Water-Level Forecasting Using Neuro-wavelet Technique

Pradnya Dixit and Shreenivas Londhe

Abstract In the support of all ocean-related activities, it is necessary to predict the actual seawater levels as accurate as possible. The present work aims at forecasting the water levels from 3 to 6 weeks in advance at three locations: Dauphin Island, AL (Gulf of Mexico); Portland, ME (Gulf of Maine); and Cordova, AK (Gulf of Alaska) of divergent oceanic environment along the US coastline using neuro-wavelet technique (NWT) which is a combined approach of wavelet transform (WT) and artificial neural network. For this, time series of water-level anomaly (difference between the observed water level and harmonically predicted tidal level) was used to develop the NWT models at respective stations to predict the water levels for three different lead times from 3 to 6 weeks ahead. For this, hourly observed water levels along with harmonic tides were obtained from the National Oceanic and Atmospheric Administration of USA. The time series of water-level anomaly was decomposed using discrete wavelet transform (DWT) into low (approximate) and high (detail) frequency components. Further, these approximate coefficients were decomposed up to the desired level of decomposition (third and fifth levels) by multiresolution analysis of WT in order to provide more detailed and approximate components which ultimately provides relatively smooth varying amplitude series to develop the NWT models. Thus, the effect of autocorrelation in time series analysis was removed by decorrelating it using WT. Neural networks were trained with these decorrelated approximate and detailed wavelet coefficients. The outputs of networks during testing were reconstructed back using inverse DWT. Network-predicted anomaly was then added to harmonic tidal level to predict the water levels. Performance of NWT models was judged by drawing the water-level plots and other error measures. The NWT models performed reasonably well for all forecasting intervals at all the locations.

Keywords Water levels · Harmonic tides · Water-level anomaly
Artificial neural networks · Neuro-wavelet technique

P. Dixit (✉) · S. Londhe
Civil Engineering, Vishwakarma Institute of Information Technology, Pune, India
e-mail: pradnya.dixit@viit.ac.in

S. Londhe
e-mail: shreenivas.londhe@viit.ac.in

1   Introduction

Accurate information of seawater levels and their variations is required for planning, construction, operation, and maintenance works of various coastal as well as off-shore structures. Generally, variations in seawater level are large enough to disturb the day-to-day operations of different coastal structures in nearby areas, particularly in shallow water depth as well as safety of maritime activities and human lives. It generates the necessity of accurate prediction of seawater levels. Traditionally, harmonic analysis is used for tide predictions, but often the values of predicted tide and the measured (observed) water levels are not identical [1]. In recent years as an alternative modeling approach to overcome the drawbacks of traditional methods, researchers have applied the domain of data-driven techniques in which applications of artificial neural networks and genetic programming are predominant. Though these techniques proved their efficiency in prediction accuracy when modeling using univariate time series is concerned the competency of these techniques becomes a question as 'phase lag' or 'time lag' occurs in most of all the forecasts. This can be attributed to the 'effect of autocorrelation' which inherently occurs in any univariate time series modeling. Many researchers earlier have used the same 'seawater anomaly' to predict the correct seawater-level anomaly, and from the predicted anomaly, further they had predicted the tides. But the use of seawater-level anomaly as input tends toward the univariate time series modeling and effect of autocorrelation plays its role which ultimately results through a 'time lag or phase lag' in prediction and then tide prediction exercise becomes ineffective one.

Captivating this as motivation, the present study aims in predicting the accurate seawater levels by removing the prediction lag not only at short lead time intervals but long as 3–6 weeks in advance at three locations: Dauphin Island, AL (Gulf of Mexico); Portland, ME (Gulf of Maine); and Cordova, AK (Gulf of Alaska) of divergent oceanic environment along the US coastline using neuro-wavelet technique (NWT) which is a combined approach of wavelet transform (WT) and artificial neural network. For this, time series of water-level anomaly (difference between the observed water level and harmonically predicted tidal level) was used to develop the NWT models at respective stations to predict the water levels for different lead times: 3 weeks (3w), 4 weeks (4w), 5 weeks (5w), and 6 weeks (6w) ahead. For this, hourly observed water levels along with harmonic tides were obtained from the National Oceanic and Atmospheric Administration of USA. The time series of water-level anomaly was decomposed using discrete wavelet transform (DWT) into low (approximate) and high (detail) frequency components. Further, these approximate coefficients were decomposed up to the desired level of decomposition (third and fifth levels) by multiresolution analysis of WT in order to provide more detailed and approximate components which ultimately provides relatively smooth varying amplitude series to develop the NWT models. Thus, the effect of autocorrelation in time series analysis was removed by decorrelating it using WT. Neural networks were trained with these decorrelated approximate and detailed wavelet coefficients. The outputs of networks during testing were reconstructed back using inverse DWT.

Network-predicted anomaly was then added to harmonic tidal level to predict the water levels. Performance of NWT models was judged by drawing the water-level plots and other error measures. The NWT models performed reasonably well for all forecasting intervals at all the locations.

2 Study Area and Data

For the present work, three tidal stations, namely Dauphin Island, AL (Gulf of Mexico); Portland, ME (Gulf of Maine); and Cordova, AK (Gulf of Alaska), were selected which are from different oceanic and meteorological environments and maintained by the National Water Level Program (NWLP) of National Oceanic and Atmospheric Administration (NOAA) of the USA. Hourly observed water-level data from the year 2000 to 2005 along with harmonic tidal data was used to train and test the model. The locations of these stations are depicted in Table 1 and Fig. 1. It can be observed from Fig. 1 that, station Dauphin Island is inside the Gulf of Mexico region which experiences very severe hurricane events every year from June to November and where the effect of hurricane winds will be of greater extent on water levels. On the other hand station Portland, ME in Gulf of Maine facing an open sea which experiences extreme cold weather conditions along with wind forcing due to tropical storms. The difference between maximum and minimum tidal levels at these locations also indicates that there is a large variation in water levels. The third station Cordova, AK, which is in Gulf of Alaska experiences a diverse wind condition than the previous two stations as it is located in the different meteorological environments than those previous stations. Thus, in the present work, the applicability of the neuro-wavelet technique in various oceanic conditions will be judged characteristically for removing the lag in prediction as well as for accurate prediction at higher lead time intervals. Readers are directed to 'http://tidesandcurrents.noaa.gov' for further details.

Table 1 Location of tidal stations

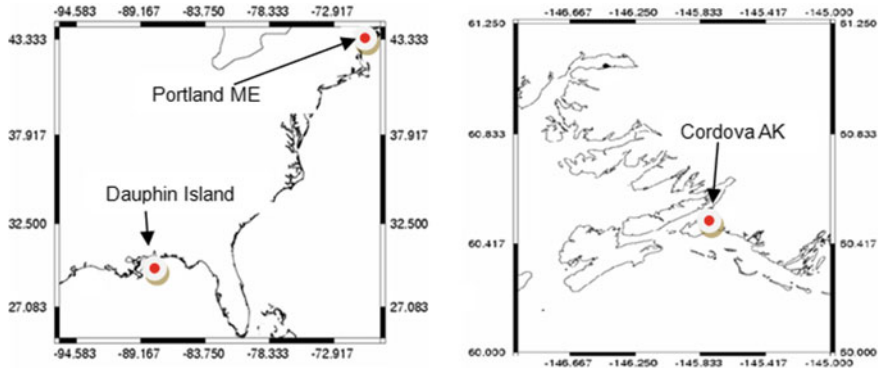| Sr. No.	Station	Location	Maximum WL (m)	Minimum WL (m)	Difference
1	Dauphin AL 8735180	30° 15′N, 88° 4.5′W	2.136	−0.405	2.541
2	Portland ME 8418150	43° 39.4′N, 70° 14.8′W	3.85	−0.948	4.798
3	Cordova AK 9454050	60° 33.5 N, 145° 45.1 W	5.172	−1.245	6.417

Fig. 1 Study area (Gulf of Mexico, Gulf of Maine, and Gulf of Alaska)

3 Neuro-Wavelet Technique

As mentioned in the introduction, a neuro-wavelet technique (NWT) is the combination of artificial neural network and the discrete wavelet transform. The discrete wavelet transform analyzes the frequency of the signal with respect to time at different scales. It decomposes time series into low (approximate) and high (detail) frequency components. The decomposition of approximate can be carried out further up to desired multiple levels in order to provide more detailed and approximate components which provides relatively smooth varying amplitude series. The neural network is thus trained with decorrelated approximate and detailed wavelet coefficients. The outputs of networks during testing are generally reconstructed back using inverse DWT. Figure 2 shows the generalized algorithm for the neuro-wavelet model. The total data set of observed water-level anomaly values is filtered into approximate (CA1) and detailed (CD1) components at the first decomposition level. Further, first approximate coefficient (CA1) is filtered into approximate (CA2) and detailed (CD2) components in the second-level decomposition. At the third-level decomposition, the second approximate coefficient (CA2) is again decomposed in approximation (CA3) and detailed (CD3) components. Finally, for third-level decomposition system, three detailed components and one approximate component are used to train the neural network (Fig. 3). This is called the multiple-level decomposition using wavelet transform. As authors have successfully applied the NWT for forecasting the ocean waves, readers are directed to Dixit et al. [2–4] for further details of neuro-wavelet technique along with the details of artificial neural network (ANN) and wavelet transforms (WTs). The choice of wavelet transformation is in fact an important part of wavelet analysis and depends very much upon both the properties of the signal under investigation and what the investigator is looking for.
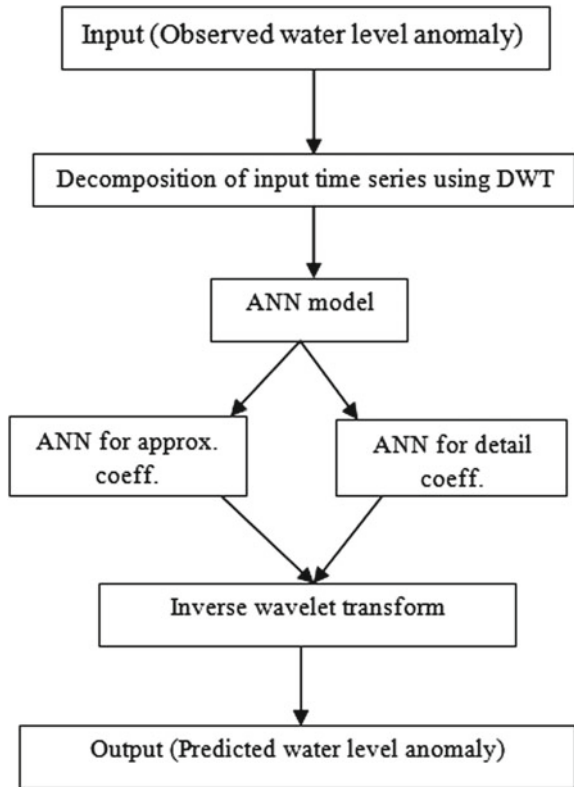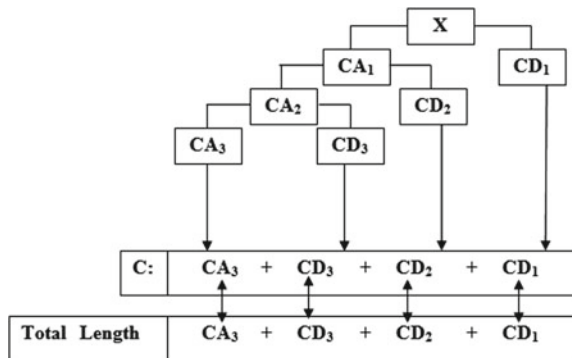
Fig. 2 Algorithm of NWT

Fig. 3 Multilevel decomposition by wavelet transform

As mentioned earlier, instead of observed water-level time series, the time series of water-level anomaly (difference between the observed water level and harmonically predicted tidal level) is used in the present work to develop the NWT models at the respective stations to predict the water levels for different lead time intervals. For this, time series of water-level anomaly was decomposed using discrete wavelet
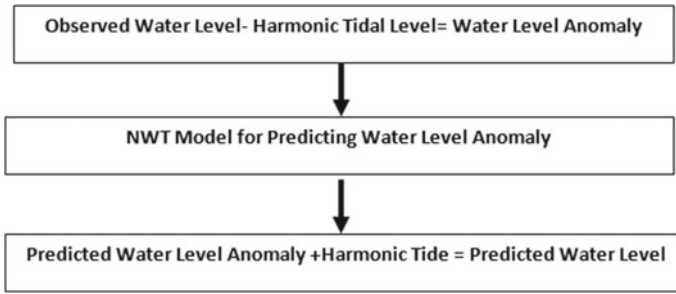
Fig. 4 Flowchart for water-level modeling

transform (DWT) into low (approximate) and high (detail) frequency components as mentioned earlier up to third-level (3d) and fifth-level (5d) decomposition system wherever necessary. This decomposition helps to remove the effect of autocorrelation in the time series analysis to get the outputs. Thus, the network-predicted anomaly was then added to harmonic tidal level to predict the water levels. This methodology is elaborated further in Fig. 4 also.

It was noticed by the authors in earlier attempt of waves using NWT, that higher order Daubechies wavelets gives better results for higher lead time intervals [2, 3], to forecast the water level anomaly from 3 to 6 week ahead Daubechies wavelet 'db35' is used in the present work to develop all the models. Though the available water-level anomaly series is of hourly basis, for the present study one time step of 1 week is used for model development. The model for 3 weeks ahead forecast consists of four values as inputs. The first input is the anomaly for the current time step 't', the second input is the anomaly obtained at 1 week behind the current time step, i.e., '$t - 1w$', and similarly, the third input is the anomaly obtained at 2 weeks behind the current time step, i.e., '$t - 2w$', while fourth input consists the anomaly obtained at 3 weeks behind the current time step, i.e., '$t - 3w$'. Therefore, the four inputs were: $t, t - 1w, t - 2w$, and $t - 3w$ where the output is the 3 weeks ahead water-level anomaly, i.e., '$t + 3w$'. Though the forecasting interval varies from 3 to 6 weeks specifically to judge the performance of NWT, a number of inputs were kept the same while developing all the models. Thus, model for 4 weeks ahead forecast consists the same inputs: $t, t - 1w, t - 2w, t - 3w$, and output will be 4 weeks ahead water-level anomaly value; likewise, 5 weeks ahead and 6 week ahead also include the same inputs, but output will be the 5 and 6 weeks ahead water-level anomaly values. These inputs and outputs of the various models can be expressed as

Input	Output
For 3 weeks ahead model: t, $t-1w$, $t-2w$, $t-3w$	t+3w
For 4 weeks ahead model: t, $t-1w$, $t-2w$, $t-3w$	t+4w
For 5 weeks ahead model: t, $t-1w$, $t-2w$, $t-3w$	t+5w
For 6 weeks ahead model: t, $t-1w$, $t-2w$, $t-3w$	t+6w

Instead of taking serially values of water-level anomaly on hourly basis with continuous time steps like 't, $t-1$, $t-2$ h…' as inputs, here the specific time steps as 't, $t-1w$, $t-2w$, $t-3w$' were selected purposely as inputs which indirectly helps to break the autocorrelation effect in removal of phase lag [5, 6]. The models were calibrated with 70% of the total data, and the remaining data is used for the testing. Separate models were developed for 3w, 4w, 5w, and 6w ahead forecasts. Separate three-layered feedforward networks are developed for both approximate and detailed components of the wave data along with Levenberg–Marquardt (LM) as an algorithm, 'log-sigmoid' and 'linear' as transfer functions. Ultimately, the performance of NWT models was judged by drawing the water-level plots and other error measures like correlation coefficient (CE), root-mean-squared error (RMSE), mean absolute error (MAE), and scatter index (SI) between the observed and corrected water levels (after adding/subtracting the predicted water-level anomalies from the tidal levels).

4 Results and Discussions

As mentioned earlier, all the developed models are tested with the unseen data sets and their performance is judged by the traditional error measures like root-mean-squared error (RMSE), mean absolute error (MAE), coefficient of efficiency (CE), scatter index (SI), and correlation coefficient 'r' along with the water-level plots and scatter plots. Tables 2 and 3 showcase the results and error measures of Portland and Cordova, respectively, where the third-level decomposition (3d) system was used to predict the water-level anomaly from 3 to 6 weeks ahead in time. From Table 2, it is clear that for Portland for all the four time intervals, i.e., from 3 to 6 weeks, 'r' values are above 0.99 which showcases the superiority of NWT over the traditional techniques. Also, range of RMSE varies from 0.0987 to 0.1104 for 3–6w ahead forecasts, whereas the range of MAE varies from 0.0093 to 0.1452 for 3–6 W forecasts, respectively. Higher values of CE and lower values of SI again indicate the good performance of developed NW models. Like Portland, NWT models developed for Cordova also proclaimed better quality results for all the time intervals from 3 to 6 weeks (Table 3). Figures 5 and 6 present the observed and predicted water levels at Portland and Cordova stations, respectively. From these figures, it is evident that

Table 2 Results of Portland

Forecasting interval	Decomposition level	r	RMSE	MAE	CE	SI
3w	3d	0.999	0.0987	0.0093	0.99	0.001
4w	3d	0.998	0.1001	0.0998	0.98	0.099
5w	3d	0.995	0.1098	0.0121	0.978	0.095
6w	3d	0.992	0.1104	0.1452	0.971	0.0910

Table 3 Results of Cordova

Forecasting interval	Decomposition level	r	RMSE	MAE	CE	SI
3w	3d	0.999	0.0977	0.0089	0.99	0.001
4w	3d	0.999	0.0976	0.0092	0.99	0.001
5w	3d	0.998	0.0999	0.0105	0.97	0.009
6w	3d	0.994	0.1120	0.0113	0.96	0.010
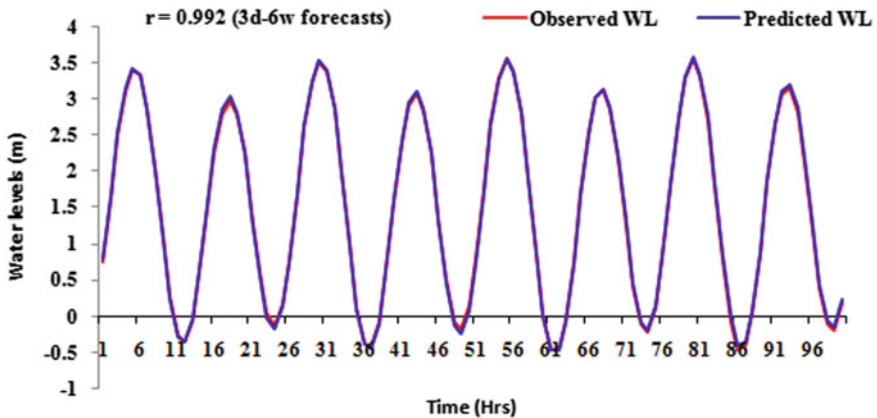
Fig. 5 Six weeks ahead forecast at Portland by third-level decomposition

the phase lag is completely removed in the prediction and due to which prediction accuracy is improvised up to 0.99: r at higher lead time interval also.

But at Dauphin, the results obtained were not in agreement with these two stations. It can be seen from Table 4 that when third-level decomposition (3d) system was used, as the lead time interval increased from 3 to 6 weeks, 'r' values decreased from 0.904 to 0.834, respectively. Therefore, to improvise the results for higher precision, it was decided to decompose the input time series up to fifth-level decomposition (5d) as higher-order decomposition facilitates more number of filters than the lower one and thus helps to remove the autocorrelation effect in the input time series which exactly helps to improve the prediction accuracy [3]. Table 4 also presents the '5d' results for all the four lead time intervals from 3 to 6 weeks, and it can be said from these results that '5d' system had dominated the '3d' system as values of 'r'
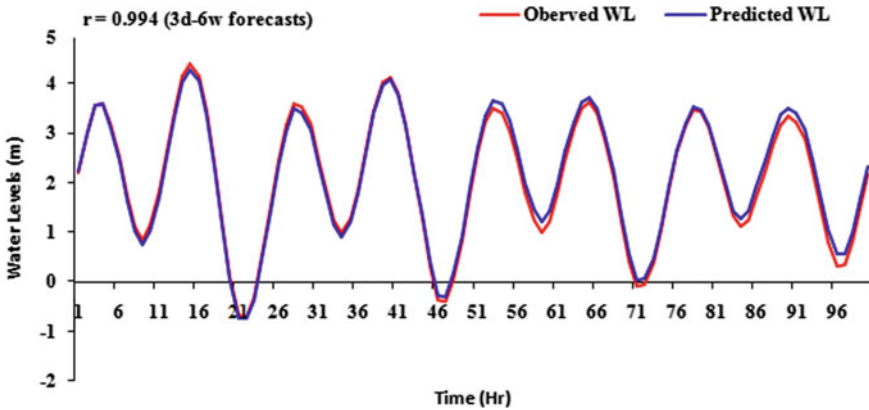
Fig. 6 Six weeks ahead forecast at Cordova by third-level decomposition

Table 4 Results of Dauphin Island

Forecasting interval	Decomposition level	r	RMSE	MAE	CE	SI
3w	3d	0.904	0.104	0.011	0.900	0.87
	5d	0.989	0.080	0.002	0.970	0.36
4w 24 h	3d	0.890	0.115	0.018	0.893	0.92
	5d	0.960	0.086	0.008	0.918	0.45
5w	3d	0.871	0.108	0.021	0.826	1.12
	5d	0.952	0.089	0.011	0.896	0.76
6w	3d	0.834	0.210	0.028	0.743	1.20
	5d	0.911	0.098	0.019	0.820	0.87

improved from 0.904 to 0.989 for 3w, 0.890 to 0.960 for 4w, 0.871 to 0.952 for 5w, and 0.834 to 0.911 (6w) ahead forecasts. As not only 'r' values were improvised but all other error measures values also improved with better range. This indicates the competency of NWT in forecasting the water levels at comparatively longer lead times from 3 weeks (504–1176 h or ¾th of a month) to 6 weeks (1176 h or 1½th of a month). It is evident from all 5d results in Table 4 that the range of SI is decreased from 3 to 5d level. Also, it can be said that as decomposition level increased the model accuracy is increased at higher lead time intervals with reduced values of SI. This authenticated the superiority of higher-level decomposition than the lower ones in precise prediction a time series. Figures 7 and 8 represent the performance of 3 and 5d decomposition models and endorse the superiority of higher-level decomposition on lower-level decomposition system clearly. Also, it is clear from all these figures that phase lag is removed completely from prediction.

These all results depict the excellent performance of the developed NWT models at each station. Though all the three stations Dauphin, Portland, and Cordova were from
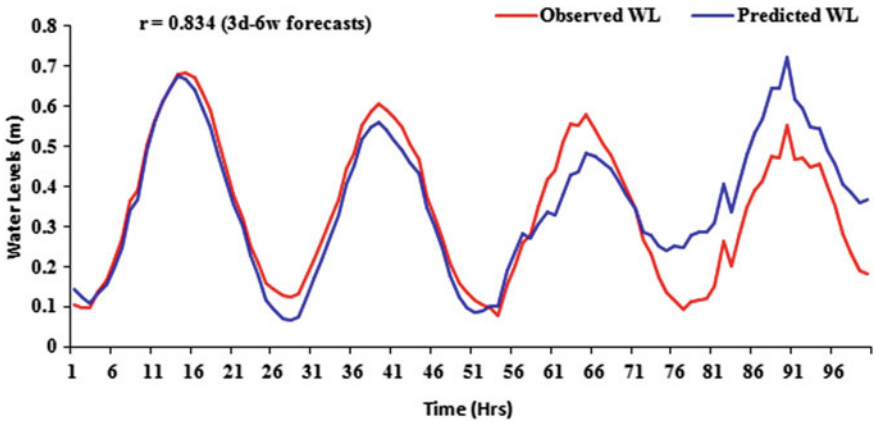
Fig. 7 Six weeks ahead forecast at Dauphin by third-level decomposition
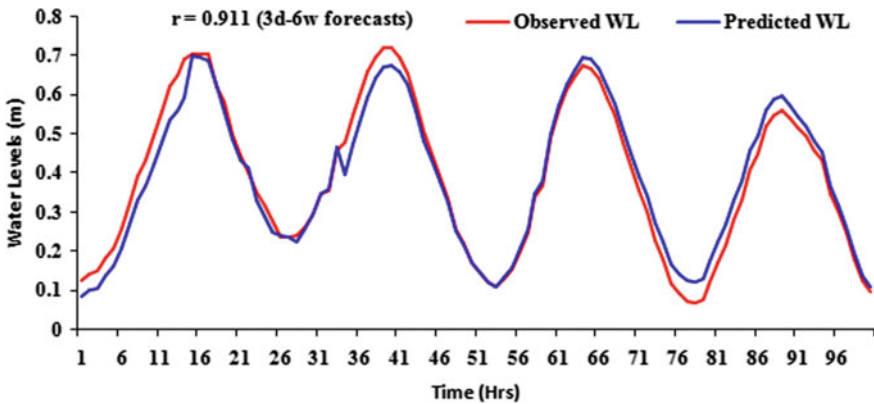

Fig. 8 Six weeks ahead forecast at Dauphin by fifth-level decomposition

different oceanic environments, NWT proved its proficiency at all the stations for all the lead time intervals with 'no phase lag'. Thus, NWT is successful in predicting the accurate water levels at longer lead time as well as with highest precision with all error measures.

Hence, it can be said that the application of NWT is proved to be worthy in the diversified oceanic environments as well.

5 Conclusions

This paper presents the application of neuro-wavelet technique (NWT) for forecasting seawater levels at a different lead time interval from 3 to 6 weeks ahead at three

US coastline stations, namely Dauphin, Cordova, and Portland. The objective of the work is to judge the performance of neuro-wavelet technique in removing the prediction lag water-level forecasting and to improve the prediction accuracy at high lead time intervals in three complete different environments.

Therefore, a combination of wavelet transform and artificial neural networks (neuro-wavelet technique, NWT) when applied for seawater level at three US stations, it is observed that the phase lag in prediction is removed completely. NWT is successful in maintaining the prediction accuracy at higher lead time intervals also. It is confirmed that higher-level decomposition is quite useful for improvising the prediction accuracy at higher lead time intervals. From all the above-mentioned results, it can be said that the performance of neuro-wavelet technique is highly satisfactory in different oceanic environments. As per authors' best knowledge, NWT is applied very first time to predict the seawater level in a view to remove the phase lag and to improve the prediction accuracy as high lead time intervals up to 6 weeks ahead. Taking into account all these facts, it can be said that the application of NWT for accurate prediction of oceanic water levels is pretty useful and can be used in time series predictions successfully as both the major impediments about the 'timing lag' problem and 'prediction' at higher lead can be overcome by neuro-wavelet technique.

References

1. Nitsure SP, Londhe SN, Khare KC (2014) Prediction of sea water levels using wind information and soft computing techniques. Appl Ocean Res 47:344–351
2. Dixit PR, Londhe SN, Dandawate YH (2015) Wave forecasting using neuro-wavelet technique. Int J Ocean Clim Syst 5(4):237–248
3. Dixit PR, Londhe SN, Dandawate YH (2015) Removing prediction lag in wave height forecasting using Neuro-wavelet modeling technique. Ocean Eng 93:74–83
4. Dixit PR, Londhe SN (2016) Prediction of extreme wave heights using neuro wavelet technique. Appl Ocean Eng 58:241–252
5. Abrahart RJ, Heppenstall AJ, See LM (2007) Timing error correction procedure applied to neural network rainfall-runoff modeling. Hydrol Sci J 52(3):414–431
6. De Vos NJ, Rientjest THM (2005) Constraints of artificial neural networks for rainfall-runoff modeling: trade-offs in hydrological state representation and model evaluation. Hydrol Earth Syst Sci (9):111–126
7. Indian National Centre for Ocean Information Services, Hyderabad, India. http://www.incois.gov.in
8. Babovic V, Keijzer M (2000) Genetic programming as a model induction engine. J Hydroinform 1(02):35–60