72	−0.001 85	0.111 40	0.983 26	124.990 196
	48	0.128 15	0.100 22	0.053 88	−0.002 82	0.128 13	0.978 21	131.101 237
Fernandina Beach	1	0.053 46	0.040 82	0.009 38	−0.000 14	0.053 46	0.997 04	137.180 740
	12	0.093 70	0.067 83	0.028 80	−0.002 01	0.093 68	0.990 85	129.573 803
	24	0.119 65	0.089 11	0.046 97	−0.004 14	0.119 59	0.985 11	154.554 347
	48	0.138 00	0.109 84	0.062 48	−0.006 54	0.137 85	0.980 49	148.775 555

5.5 *Performance of tidal level prediction*

To validate the forecasting accuracy, universality and generalization ability for longer time scope tidal level predictions, the simulation experiments are preformed based on the observed tidal level data of five different tidal stations located on the east coast of the Georgia and South Carolina in America. The simulation results of multistep tidal prediction based on the same simulation environment and the same simulation parameters are shown in Table 8. The time duration of the observed tidal level data of the five tidal stations is the same as the Fort Pulaski Tidal Station. It is revealed in Table 8 that the proposed approach can achieve stable and accurate tidal predictions for the observed tidal data of the five different tidal stations which are shown in Fig. 4. In addition, simulation results also illustrate that the proposed the hybrid prediction method is able to provide stable and reliable prediction results for the tidal level forecast with satisfactory prediction accuracy.

6 Conclusions

In this study, three different FIS generation approaches named as GP, FCM and SC are employed inthe establishment of the ANFIS based precise tidal level prediction models. The simulation experiments demonstrate that the GP algorithm is the best approach for the generation of the FIS in the three FIS-producing approaches. Besides, the GP fuzzy system possesses the highest prediction accuracy, which can be concluded from Table 4. Furthermore, a hybrid tidal level prediction model is proposed based on the combination of the harmonic analysis method and the ANFIS-GP4 method, which takes both advantages of them. The harmonic analysis model and the ANFIS-GP4 model denote the effects of the celestial bodies and the meteorological elements, respectively. The effectiveness and feasibility of the proposed model is demonstrated by the experimental results of the precise tidal level prediction. In addition, the experimental results also confirm that the proposed novel model can produce accurate predictions with high operating speed compared with the conventional AR, BP and the ANFIS-GP4 model, which is illustrated in Tables 5 and 7.

In general, application of the combination of the ANFIS-GP4 model and the harmonic analysis model in topology design of the hybrid prediction model testified that the proposed prediction model can be utilized successfully for the prediction of precise tidal level in the field of marine engineering.

References

Cakmakci M. 2007. Adaptive neuro-fuzzy modelling of anaerobic digestion of primary sedimentation sludge. Bioprocess and Biosystems Engineering, 30(5): 349–357

Chang F J, Lai H C. 2014. Adaptive neuro-fuzzy inference system for the prediction of monthly shoreline changes in northeastern Taiwan. Ocean Engineering, 84: 145–156

Chiu S L. 1994. Fuzzy model identification based on cluster estimation. Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology, 2(3): 267–278

Clue V. 2004. Harmonic analysis. In: Proceedings of the 2004 IEEE Electro/Information Technology Conference. Milwaukee USA, 53–58

Fang Guohong, Zheng Wenzhen, Chen Zongyong, et al. 1986. Analysis and Prediction of Tide and Tidal Currents (in Chinese). Beijing: China Ocean Press, 1–20

Günaydın K. 2008. The estimation of monthly mean significant wave heights by using artificial neural network and regression methods. Ocean Engineering, 35(14–15): 1406–1415

Haykin S S. 1999. Neural Networks: A Comprehensive Foundation. 2nd ed. New Jersey, USA: Prentice Hall, 33–102

He Shijun, Zhou Wenjun, Zhou Ruyan, et al. 2012. Study of tide prediction method influenced by nonperiodic factors based on support vector machines. Acta Oceanologica Sinica, 31(5): 160–164

Huang Wenrui, Murray C, Kraus N, et al. 2003. Development of a regional neural network for coastal water level predictions. Ocean Engineering, 30(17): 2275–2295

Jain P, Deo M C. 2007. Real-time wave forecasts off the western Indian coast. Applied Ocean Research, 29(1–2): 72–79

Jang J S R. 1993. ANFIS: adaptive-network-based fuzzy inference system. IEEE Transactions on Systems, Man, and Cybernetics, 23(3): 665–685

Jang J S R, Sun C T. 1995. Neuro-fuzzy modeling and control. Proceedings of the IEEE, 83(3): 378–406

Lee T L. 2004. Back-propagation neural network for long-term tidal predictions. Ocean Engineering, 31(2): 225–238

Lee T L. 2006. Neural network prediction of a storm surge. Ocean Engineering, 33(3–4): 483–494

Lee T L. 2008. Back-propagation neural network for the prediction of the short-term storm surge in Taichung harbor, Taiwan. Engineering Applications of Artificial Intelligence, 21(1): 63–72

Lee T L, Jeng D S. 2002. Application of artificial neural networks in tide-forecasting. Ocean Engineering, 29(9): 1003–1022

Lin Chunsheng, Deng Daxin, Ren Dekui. 2004. Adaptive AR model prediction filtering for ship hydraulic pressure signal on wind wave background. Haiyang Xuebao (in Chinese), 26(4): 133–138

Lu Xiaopeng, Ye Qingwei, L Cuilan. 2015. Tidal current prediction based on the sparse AR model. Journal of Marine Sciences, 33(2): 14–18

Mabrouk A B, Abdallah N B, Dhifaoui Z. 2008. Wavelet decomposition and autoregressive model for time series prediction. Applied Mathematics and Computation, 199(1): 334–340

Mekanik F, Imteaz M A, Talei A. 2016. Seasonal rainfall forecasting by adaptive network-based fuzzy inference system (ANFIS) using large scale climate signals. Climate Dynamics, 46(9–10): 3097–3111

Nezlin N P, Li Bailian. 2003. Time-series analysis of remote-sensed chlorophyll and environmental factors in the Santa Monica-San Pedro Basin off Southern California. Journal of Marine Systems, 39(3–4): 185–202

Ruano A E. 2005. Intelligent Control Systems Using Computational Intelligence Techniques. London: Institution of Engineering and Technology, 219–252

Stefanakos C. 2016. Fuzzy time series forecasting of nonstationary wind and wave data. Ocean Engineering, 121: 1–12

Takagi T, Sugeno M. 1985. Fuzzy identification of systems and its applications to modeling and control. IEEE Transactions on Systems, Man, and Cybernetics, SMC-15(1): 116–132

Xiao Bin, Qiao Fangli, Shu Qi. 2016. The performance of a z-level ocean model in modeling the global tide. Acta Oceanologica Sinica, 35(11): 35–43

Yetilmezsoy K, Özkaya B, Cakmakcı M. 2011. Artificial intelligence-based prediction models for environmental engineering. Neural Network World, 21(3): 193–218

Yin Jianchuan, Wang Nini. 2013. Online grey prediction of ship roll motion using variable RBFN. Applied Artificial Intelligence, 27(10): 941–960

Yin Jianchuan, Wang Nini. 2015. A variable multidimensional fuzzy model and its application to online tidal level prediction. Journal of Computational and Theoretical Nanoscience, 12(7): 1384–1394

Yin Jianchuan, Wang Nini. 2016. An online sequential extreme learning machine for tidal prediction based on improved Gath-Geva fuzzy segmentation. Neurocomputing, 174: 85–98

Yin Jianchuan, Zou Zaojian, Xu Feng. 2013. Sequential learning radial basis function network for real-time tidal level predictions. Ocean Engineering, 57: 49–55

Yin Jianchuan, Zou Zaojian, Xu Feng, et al. 2014. Online ship roll motion prediction based on grey sequential extreme learning machine. Neurocomputing, 129: 168–174

Young P, Shellswell S. 1972. Time series analysis, forecasting and control. IEEE Transactions on Automatic Control, 17(2): 281–283

Zhao Yingjie, Li Mingchang, Li Guanglou, et al. 2014. Time series correlation analysis of pollution in marine environment waters. Applied Mechanics and Materials, 522–524: 52–55