Global Journal of
Computer Sciences:
Theory and Research

Volume 9, Issue 1, (2019) 010-020

www.gjcs.eu

Prediction of geostrophic currents using big weather data archive and neural networks for the Aegean Sea

Timur Inan*, Computer Programming Programme, Vocational High School, Istanbul Arel University, Buyukcekmece, Turkey
Ahmet Fevzi Baba, Electrics-Electronics Engineering, Technology Faculty, Marmara University, Istanbul, Turkey

Suggested Citation:
Inan, T. & Baba, A. F. (2019). Prediction of geostrophic currents using big weather data archive and neural networks for the Aegean Sea. *Global Journal of Computer Sciences: Theory and Research*. *9*(1), 010–020.

Received September 11, 2018; revised January 13, 2019; accepted March 15, 2019.
Selection and peer review under responsibility of Prof. Dr. Dogan Ibrahim, Near East University, Cyprus.

Abstract

Prediction of sea and weather environment variables like wind speed, wind direction, wave height, wave direction, sea surface current direction and magnitude has always been an important subject in marine engineering as they effect on ship speed and effect the time of arrival to destination point as well. In this study, we propose a neural network that can predict the latitudinal and longitudinal components of sea surface currents in the Aegean Sea. The system can predict the sea surface currents components using the wind components which are gathered from the INMARSAT weather report system. The neural network is trained using the historical data which is gathered from UCAR historical weather database and historical surface current data which is gathered from IFREMER database.

Keywords: Sea surface current, weather report, prediction, neural network, big data archive.

* ADDRESS FOR CORRESPONDENCE: **Timur Inan**, Computer Programming Programme, Vocational High School, Istanbul Arel University, Buyukcekmece, Turkey
E-mail address: timurinan@arel.edu.tr / Tel.: +850-8502735

1. Introduction

For the voyage efficiency of the ships, knowing the sea surface currents can be helpful, because, sea current effects the ship's speed. This effect can be little for the short voyages but can be greater for the long voyages. From past to now, the captains used their experience or sea current charts to predict the effect of the sea currents to the ships' speed. A geostrophic current is an oceanic flow in which the pressure gradient force is balanced by the Coriolis effect. The direction of geostrophic flow is parallel to the isobars, with the high pressure to the right of the flow in the Northern Hemisphere, and the high pressure to the left in the Southern Hemisphere [1]. For computer calculation of circulation prediction, generally Princeton Ocean Model (POM) [2] have been used and some prediction services for sea circulation estimation like Copernicus weather service [3] and Poseidon system [4] exist. Although the services exist, gathering the predictions from the services needs Internet connection and parsing the files to find the necessary information for the sailing route is hard and also the usage of the circulation model needs programming knowledge. These factors make the services' information hard to use for the sailing vessels. In this study, we propose a prediction system to predict the latitudinal and longitudinal components of sea surface currents for the Aegean Sea. We use historical weather data and historical sea surface data to build an artificial neural network for prediction of sea surface current for sailing ships. The system needs only wind prediction data to make sea current prediction. Wind data are available from INMARSAT system or from NAVTEX equipment for all sailing ships. The proposed system is able to make a good prediction of sea current components receiving the information from INMARSAT or NAVTEX system. In order words, the proposed system can be used on-board to inform the watch officers for the sea current components for the areas to be sailed. The overall scheme of the system can be seen in Figure 1.

2. Literature review

From past to now, there have been many studies which examine the observations or predictions of sea and weather conditions of the Aegean sea by means of wind, current and wave. Horton et al. described a nowcast/forecasty stemf or the entire Mediterranean Sea, designed for real-time forecasts and closely resembling operational numerical weather prediction systems[5]. Kourafalou and Barbopoulos proposed a high resolution simulation for the circulation of the North Aegean Sea[6]. The model for this study is based on the POM. The seasonal characteristics of the circulation in the North Aegean sea are examined. Nittis, from Hellenic Meteorological Research described the POSEIDON system, buoys and the facilities of the buoys including their sensors which are used for real-time observations[7]. Nittis et al. examined the efficiency and functionality of POSEIDON system by means of the data quality. The limitations and forecasting skill of the buoy system is discussed[8]. Olson et al. described a pilot experiment using an array of 45 drifters to explore the circulation in the north and Central Aegean Sea[9]. In the study, 45 drifters are used and the drifters send position and circulation data over the Argos system. Zervakis et al.

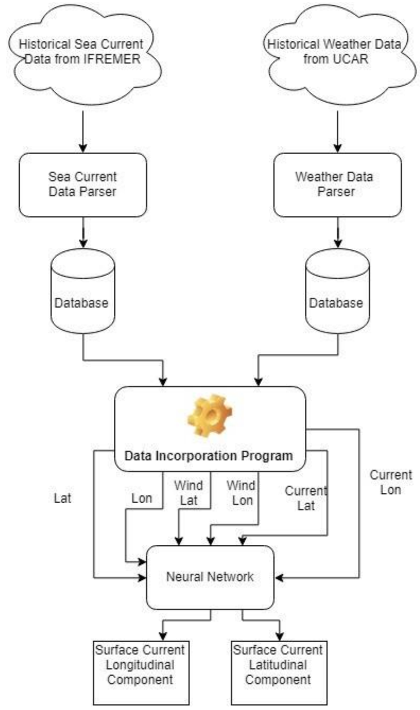

Figure 1. The overall system

Examined the seasonal variability and geostrophic circulation through a series of XBT transects from Piraeus, Greece to Alexandria, Egypt, extending from October 1999 to October 2000 on board Voluntary Observing Ships in the framework of the Mediterranean Forecasting System Pilot project[10]. Oddo et al. proposed a general circulation model for Atlantic-Mediterranean seas for operational forecasting. Also there have been some studies for prediction of sea and weather conditions using various prediction methods[11]. Makarynskyy et al. proposed a system for prediction of sea level variations using neural network at Hillarys Boat Harbour, Western Australia[12]. Altunkaynak[13] proposed a neural network system for prediction of surface water level fluctuations of Lake Van. Hong et al.[14] introduced the dynamic neuro-fuzzy local modeling system that is based on a dynamic TakagiSugeno type fuzzy inference system with on-line and local learning algorithm for complex dynamic hydrological modelling tasks. The system is to forecast the flow of Waikoropupu Springs, located in the Takaka Valley, South Island, New Zealand and the influence of the flow change to the Megawatt Cobb hydropower station on spring flow. Korres et al. proposed a system that uses extended kalman filter to correct the forecast state of a 1/30° POM of the Aegean Sea on a weekly basis. The study is based on the POSEIDON system data[15]. Karimi et al. used Neuro-Fuzzy and Artificial Neural Network techniques for predicting the sea level in Darwin Harbor, Australia[16]. Inan and Baba proposed a set of artificial neural network to predict wind, wave and circulation parameters using the historical data of E1M3A float which is a part of POSEIDON system[17]. Deo et al. proposed a feed forward neural network to predict the wave height in east coast of India[18]. Yasseri et al. proposed an artificial neural network to predict the significant wave height along the US Coast; the network was trained using US National Oceanographic Data Centre Buoy46005 data[19]. Mahjoobi et al. presented alternative hindcast models based on Artificial Neural Networks, Fuzzy Inference System (FIS) and Adaptive-Network-based FIS[20]. The data set used in the study comprised wave and wind data gathered from deep water location in Lake Ontario. Rao and Mandal introduced a system for estimation of wave heights and periods, proposed back propagation neural network with three updated algorithms, namely, Rprop, Quickprop and superSAB[21]. Tur proposed a fuzzy-neural system for prediction of significant wave heights for Filyos area, Turkey[22].

3. Methods

In order to estimate the current direction and velocity, historical wind and current information were used. The historical wind information was gathered from NCAR-UCAR research data archive[23], the International Comprehensive Ocean-Atmosphere Data Set Release 3, Individual Observations data set was used. The data set has the historical sea and weather environment data between the years 1828 and present. The historical geostrophic current data was gathered from Glob Current-Data Catalogue website[24].

The historical current data has geostrophical current data between 2000 and 2012 years. The wind and current data are recorded daily. The proposed system is trained using the wind data from the database which is gathered from NCAR-UCAR and the current data from the database which is gathered from Glob Current-Data Catalogue. A neural network is trained using these data to make a prediction of circulation in desired latitude and longitude. The coordinates are gathered from historical current data database.

3.1. NCAR-UCAR research data archive and data parser

The gathered data archive has .csv extension. We turned it into .xlsx file format. We parsed the data to get the latitude, longitude, date, wind direction and magnitude data. The result file has many columns regarding to the International Maritime Meteorological Archive (IMMA) Format. The parsed data regarding to the IMMA format can be seen in Table 1. After

Table 1. Gathered data format

No.	Length	Abbr.	Element description	Scaled min.	Scaled max.	Units code
1	4	YR	Year UTC	1,600	2024	(AAAA)
2	2	MO	Month UTC	1	12	(MM)
3	2	DY	Day UTC	1	31	(YY)
4	5	LAT	Latitude	90.00	90.00	0.01N
5	6	LON	Longitude	179.99	359.99	0.01E
6	3	D	Wind direction	1	360	
7	3	W	Wind speed	0	99.9	0.1 m/s

parsing the data to an excel file, we turned into a mat file to use in MATLAB application. The detail of IMMA format and detail of the data can be seen in[25]. We turned wind direction and wind speed into 'northward component of wind' and 'eastward component of wind' using Eqs. (1) and (2).

$$w_e = w_v x\cos(\mathrm{rad}(w_d)) \tag{1}$$

$$w_n = w_v x\sin(\mathrm{rad}(w_d)) \tag{2}$$

Where w_e denotes eastward component of wind velocity, w_n denotes northward component of wind velocity, w_v denotes the wind velocity, w_d, *denotes* wind direction in angles. The result data has 591158 rows and 7 columns.

3.2. IFREMER-GlobCurrent data and data parser

GlobCurrent data can be gathered freely from IFREMER website. The data contain daily geostrophic current data in resolution 0.1° between the years 2002 and 2014. The file format is NetCDF which is also known as unidata network common data form files. The eastward and northward current velocity data were extracted from .nc file to .mat file using MATLAB functions.

Table 2. netCDF file components

netCDFFile contents	Description	Units
Eastward geostrophic current velocity	Estimated northward component of the geostrophic current vector.	m/s
Northward geostrophic current velocity	Estimated eastward component of the geostrophic current vector.	m/s

The data format can be seen in Table 2. The detailed information about the data and the data format can be found in[26]. The current data for each day between 2002 and 2014 have been transformed into two matrices. First matrix has northward component of current velocity and the second one has the eastward component of current velocity. Both matrices have 67 rows and 67 columns. In Table 3, a part of a matrix which has daily current data can be seen. The first column on the left has the longitude data and

Table 3. Example of sea circulation data matrix

	34.06	34.18	34.31	34.43
19.93	0.015	0.09	0.07	−0.01
20.06	0.05	0.06	0.03	0.01
20.18	0.06	0.04	0.02	0.02
20.31	0.03	0.04	0.01	0.02

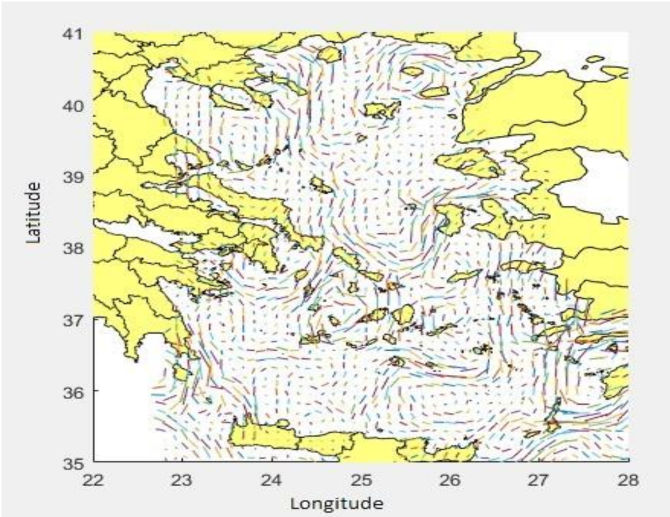

Figure 2. An example of circulation on the Aegean Sea

the first row has the latitude data. In the cells formed by the combination of rows and columns, there are current data in m/s. In Figure 2, an example of circulation on the Aegean sea can be seen. The example is for the day 01/01/2002.

3.3. Inmarsat weather report system

While incorporating the wind and current data, we used Inmarsat Weather Report System meteorological area parts for METAREA-III. Meteorological areas can be seen in Figure 3[27]. The forecast are codes and names can be seen in Table 4. The forecast areas had to be used because location of wind data and current data are not exactly the same locations. To create a relationship between the locations, we searched if they were in the same forecast area.

Table 4. Names and codes of forecast areas for the Aegean Sea

Area name	Area code	Area name	Area code
Rodhos Sea	38	South Evvoikos	46
Karpathio	39	Kefireas Strait	47
West Cretan	40	Central Aegean	48
East Cretan	41	NW Aegean	49
SW Aegean	42	NE Aegean	50
SE Aegean	43	Thrakico	51
Samos Sea	44	Thermaicos	52
Saronikos	45		

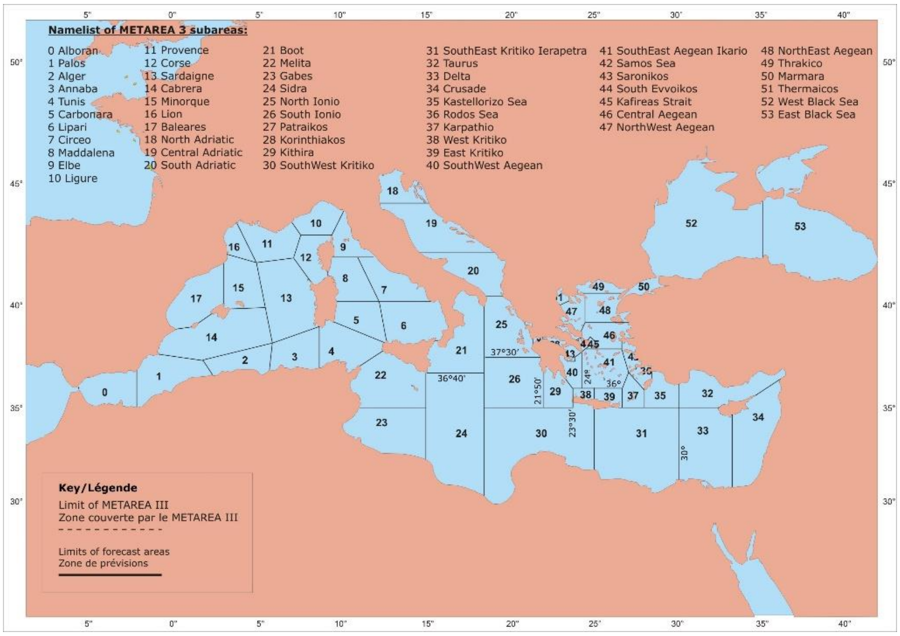

Figure 3. METAREA-III forecast areas

3.4. Data incorporation program

The data gathered from UCAR-NCAR and from GlobCurrent database had to be incorporated to train the artifical neural network. The wind velocity and direction information had to be taken from NCAR-UCAR data; the northward and eastward velocity of geostrophic current had to be taken from IFREMER GlobCurrent data. The locations of both data are associated using Inmarsat weather forecast area system. The data were incorporated using Algorithm 1.

Algorithm 1. Algorithm for incorporation of data

```
get the wind data;
get the boundaries of forecast areas;
get the matrices which contain current data;
for year=2002 to 2014 do
 for month=1 to 12 do
  for day=1 to 31 do
   if current data exists for the date  then
    get ce and cn ;
    for i=2 to 67 do
       for j=2 to 67 do
        lon=cei,1 ;
        lat=ce1,j ;
        if lat and lon value are inside any forecast area boundaries then
          obtain the area code of the location using Table 4;
        for t=1 to 591158  do
            if data exists for the same area code and same date then
              save latitude, longitude, we, wn, cei,j ,cni,j ;
            end if
      end for
      end if
     end for
   end for
end if
end for
end for
end for
```

Here, *ce* denotes eastward current velociyt matrix, *cn* denotes northward current velociyt matrix, w_e denotes eastwards component of wind velocity and w_n denotes northwards component of wind velocity.

3.5. Neural network

The proposed neural network created using M A T L A B neural network toolbox. The data created by the incorporation program has 6 rows and 1025866 rows. Because the data was big, heap memory was low to take the whole data so the data were segmented into pieces for training the network. The network parameters can be seen in Table 5. The graphic presentation of the proposed neural network can be seen in Figure 4.

Table 5. The network parameters

Neuron size	50
Inputs	Latitude, Longitude, Northwards Component of Wind, Eastwards Component of Wind
Outputs	Northwards Component of Current, Eastwards Component of Current
Testing percentage	15
Training percentage	75
Validation percentage	15
Hidden layer function	Tan-sigmoid
Output layer function	Linear
Training algorithm	Levenberg-Marquardt

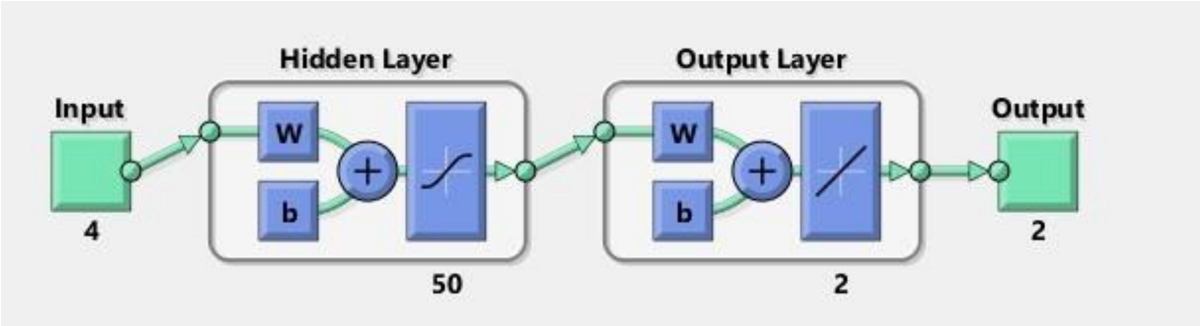

Figure 4. Proposed neural network

4. Results and discussion

The neural network proposed in this study was created using MATLAB environment. The computer that ran the system was a computer with a Windows operating system, an Intel core i7 processor at 3.6 GHz and an 8-GB RAM. The proposed neural network was trained using the wind and current data described in the Methods section. The performance of the proposed neural network was evaluated on two criteria. The criteria were mse (mean squared error) and regression values. The mean squared error of the proposed network was calculated using Eq. (3).

$$\text{mse} = \frac{\sum_{t=1}^{n}(e_t - 0_t)^2}{n} \tag{3}$$

The regression between the estimated data and the observed data are calculated using Eq. (4).

$$R = \frac{\sum_{t=1}^{N}\left(e_t - \overline{e_t}\right) . \left(0_t - \overline{0_t}\right)}{\sqrt{\sum_{t=1}^{N}\left(e_t - \overline{e_t}\right) . \sqrt{\sum_{t=1}^{N}\left(0_t - \overline{0_t}\right)}}} \tag{4}$$

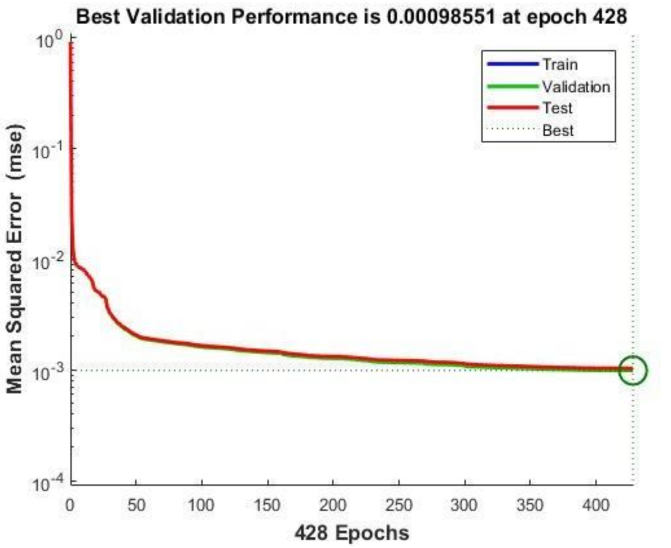

Figure 5. Performance of the proposed neural network

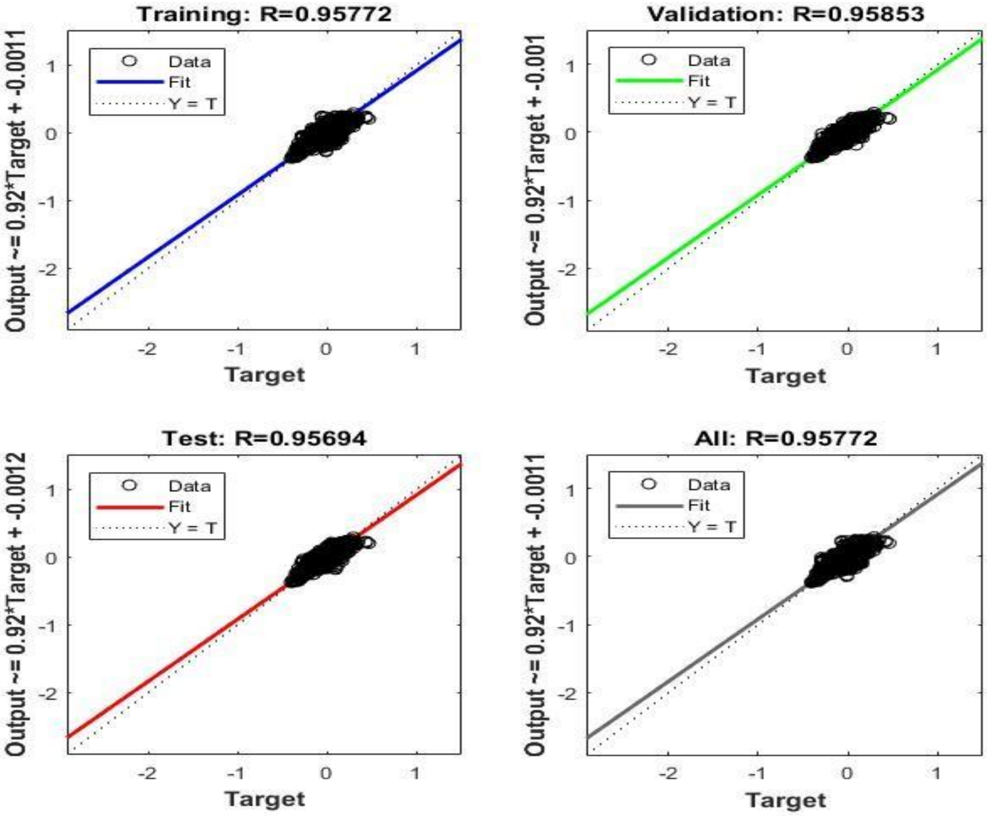

Figure 6. Regression of the proposed neural network

where e_t *denotes* estimated value at t*th step*, o_t *denotes* the observed value at t*th* step, and *t* denotes the row number of the input data that neural network used for training. The mean squared error is found 0.00098551 and regression value for testing 0.95694, for validation 0.95853, for training 0.95772 and for all 0.95772. The graphic presentation of performance and regression values can be seen in Figures 5 and 6.

5. Conclusion

In this paper, we proposed a neural network that can predict geostrophic current on the Aegean Sea when giving the latitude, longitude, wind speed and direction. The neural network was trained using NCAR-UCAR historical database for wind data and IFREMER-GlobCurrent historical database for sea current data. The proposed network can estimate the sea current components of the given coordinate successfully. The mean squared error of the network is calculated as 0.00098551 and regression value of the network is calculated as 0.95722. The performance of any neural system can be determined by these values. A higher regression value and lower mean squared error value show that the system is good at the prediction of desired geostrophic current component values.

References

[1] Wikipedia. (2018). *Wikipedia*. [Cevrimici]. Retrieved November 22, 2018, from, https://en.wikipedia.org/wiki/Geostrophic_current

[2] Princeton. (2018). *The Princeton Ocean Model*. [Cevrimici]. Retrieved November 13, 2018, from, http://www.ccpo.odu.edu/POMWEB/

[3] C. M. E. System. (2018). *Seasonal Forecasts*. [Cevrimici]. Retrieved November 9, 2018, from http://www.copernicus.eu

[4] Poseidon. (2018). *Poseidon System*. [Cevrimici]. Retrieved November 13, 2018, from http://poseidon.hcmr.gr/ocean_forecast.php?area_idaegselectParameter

[5] Horton, C., Clifford, M., Schmitz, J. & Kantha, L. (1997). A real-time oceanographic nowcast/forecast system for the mediterranean sea. *Journal of Geophysical Research: Oceans, 102*(C11), 123–156.

[6] Kourafalou, V. H. & Barbopoulos, K. (2003). High resolution simulations on the North Aegean Sea seasonal circulation. *Annales Geophysicae, 21*(1), 251–265.

[7] Nittis, K. & Perivoliotis, L. (2002). Circulation and hydrological characteristics of the North Aegean Sea: a contribution from real-time buoy measurements. *Mediterranean Marine Science, 3*(1), 21–31.

[8] Nittis, K., Zervakis, V., Perivoliotis, L., Papadapoulos, A. & Chronis, G. (2001). Operational monitoring and forecasting in the Aegean Sea: system limitations and forecasting skill evaluation. *Marine Pollution Bulletion, 43*(7–12), 154–163.

[9] Olson, D. B., Johns, W. E., Veneziani, M., Samuels, G. & Koruafalou, V. H. (2006). Aegean surface circulation from a satellite-tracked drifter array. *Journal of Physical Oceanography, 37*, 1898–1917.

[10] Zervakis, V., Papadoniou, G., Tziavos, C. & Lascara, A. (2003). Seasonal variability and geostrophic circulation in the eastern Mediterranean as revealed through a repeated XBT transect. *Annales Geophysicae, 21*(1), 33–47.

[11] Oddo, P., Adani, M., Pinardi, N., Fratianni, C., Tonani, M. & Pettenuzzo, D. (2009). A nested Atlantic-Mediterranean Sea general circulation model for operational forecasting. *Ocean Science, 5*, 461–473.

[12] Makarynskyy, O., Makarynska, D., Kuhn, M. & Featherstone, W. E. (2004). Predicting sea level variations with artificial neural networks at Hillarys Boat Harbour, Western Australia. *Estuarine, Coastal and Shelf Science, 61*, 351–360.

[13] Altunkaynak, A. (2007). Forecasting surface water level fluctuations of lake van by artificial neural networks. *Water Resource Management, 21*, 399–408.

[14] Hong, Y.-S. T. & White, P. A. (2009). Hydrological modeling using a dynamic neuro-fuzzy system with on-line and local learning algorithm. *Advances in Water Resources, 32*, 110–119.

[15] Korres, G., Nitti, K., Perivoliotis, L., Tsiaras, K., Papadopoulos, A., Triantafyllou, G., . . . Abdullah, K. (2010). Forecasting the Aegean Sea hydrodynamics within the POSEIDON-II operational system. *Journal of Operational Oceanography, 3*(1), 37–49.

[16] Karimi, S., Kisi, O., Shiri, J. & Makarynskyy, O. (2013). Neuro-fuzzy and neural network techniques for forecasting sea level in Darwin Harbor, Australia. *Computers & Geosciences, 52*, 50–59.

[17] Inan, T. & Baba, A. F. (2018). Building artificial neural networks to predict direction and magnitude of wind, current and wave for sailing vessels. *Global Journal of Information Technology: Emerging Technologies, 8*(2), 55–61.

[18] Deo, M. C. & Naidu, S. C. (1999). Real time wave forecasting using neural networks. *Ocean Engineering, 26*, 191–203.

[19] Yasseri, S. F., Bahai, H., Bazargan, H. & Aminzadeh, A. (2010). Prediction of safe sea-state using finite element method and artificial neural networks. *Ocean Engineering, 37*, 200–207.

[20] Mahjoobi, J., Etemad-Shahidi, A. & Kazeminezhad, M. H. (2008). Hindcasting of wave parameters using different soft computing methods. *Applied Ocean Research, 30*, 28–36.

[21] Raoa, S. & Mandal, S. (2005). Hindcasting of storm waves using neural networks. *Ocean Engineering, 32*, 667–684.

[22] Tur, R. (2016). Belirgin Dalga Yuksekliginin (Significant Wave Height=Hs) Bulanık Sinir Aglari (ANFIS) ile Tahmini. *Mehmet Akif Ersoy Universitesi Fen Bilimleri Enstitusu Dergisi, 7*(2), 175–183.

[23] NCAR-UCAR. (2016). *Research data archive*. [Cevrimici]. Retrieved November 13, 2016, from https://rda.ucar.edu/

[24] Ifremer. (2016) *Globcurrent-data catalogue*. [Cevrimici]. Retrieved November 13, 2016, from http://globcurrent.ifremer.fr/products-data/data-catalogue

[25] ICOADS. (2016). *International Comprehensive ocean-atmosphere data set*. [Cevrimici]. Retrieved November 13, 2016, from https://rda.ucar.edu/datasets/ds548.0/docs/R3.0-imma1_short.pdf

[26] Ifremer. (2016). *Glob current-glob current products*. [Cevrimici]. Retrieved November 22, 2016, from http://globcurrent.iferemer.fr/products-data

[27] GMDSS. (2017). *The global martime distress and safety system (gmdss) services*. World Maritime Organisation.