

Using Gridded Multi-mission Sea Surface Height Data to Estimate Tidal Heights at Columbia River Estuary

Vladimir J. Alarcon(✉)

Civil Engineering Department, Universidad Diego Portales,
441 Ejercito Ave., Santiago, Chile
vladimir.alarcon@udp.cl

Abstract. Measured tidal height time-series are critical for establishing initial and boundary conditions for hydrodynamic models of estuaries. The inexistence of tidal stations in the developing world is more evident than in other parts of the world. The lack of tidal height time-series data oftentimes forces modelers to interpolate or extrapolate these critical data, introducing important uncertainty bounds in the output of hydrodynamic models of estuaries. This paper assesses the feasibility of using gridded multi-mission sea surface height data for estimating tidal heights for the Columbia River estuary located in Northwestern USA. Ocean surface anomaly and geostrophic velocities gridded data, along with historical measured tidal heights are used for training and calibration of the ANN model. A nonlinear autoregressive exogenous neural network is used to predict tidal heights at a 14-min time step. Statistical comparison between measured and predicted data showed that the quality of predicted values was good. Regression analysis for goodness of fit showed the applicability of the proposed method: $R^2 = 0.93$, significance-F = $4.57 * 10^{-18}$, F = 370.80, P-value for the intercept = 0.13, standard error = 0.26 m. Overall, the neural network provides good estimations of tidal heights, considering that it uses coarse-spatial-resolution ocean surface and water velocity data with daily temporal resolution.

Keywords: NARX · Neural network · Tidal heights · Prediction · Gridded multi-mission · Ocean surface topography

1 Introduction

Measured tide data are critical for establishing initial and boundary conditions for small-scale estuary hydrodynamic models. Although coastal waters (estuaries, bays) receive fresh water inputs from rivers draining to the water body, tidal effects are present kilometers upstream from the estuary along the incoming rivers. Hence, the water surface elevations and velocities calculated by hydrodynamic models for estuaries/bays and coastal rivers are also strongly dependent on boundary conditions provided at open-ocean boundaries. However, the lack of tidal stations throughout of the developing world precludes the correct characterization of the hydrodynamics of coastal water bodies. For example, when modelers are put to the task of developing an

© Springer Nature Switzerland AG 2019
S. Misra et al. (Eds.): ICCSA 2019, LNCS 11621, pp. 602–611, 2019.
https://doi.org/10.1007/978-3-030-24302-9_43

estuary or bay hydrodynamic model, interpolation or extrapolation of these critical data is performed, generating undesired uncertainty in the output of hydrodynamic models.

Several researches have explored the use of satellite altimeter data for hydrodynamic modeling. Le Provost et al. [1] assimilated satellite altimetry into a hydrodynamic model (FES94.1) assimilating empirical Topex/Poseidon (T/P) CSR2.0 tidal solution into the hydrodynamic model. Matsumoto et al. [2] assimilated T/P data into the hydrodynamic model NAO99.0 for modeling ocean tides around Japan. Egbert and Erofeeva [3] presented a computationally efficient relocatable system for generalized inverse modeling of barotropic ocean tides which required repeated solution of the forward and adjoint linearized shallow water equations (the same equations that are solved by hydrodynamic models). Foreman et al. [4] assimilated T/P and other satellite altimeter data into the FUNDY5 hydrodynamic code for removing tides from satellite altimeter observations. Han et al. [5] enriched a 3-D hydrodynamic model (Princeton Ocean Model, POM) with multi-mission satellite altimeter data. The abundant literature in this topic shows that assimilation OSTM data into ocean hydrodynamic models is feasible. However, using OSTM data for enriching estuary or bay hydrodynamic models is not as common.

This paper assesses the use of gridded multi-mission ocean surface topography data for estimating tidal heights for the Columbia River estuary located in Northwestern USA. Although the estuary tidal regime is fairly documented through several tidal stations in the area, this study area was chosen precisely because tidal gauge data was necessary to compare the estimations produced in this study. A range of algorithms exist for estimating tidal heights ranging from statistical methods to harmonic analysis. From those available options, artificial neural networks seem to provide opportunities for implementing a system that would combine spatial and temporal data. The novelty of this research is the use of a combination of tidal time-series data with gridded spatial data of ocean surface topography and water velocities.

2 Methods

2.1 Review on the Application of Artificial Neural Networks for Estimating Tidal Heights

The artificial neural network (ANN) approach to prediction of tidal heights is relatively widespread. Qiang et al. [6] developed a forecasting method for tidal heights from observed data and ANN. An artificial neural network was developed for estimating tidal heights from observed tide heights before their current time node. In this research, the neural network was optimized for one year and using multiple stations and artificial neural networks of each observation station is proposed. Wang and Deng [7], proposed a novel remote sensing algorithm for the retrieval of spatially distributed gage height data in nearshore waters using the Moderate Resolution Imaging Spectroradiometer (MODIS Aqua) satellite data and ground truth measurements collected daily from U.S. Geological Survey stations along the Louisiana Gulf Coast. The proposed methodology produced daily, spatially distributed gage height data. Ahn et al. [8] proposed a fuzzy neural network realizing Takagi-Sugeno-Kang inference mechanism employing meta-

cognitive learning algorithm for wave prediction/forecasting in coastal regions. Kaveh et al. [9] used ANN for postprocessing of water level predictions (produced by a hydrodynamic model). A back-propagation artificial neural network technique was applied. The proposed method was used to improve surge predictions of the hydrodynamic model in the region. López et al. [10] present an infra-gravity water wave hindcast methodology through using ANN to construct and analyze the single and the joint probability density functions of two characteristic wave-heights, highlighting the potential of the methodology to characterize the infra-gravity wave conditions inside a port basin and its suitability to study other coastal problems in which these waves are involved. Kim et al. [11] proposed a methodology to estimate damage of breakwater armor blocks incorporating a wave height prediction ANN into a Monte Carlo simulation. The waves predicted by the ANN were comparable to those from a wave transform analysis. Chen et al. [12] explored the use of an artificial neural network (back propagation neural network) and adaptive neuro-fuzzy inference system algorithms to improve incorrect calculations produced by a two-dimensional hydrodynamic model, in predicting storm surge height during typhoon events. Sertel et al. [13] compared five different methods to estimate daily mean sea level heights. These researchers compared: the least squares estimation of sea level model, the multilinear regression (MLR) model, feed forward back propagation ANN (FFBP), radial basis function ANN (RBF), and generalized regression ANN (GRNN). Results of the study illustrated that the ANN and MLR models provided comparatively better results than the conventional model used for estimating sea level, least squares estimation. FFBP, RBF, and MLR algorithms produced significantly better results than the GRNN method, and the best performance was obtained using the FFBP algorithm. El-Rabbany and El-Diasty [14] explored using a neural network-based model for sequentially predicting tidal heights using tide data series collected at various tide gauges. A modular, three-layer feedforward neural network trained using the back-propagation algorithm is used for this purpose. The results of this research showed that the accuracy level of the tidal prediction has improved by a factor of 5 when using the neural network model.

However, in most of the above summarized applications of ANN to tidal heights estimations, the researchers either used data from multiple stations, or produced coarse averaged daily tidal heights, or combined ANN with hydrodynamic models, fuzzy sets, genetic algorithms, introducing complexity and uncertainty to the estimations. None of those researches produced tidal heights estimations at under-an-hour temporal resolution.

2.2 Study Area

The Columbia River Estuary is a water body located on Northwestern USA on the Oregon–Washington border. The Columbia River is the main river feeding the estuary; it is the largest river in the region (longitude: 2000 km approximately) draining a watershed of around 670000 km^2. The NOAA National Ocean Services provides data from several tidal stations in the area. In this research, tidal data from the Astoria station (Fig. 1) were used for comparison to simulated tidal heights.

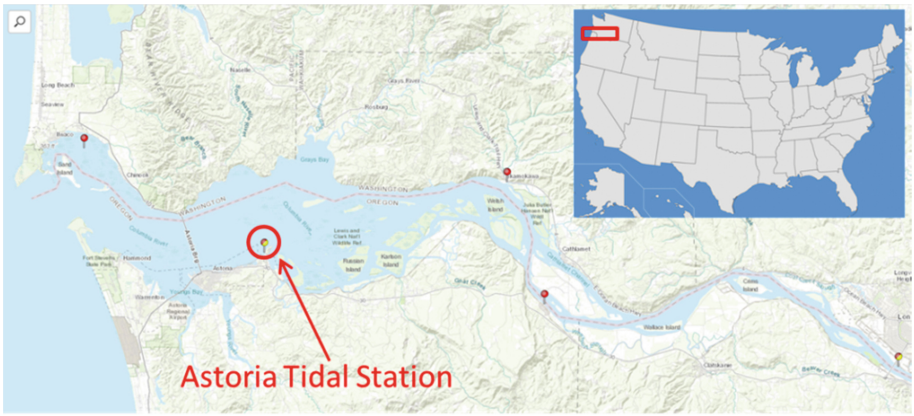

Fig. 1. Study area. Columbia River estuary and tidal stations in the area. Tidal time-series data from Astoria tidal gauge were used in this study.

Tidal time-series measured at Astoria tidal station for the period January 1 through January 31 were downloaded from NOAA data portals. The data provided by NOAA has a temporal resolution of 6 min. The original temporal time-step was resampled to 14 min to avoid long processing times. Results simulated by the model developed in this study were compared to tidal heights reported at Astoria Station during the calibration and validation processes.

2.3 Gridded Multi-mission Ocean Surface Topography Data

The multi-mission gridded altimeter fields with enhanced coastal coverage data product provided by NASA [15] consists of sea surface height anomalies and zonal and meridional geostrophic velocities for the US west coast. The geographical zone currently covered by this product goes from 35.25 to 48.5° latitude North, and from 227.75 to 248.5° longitude East [15]. Platforms and sensors used in the generation of the gridded data were: Jason-1/Poseidon-2, Topex/Poseidon/Poseidon-altimeter, OSTM/Jason-2/ Poseidon-3 [15]. In this research, daily data from January 1 through January 19, 2011, were downloaded from the Jet Propulsion Laboratory website [15]. Figure 2 shows the gridded sea surface height anomalies and water surface velocities for the study area for January 1, 2011. Data are provided as a 0.25-degrees spatial-resolution grid. For brevity, Fig. 2 shows the data in a 1-degree grid.

2.4 Artificial Neural Network

A nonlinear autoregressive exogenous neural network was used in this research. This type of neural network is provided with feedback connections and can use past values of predicted or true (measured) time-series (Boussaada et al. [16]). This key feature makes them applicable for time-series prediction and they are designed for using one or more time-series as input data as well. However, the data that are available for this research are incongruent in temporal resolution. While tidal time-series were resampled

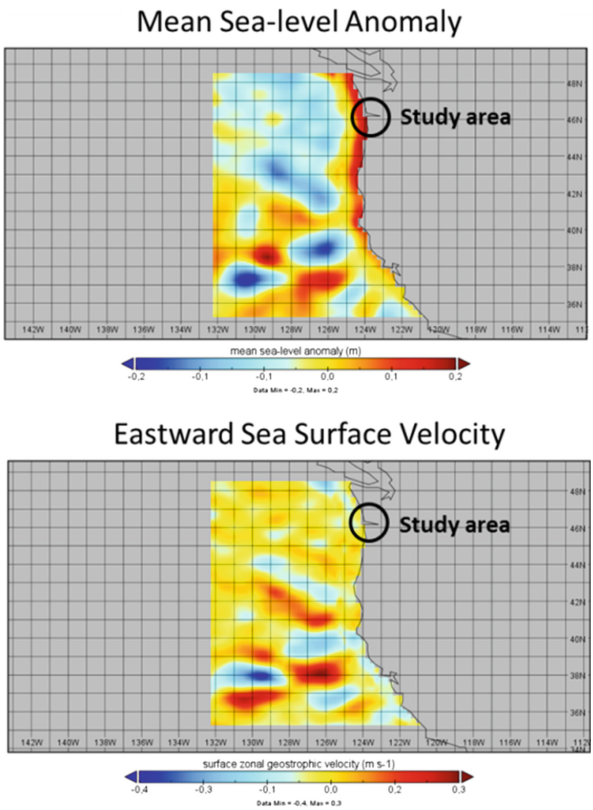

Fig. 2. Mean sea surface height anomaly and zonal and meridional geostrophic eastward velocities for the US west coast. Data are visualized for January 1, 2011. Daily data for the period 1/1/2011–1/19/2011, were downloaded from [15].

to a 14-min time step, multi-mission OST and geostrophic velocities have a 1-day temporal resolution. Clearly, daily gridded data would not be appropriate to model and simulate 14-min tidal time-series. To overcome this hurdle, spatial data in the east direction were used as time-series data. It is hypothesized that surface anomalies and geostrophic velocities far away from the coast are precursors of tides occurring at the estuary hours or days later.

An open-loop architecture is proposed in which the future value of tidal heights at cell i occurring at time $t+n$, TH_i^{t+n}, is predicted from the measured value of ocean surface topography at cell $i-k$, at time t, OST_{i-k}^{t}, eastward geostrophic velocity, v_{i-k}^{t} s, and the measured past values of tidal heights at cell i occurring at time $t-m$, TH_i^{t-m}. Figure 3 presents a conceptual description of the neural network configuration used in this research.

All computations were performed using MATLAB. The Bayesian regularization backpropagation algorithm was used.

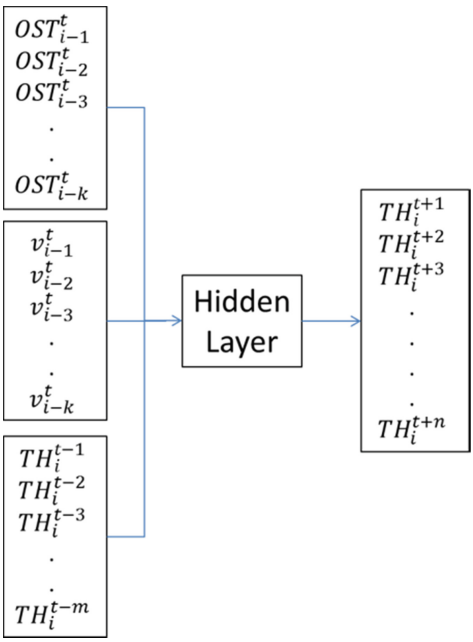

Fig. 3. Conceptual description of the open-loop neural network used in this research. Ocean surface topography (OST) spatial-series, eastward geostrophic water velocity (v) spatial-series, and past tidal height time-series (TH), were fed to the neural network to predict future TH values.

3 Results

3.1 Assessment of NOAA-Predicted Tidal Height Data at Astoria Station

An assessment of measured tidal heights at Astoria station was performed in relation to predicted NOAA-predicted tidal heights. The data provided by the station's website include predictions made with traditional harmonic analysis. Figure 4 shows results of the comparison. Regression analysis was performed to ascertain whether the predictions were statistically consistent with measured values. The analysis showed that the indicators: F (55798.84), significance-F ($2.87 * 10^{-284}$), and P-values for the intercept ($5.69 * 10^{-30}$) and the independent variable ($2.87 * 10^{-284}$) were well within the statistical requirements for goodness of fit. The standard error was 0.016 m (0.054 ft). As shown in the figure, although those indicators are very good, and the coefficient of determination is high ($R^2 > 0.9$), predicted values overestimate tidal heights consistently for all simulation times.

3.2 Tidal Heights Predicted Using the Neural Network

Sensitivity analysis using number of nodes, number of iterations, feedback delays, forward delays, size of training data (number of spatial grid cells and number of past tidal height values). The objective of the sensitivity analysis was to achieve the most

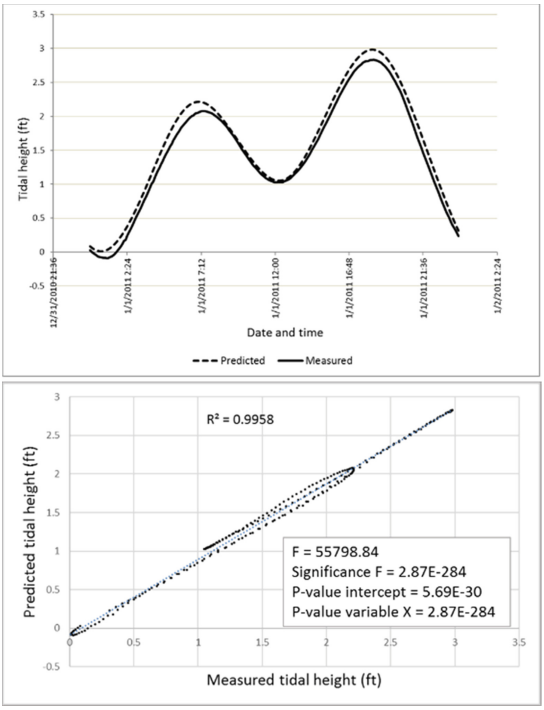

Fig. 4. Assessment of NOAA-predicted tidal height data at Astoria station. Tidal heights predictions through traditional harmonic analysis consistently overestimate measured tidal heights.

appropriate neural network configuration that provided the best fit of predicted tidal heights vs. measured tidal heights. The resulting neural network consisted in an open-loop non-linear autoregressive exogenous neural network, provided with a 2-node hidden layer. The number of iterations that provided good predicted results was 50, with fast input delay (1:1) and fast feedback delay (1:2). The network used Bayesian regularization backpropagation. The training size consisted in 33 spatial grid values for ocean surface topography (OST) and eastward geostrophic velocity (v). These spatial data were fed to the neural network in conjunction with a time-series of past tidal heights measured at Astoria Station, resampled every 14 min (raw data were provided with 6-min interval). Past tidal heights for one day (103 values) were used for training. The neural network was set up to predict 103 values for the next day. Figure 5 shows the results of this process.

Figure 5 shows a comparison of tidal heights predicted by the neural network for a daily event in which the maximum-height tide reached more than 2.7 m (9.2 ft). The neural network predicts tidal heights with a 14-min interval, generating 103 tidal height values per day. The comparison with measured data shows that the statistical fit between predicted and measured values is good. The regression analysis performed with predicted and measured values showed that: significance-F = $4.57 * 10^{-18}$,

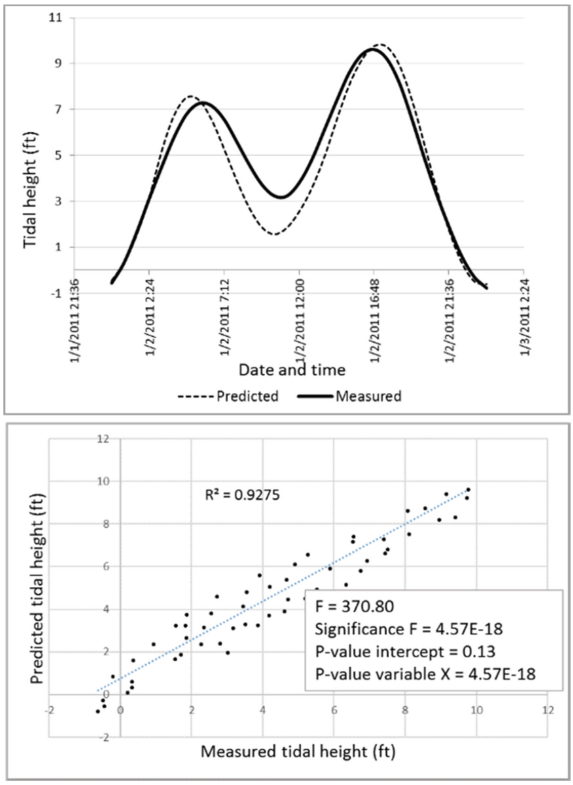

Fig. 5. Tidal heights predicted using the neural network.

F = 370.80, P-value for the intercept = 0.13, and P-value for the independent variable = 4.57 * 10^{-18}. These indicators are within the statistical requirements for goodness of fit. As shown above, the P-value of the intercept is slightly higher than the P-value corresponding to the comparison of tidal heights calculated using harmonic analysis. The standard error calculated for the neural network predicted values was 0.26 m (0.88 ft). It is widely known that the standard error provides the absolute measure of the typical distance that the data points fall from the regression line; in the case of the tidal values calculated by the neural network, they fall 0.26 m (in average) from the line shown in Fig. 5. Tidal peaks and tidal minima are the regions in which the neural network predictions differs more noticeably and are responsible of the increase in the standard error. The predicted values are also slightly biased to the left for hours before noon, and to the right for tidal heights corresponding to times after noon. Overall, the neural network provides good estimations of tidal heights, considering that it uses coarse-spatial-resolution ocean surface topography and velocity data and daily temporal resolution.

4 Conclusions

This prospective computational experiment shows that a nonlinear autoregressive exogenous neural network is able to predict tidal heights occurring at the Columbia River estuary with a 14-min temporal resolution. Most important, the neural network successfully uses gridded spatial data of ocean surface topography and eastward geostrophic water velocities (extracted from multi-mission sea surface height data) along with past time-series tidal height data. Through sensitivity analysis the resulting neural network consisted in an open-loop neural network, consisting of a 2-node hidden layer, with 50 iterations, fast input delay (1:1), and fast feedback delay (1:2). The network used Bayesian regularization backpropagation. The training data consisted in 33 spatial grid-values for ocean surface anomalies and 33 grid-values of eastward geostrophic velocity, along with 103 historical tidal height values. Statistical comparison between measured and predicted data showed that the statistical fit between measured and predicted tidal heights is good. Regression analysis was performed, and the resulting indicators of fit were within the statistical requirements for goodness of fit: $R^2 = 0.93$, significance-F = $4.57 * 10^{-18} < F = 370.80$, P-value for the intercept = 0.13, standard error = 0.26 m. Overall, the neural network provides good estimations of tidal heights, considering that it uses coarse spatial data with daily temporal resolution.

References

1. Le Provost, C., Lyard, F., Molines, J.M., Genco, M.L., Rabilloud, F.: A hydrodynamic ocean tide model improved by assimilating a satellite altimeter-derived data set. J. Geophys. Res. **103**, 5513–5529 (2018)
2. Matsumoto, K., Takanezawa, T., Ooe, M.: Ocean tide models developed by assimilating TOPEX/POSEIDON altimeter data into hydrodynamical model: a global model and a regional model around Japan. J. Oceanogr. **56**(5), 567–581 (2018). https://doi.org/10.1023/A:1011157212596
3. Egbert, G.D., Erofeeva, S.: Efficient inverse modelling of barotropic ocean tides. J. Atmos. Ocean. Technol. **19**, 183–204 (2002). http://volkov.oce.orst.edu/tides/
4. Foreman, M.G.G., Thomson, R.E., Smith, C.L.: Seasonal current simulations for the western continental margin of Vancouver Island. J. Geophys. Res. **105**(C8), 19665–19698 (2000)
5. Han, G., Paturi, S., de Young, B., Yi, Y., Shum, C.-K.: A 3-D data-assimilative tidal model of the Northwest Atlantic. Atmos. Ocean **48**(1), 39–57 (2010). https://doi.org/10.3137/OC303.2010
6. Qiang, L., Bing-Dong, Y., Bi-Guang, H.: Calculation and measurement of tide height for the navigation of ship at high tide using artificial neural network. Pol. Marit. Res. **25**(s3), 99–110 (2018). https://doi.org/10.2478/pomr-2018-0118
7. Wang, J., Deng, Z.: Development of a MODIS data-based algorithm for retrieving gage height in nearshore waters along the Louisiana Gulf Coast. J. Coast. Res. **34**(1), 220–228 (2018). https://doi.org/10.2112/jcoastres-d-16-00161.1
8. Anh, N., Prasad, M., Srikanth, N., Sundaram, S.: Wave forecasting using meta-cognitive interval type-2 fuzzy inference system. Procedia Comput. Sci. **144**, 33–41 (2018). https://doi.org/10.1016/j.procs.2018.10.502

9. Kaveh, N.A., Ghaheri, A., Chegini, V., Nazarali, M.: Application of a hybrid approach for tide-surge modeling in the Persian Gulf. J. Coast. Res. **32**(5), 1126–1134 (2016). https://doi.org/10.2112/JCOASTRES-D-15-00033.1
10. López, M., López, I., Iglesias, G.: Hindcasting long waves in a port: an ANN approach. Coast. Eng. J. **57**(4), Article no. 1550019 (2015). https://doi.org/10.1142/s0578563415500199
11. Kim, D.H., Kim, Y.J., Hur, D.S.: Artificial neural network based breakwater damage estimation considering tidal level variation. Ocean Eng. **87**, 185–190 (2014). https://doi.org/10.1016/j.oceaneng.2014.06.001
12. Chen, W.-B., Liu, W.-C., Hsu, M.-H.: Predicting typhoon-induced storm surge tide with a two-dimensional hydrodynamic model and artificial neural network model. Nat. Hazards Earth Syst. Sci. **12**(12), 3799–3809 (2012). https://doi.org/10.5194/nhess-12-3799-2012
13. Sertel, E., Cigizoglu, H.K., Sanli, D.U.: Estimating daily mean sea level heights using artificial neural networks. J. Coast. Res. **24**(3), 727–734 (2008). https://doi.org/10.2112/06-742.1
14. El-Rabbany, A., El-Diasty, M.: A new approach to sequential tidal prediction. J. Navig. **56**(2), 305–314 (2003). https://doi.org/10.1017/S0373463303002285
15. NASA Jet Propulsion Laboratory: Physical Oceanography Distributed Active Archive Center (PODAAC). NASA EOSDIS PO.DAAC, Pasadena, CA (2015). https://podaac.jpl.nasa.gov/
16. Boussaada, Z., Curea, O., Ahmed, R., Camblong, H., Najiba, M.B.: A nonlinear autoregressive exogenous (NARX) neural network model for the prediction of the daily direct solar radiation. Energies **11**, 620 (2018). https://doi.org/10.3390/en11030620