Journal of Soft Computing in Civil Engineering 3-2 (2019) 54-64

Contents lists available at SCCE

Journal of Soft Computing in Civil Engineering

Journal homepage: www.jsoftcivil.com

Machine Learning Method for Predicting the Depth of Shallow Lakes Using Multi-Band Remote Sensing Images

A. Jalilzadeh[1], S. Behzadi[2*]

1. M.Sc. Student in Geographic Information Systems, Department of Civil Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran
2. Assistant Professor in Surveying Engineering, Department of Civil Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran

Corresponding author: *behzadi.saeed@gmail.com*

https://doi.org/10.22115/SCCE.2019.196533.1119

ARTICLE INFO

Article history:
Received: 31 July 2019
Revised: 21 October 2019
Accepted: 25 October 2019

Keywords:
Remote sensing;
Geographic information systems (GIS);
Artificial computation;
Bathymetry.

ABSTRACT

Knowing the lake's characteristics information such as depth is an essential requirement for the water managers; however, conducting a comprehensive bathymetric survey is considered as a difficult task. After the advent of remote sensing, and satellite imagery, it has been recognized that water depth can be estimated in some way over shallow water. There are many models that can evaluate relationships between multi-band images, and depth measurements; however, artificial computation methods can be used as an approximation tool for this issue. They are also considered as fairly simple and practical models to estimate depth in shallow waters. In this paper, different methods of artificial computation are used to calculate the depth of shallow lake, then these methods are compared. The results show that Artificial Neural Network (ANN), Adaptive Neuro Fuzzy Inference System (ANFIS), and regression learner are best methods for this issue with RMSE 0.8, 1.47, and 0.96 respectively.

1. Introduction

The sea surface similar to the earth's surface is very complex with full of hills and valleys [1]. Determination of the water depths, knowing the detailed structure of the bottom and estimating

How to cite this article: Jalilzadeh A, Behzadi S. Machine learning method for predicting the depth of shallow lakes using multi-band remote sensing images. J Soft Comput Civ Eng 2019;3(2):54–64. https://doi.org/10.22115/scce.2019.196533.1119.

the volume of the lakes are essential requirements for the majority of coastal engineering, coastal science applications, managers and decision-makers of water management sector [2–4]. Many attempts have been undertaken by researchers over the past few decades to assembly depth data. These data are used for marine transportation. In addition, water bodies have important defects in environmental studies such as climate change. Therefore, checking sea levels in different time periods provides useful information. Moreover, these data are used to study sea life such as plants and fish which live in water [5,6]. Different methods can be used to illustrate the shape of underwater terrain for better understanding. Bathymetry is a common method to measure the water depth to illustrate the land that lies underwater. Bathymetric maps are the most used data for mapping the seafloor [7]. There are different sources of depth data. These sources can be classified into three different classes. 1) Hydrologic surveys, 2) Echo sounding (sonar) and 3) satellites. In the Hydrologic surveys which is an early technique, a long, heavy pre-measured rope or plummet, is used to measure the depth of the points. However, today it is not applicable due to many errors and limitations such as time-consuming and being expensive [8].

Sonar and LiDAR are two common equipment to measure the depth of seawater and lakes. Both sonar and LiDAR are mounted on the boat; in sonar, a sound wave is used for depth measuring while in the LiDAR sensor, the electromagnetic wave is widely used. Despite being high-precision, the echo sounding method is time-consuming, costly and even is impractical in deep depth. It is suitable for low and medium depths. Satellite Gravity data can also be useful in estimating depth in the whole basin. For example, GRACE mission with twin satellites are making detailed measurements of earth's gravity [9] and SWOT MISSION (Surface Water Ocean Topography) [10] make a global survey of earth’s surface water. In addition, satellite radar altimetry has been used successfully to derive water levels of continental surface water bodies [8,11]. Although bathymetric information is an obvious need in many remote areas, applications of bathymetry mapping in coastal areas are beneficial for a wide range of people and researchers. Because of the limitations, the number of measuring points cannot be too large and cannot cover the whole entire area. So the measured data have general limitations in representing the true terrain of the lake bottom [12]. To overcome such challenges, it is essential to develop methods for depth estimation in uncertain points. Numerous studies have proposed different techniques to estimate terrain of water bodies from measured points, so mapping coastal water from optical remote-sensing techniques has become an interesting method [13].

The theory of using remote sensing techniques for mapping water depth and bottom features was developed by Lyzenga [2] for the first time after that various depth estimation methods are developed and expanded by many researchers based on optical remote sensing. For instance, Bramante, Raju et al. [14], Stumpf, Holderied et al. [15] and Clark, Fay et al. [16] used this method to obtain depth data. In all these researches, they attempted to evaluate the relationships between multi-band images and depth measurements. For instance, Karimi, Bagheri et al [17] developed an equation for evaluating the relationships between Urmia lake depth and Landsat image bands. Obtaining the best equation is not a simple work, so artificial intelligence can be a great help for this purpose. In [3] artificial neural network is used to evaluate water depth in Foca bay lake in Turkey.

For over 40 years, the Landsat mission collects and provides information about earth, and it helps to understand earth better. Landsat sensors receive natural light and thermal radiation from the earth's surface. It can also measure water-leaving reflectance in water-bodies, which can be used to estimate depth. There are several factors that affect the water leaving radiance especially in shallow waters such as depth, the degree of transparency of the water, reflection from the surrounding area, colored dissolved organic matter, nature and material of the bottom and water turbidity. In a study area with a low operating radius, the depth factor can be more effective than other factors [13,18,19].

There are different models and methods in machine learning algorithms including Artificial Neural Network (ANN) and ANFIS and classify method. The purpose of this study is the implementation of these methods and demonstrating their validity, and problems involved with using these methods. The remainder of the paper is structured as follows: Section 2 presents the available data. In section 3, the implementation of Artificial Neural Network, ANFIS and Classify methods are discussed. Section 4 shows the results obtained with simulated and real data, and finally, section 5 draws conclusions and Discussions.

2. Study area

The Caspian Sea, located in the west of Asia, is the world's largest inland water body and bordered by 5 countries including: in the northeast by Kazakhstan, in the southeast by Turkmenistan, in the south by Iran, in the southwest by Azerbaijan, and in the northwest by Russia. The study area is located in Mazandaran province. The geographical location of the study area is 36° 50′ 50” N to 36° 51′ 17” N latitude and 53° 16′ 17” E to 54° 50′ 49” E longitude.

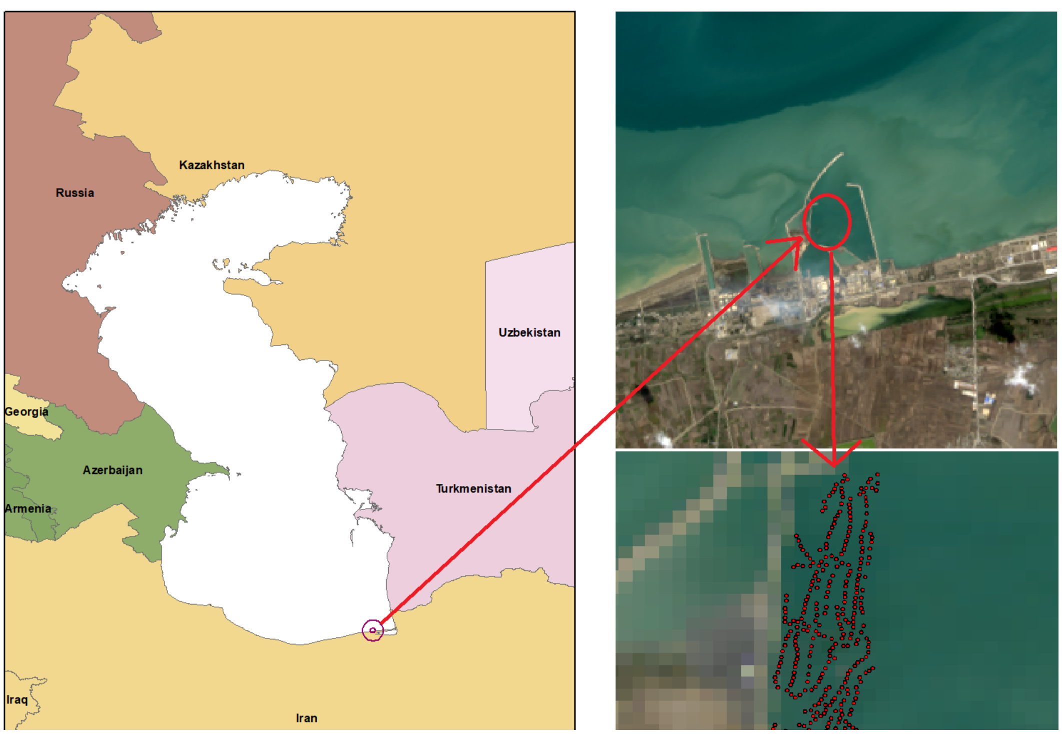

Fig. 1. The Study area and collected points.

3. Dataset

In this research, several points with known depth are needed to extract depth information from satellite images of shallow waters, and also for calibration. The obtained data about the lake's depth is scarce. Considering the available data of the Caspian Lake, the image obtained from the LANDSAT satellite is used. The depth data is gathered with field observation and collected in June 2013. The operation was designed in a way that the satellite's passage time is somewhat the same as the data gathering time. The number of points is 456 and the minimum depth is -2.2800 and the maximum is -10.2100. Landsat satellite provides images with two instruments: the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS). OLI consists of nine spectral bands (Coastal, Blue, Green, Red, Near Infrared (NIR), - Shortwave Infrared (SWIR) 1, Shortwave Infrared (SWIR) 2, Panchromatic, Cirrus) and TIRS consists of two thermal bands. More information about Landsat bands is shown in Table 1.

Table 1
Landsat8 properties.

Landsat 8	Bands	Wavelength (micrometers)	Resolution (meters)
Operational Land Imager (OLI)	Band 1 - coastal	0.435 - 0.451	30
	Band 2 - Blue	0.452 - 0.512	30
	Band 3 - Green	0.533 - 0.590	30
	Band 4 - Red	0.636 - 0.673	30
	Band 5 - Near Infrared (NIR)	0.851 - 0.879	30
	Band 6 - Shortwave Infrared (SWIR) 1	1.566 - 1.651	30
	Band 7 - Shortwave Infrared (SWIR) 2	2.107 - 2.294	30
	Band 8 - Panchromatic	0.503 - 0.676	15
	Band 9 - Cirrus	1.363 - 1.384	30
Thermal Infrared Sensor (TIRS)	Band 10 - Thermal Infrared (TIRS) 1	10.60 - 11.19	100 * (30)
	Band 11 - Thermal Infrared (TIRS) 2	11.50 - 12.51	100 * (30)

4. Methods

In this paper, attempt to estimate the structure of the sea bottom with different methods, then these methods are compared. The used methods are respectively: 1) Artificial Neural Network 2) ANFIS, and 3) Regression Learner. In these methods, input and output data play a main rule in training and testing. Eleven bands of Landsat8 is considered as input data. Band No. 8 is not involved in the model due to difference in pixel size. The surface temperature is also added as input data. The formula for calculating the surface temperature is shown as Equatin.1:

$$TE = (K2 \div log(K1 / R + 1)) - 273.5 \quad (1)$$

Where TE is the temperature in Celsius, and K2 and K1 are constant value for thermal bands (Band 10 and 11).

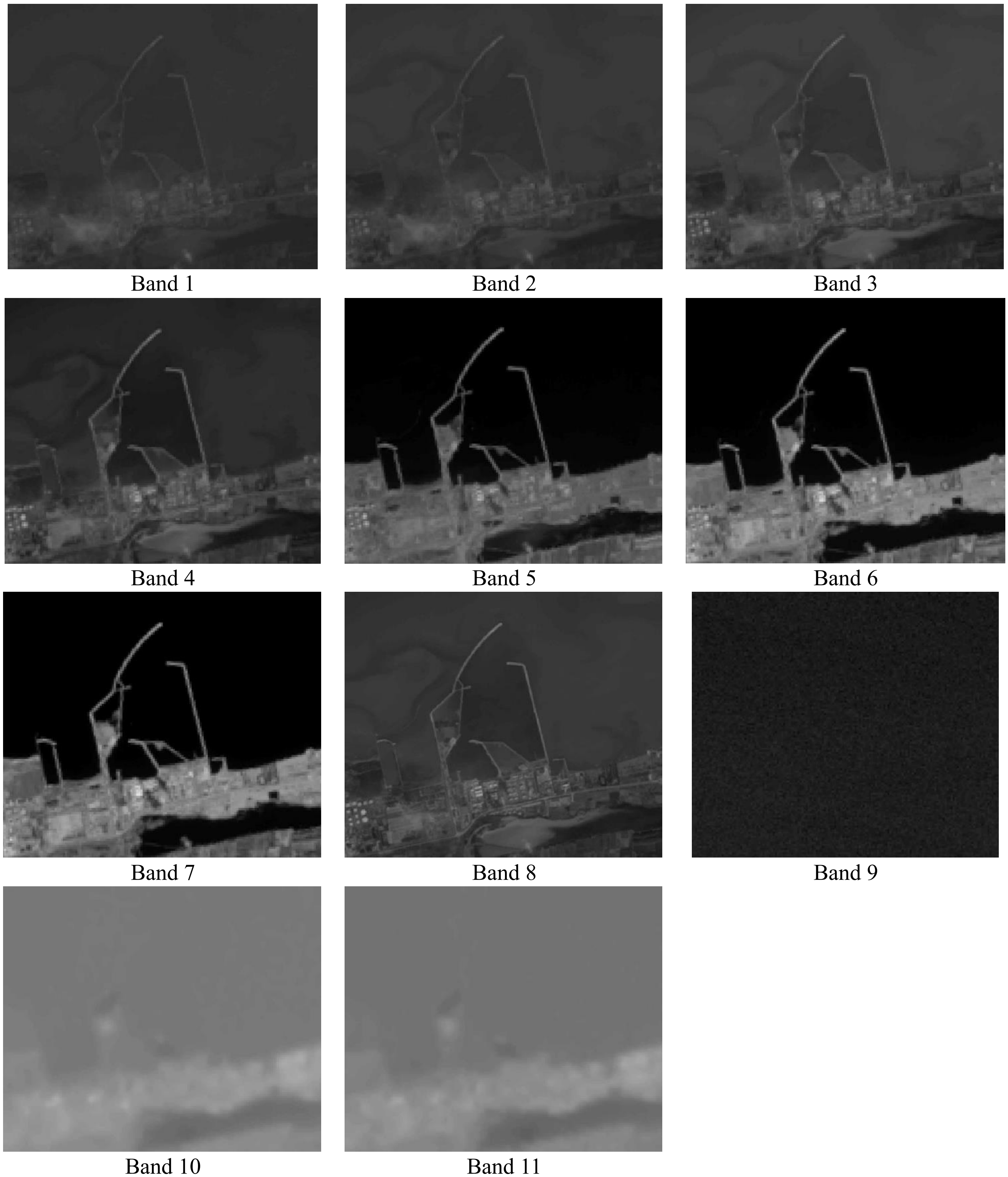

Fig. 2. Image of study area in each band.

The depth points which is collected through field observation is used for training and testing. So, 10 bands from landsat8 and surface temperature is the input data. The input points are the pixels with the same coordinates with collected depth points.

4.1. Neural network

Artificial Neural Network (ANN) as an approximation tool has found a fundamental role in solving problems in various sciences. ANN can arrive at solutions by taking some data samples rather than entire data. ANN has three interconnected layers: Input, hidden and output. The first layer is for input data. The input data include water-leaving radiance in different bands. The hidden layer may consist of numerous layers, and the output data include measured depth points.

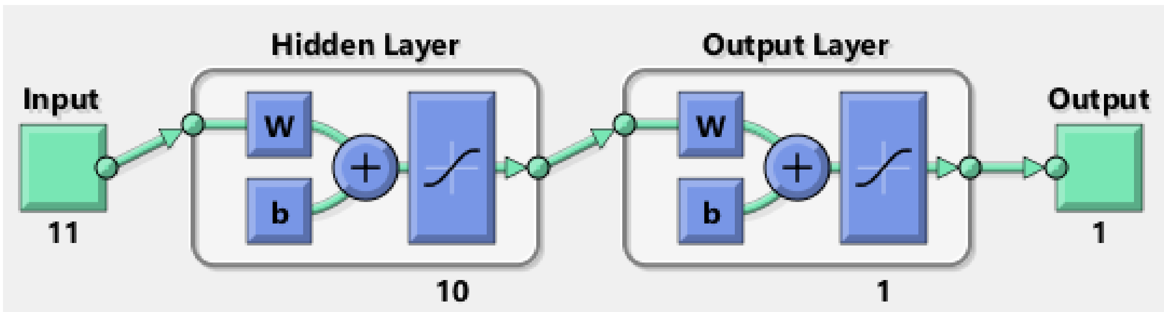

Fig. 3. Structure of used neural network (Taken from Matlab).

In this research, ten bands among eleven bands of Landsat satellite are used (the data of band No. 8 or panchromatic band is not involved in input data because of its different structure). The test areas include 456 points; these data are divided randomly into two groups: training and testing. 70% of the entire data set is utilized for training, and 30% remaining data is used for validating and testing.

ANN is programmed with 1 hidden layer and trained with feed-forward backprop. The result is shown in Table 2.

Table 2
Obtained result from neural network.

Network type	Number of layers	Number of neurons	Iterations	RMSE
feed-forward backprop	1	10	6	0.8

4.2. ANFIS

Fuzzy logic was first advanced by [20] and has come of age its applications have grown in number and variety. The Adaptive neuro-fuzzy inference system or adaptive network-based fuzzy inference system (ANFIS) is a subset of artificial intelligence, and it combines both neural networks and fuzzy logic principles.

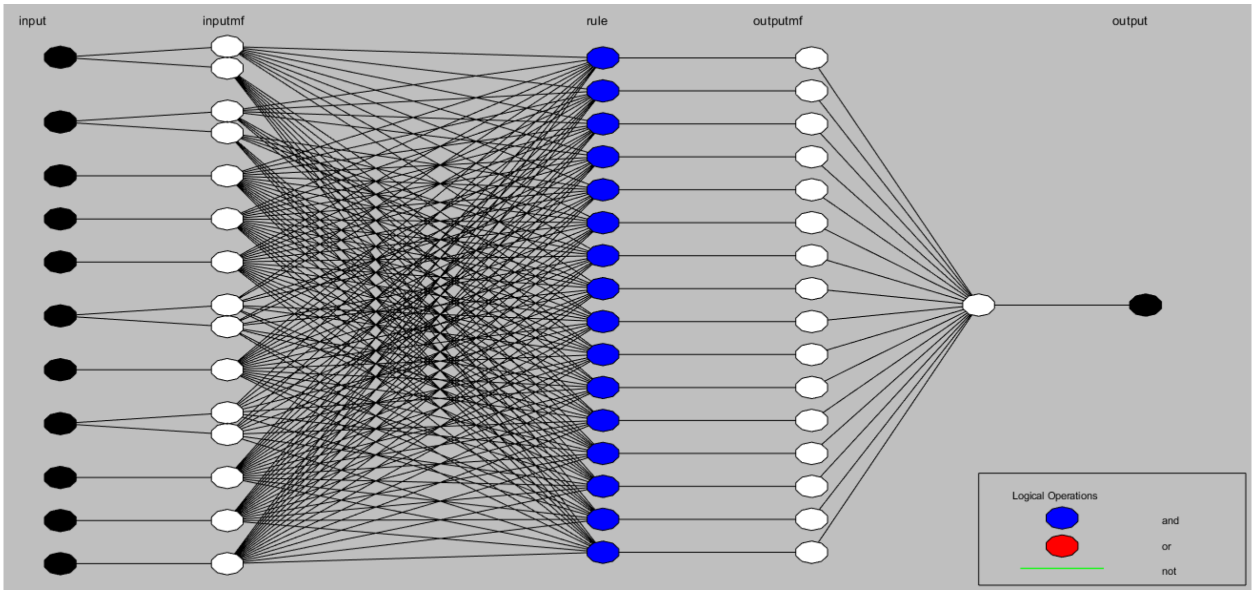

Fig. 4. ANFIS model structure for 2 bands (Taken from Matlab).

ANFIS Model generated with gauss2mf (combining two Gaussian membership functions for computes fuzzy membership values) and 3 MFs (membership function (MF) represents the membership value for each point) to each input. 350 points are used for training and 106 points are also used for testing.

Fuzzy Inference System (FIS) structure is generated from data using grid partition. Input and output data are the same as ANN. The data is divided into two groups: training and testing. 350 points are assigned for training, and 106 points for testing. For input data type of fuzzy Membership Function (MF) Type is Triangular (trimf) and number of MFs is eleven as follow [2 2 1 1 1 2 1 2 1 1 1] and for output MF type is linear. In this paper, optimization method is hybrid for Train FIS. The result of implementing ANFIS on data and some extra information is shown in Table 3. Scatter plots against training and testing data are showed in figure 4 and 5.

Table 3
ANFIS info and Obtained result.

Number of training data pairs	Number of nodes	Number of linear parameters	Number of nonlinear parameters	Total number of parameters	Number of fuzzy rules	Epochs	RMSE
350	76	192	45	237	16	10	1.47

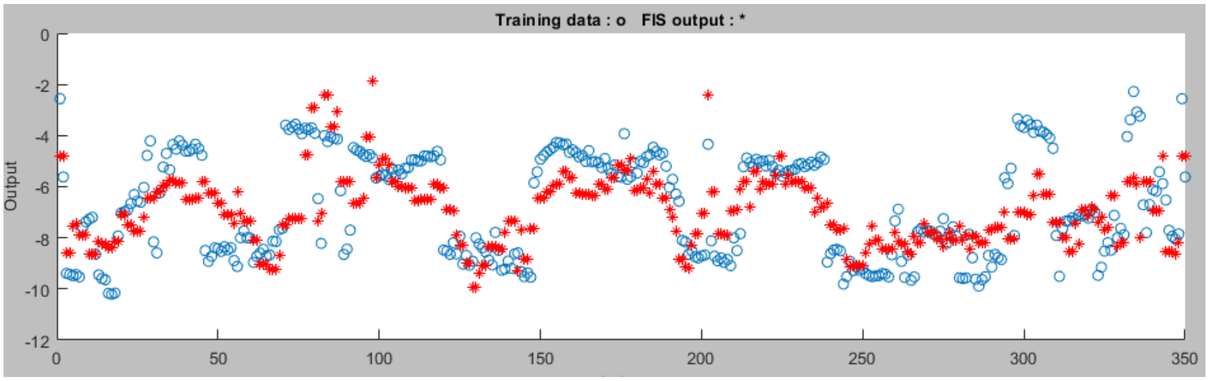

Fig. 5. Plot against training data.

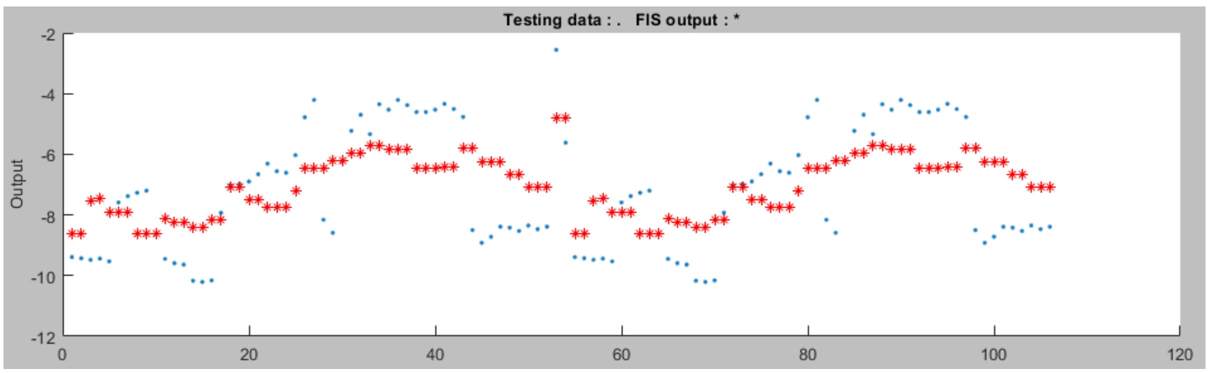

Fig. 6. Plot against training data.

4.3. Regression learner

The Regression Learner app is used to predict data with trained regression models. These models are including linear regression models, regression trees, Gaussian process regression models, support vector machines, and ensembles of regression trees. After training models, their RMSE is compared, which represented in Table 4 to search for the best regression model type. The best answer belongs to Gaussian Process Regression Model and Exponential type with RMSE equal to 0.96.

Table 4

Obtained results from regression learner models.

Linear Regression Models	RMSE
Linear	1.404
Interactions Linear	5.63
Robust Linear	1.43
Stepwise Linear	1.48

Regression Trees	RMSE
Fine Tree	1.05
Medium Tree	1.13
Coarse Tree	1.29

Support Vector Machines	RMSE
Linear SVM	1.62
Quadratic SVM	2.15
Cubic SVM	10.34
Fine Gaussian SVM	1.03
Medium Gaussian SVM	1.28
Coarse Gaussian SVM	1.55

Ensembles of Trees	RMSE
Boosted Trees	1.04
Bagged Trees	1.04

Gaussian Process Regression Models	RMSE
Rational Quadratic	.97
Squared Exponential	.99
Matern 5/2	.98
Exponential	.96

5. Results comparison

As seen, the neural network with RMSE Of 0.8 is the best answer in compare to ANFIS with RMSE equal to 1.45, and regression learning with RMSE Of 0.96. The length of processing time is another issue that can be used to compare these algorithms. The neural network responds with the shortest time and can easily apply and test different structures. Although ANFIS takes the advantages of fuzzy logic and neural network, its processing time is so long and even the hardware of computer cannot respond with increasing the volume of data or complexity of structure. However, the obtained result through ANFIS is not satisfactory. Regression learning responds at a high-speed when only one model is running, but in order to find the best regression model, all models need to be tested, which reduces the processing speed slightly. However, the obtained result and processing time is satisfactory. Table 5 provides a comparison of these models.

Table 5
Comparing used estimation methods.

	Processing time	RMSE	Simplicity	Easily at structure changing
Neural Network	1	1	1	1
ANFIS	3	3	3	2
Regression learning	2	2	2	3

The results show that ANN performs better than the others, so it has been tried to create a bathymetry map and 3D view of the whole study area.

In the study area, there are some pixels which are not water, so these data needs to be filtered. Water Ratio Index (WRI) is selected as a proper index to detect water pixels (Eq.2). in this Equation, the range bigger than 1 is considered as water area [21].

$$WRI = \frac{(T3+T4)}{(T5+T6)} \tag{2}$$

The right image in figure 4 gives information of depth in the study area. In this figure, lighter blue pixels present shallow areas and darker ones are deeper.

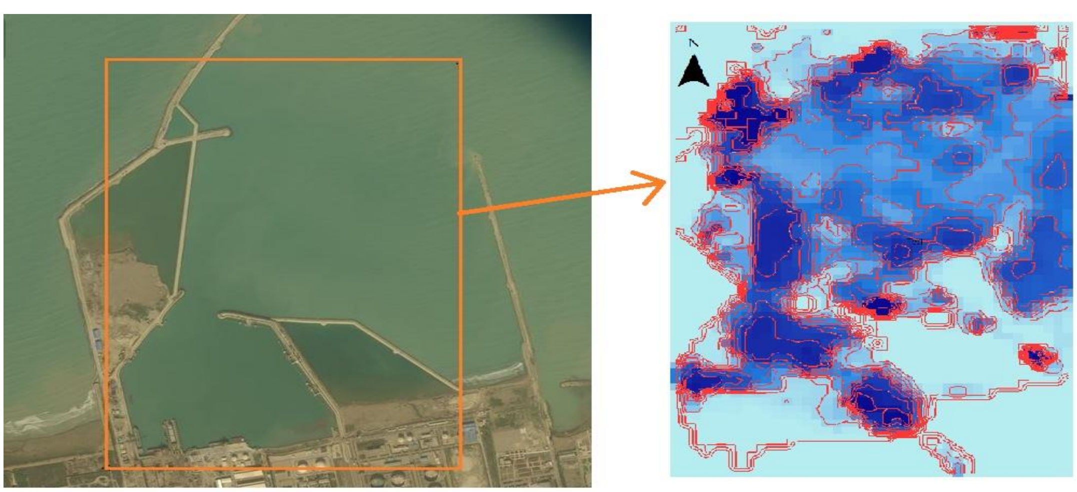

Fig 7. The study area and counters and bathymetry map.

6. Conclusion and discussion

Although the answer is somewhat satisfactory, these responses are greatly improvable. The most important issue in improving the final result is the number of depth data as a trainer to the algorithms. The satellite image as the input data is selected based on available depth data. For example, Landsat satellite is used in this article, but there are other satellites that can be used, such as Sentinel which is free and have a better resolution in comparison with Landsat. Depth data is scarce in literature or available data is considered classified in many countries. Although there is depth information on the sites, most of them are for deep points and this method does not respond at very deep points when the bottom influence becomes negligible. Given that the pixel resolution of the Landsat satellite is 30 meters, so some measured points were in the same pixel, which can partly affect the training and response. Despite all these issues, artificial intelligence has shown that it can overcome these difficulties, and this method can conveniently reduce the labor needed and saves both time and money and provides a fast and practical solution for depth estimation in shallow waters. The main advantage of using this method is it estimates depth without removing the many troublemaker factors that exist.

References

[1] Cazenave A, Altimetry S. Reference module in earth systems and environmental sciences 2018.

[2] Lyzenga DR. Passive remote sensing techniques for mapping water depth and bottom features. Appl Opt 1978;17:379. doi:10.1364/AO.17.000379.

[3] Ceyhun Ö, Yalçın A. Remote sensing of water depths in shallow waters via artificial neural networks. Estuar Coast Shelf Sci 2010;89:89–96.

[4] Elsahabi M, Makboul O, Negm A. Lake Nubia sediment capacity estimation based on satellite remotely sensed detected bathymetry. Procedia Manuf 2018;22:567–74.

[5] Dörnhöfer K, Oppelt N. Remote sensing for lake research and monitoring–Recent advances. Ecol Indic 2016;64:105–22.

[6] https://oceanservice.noaa.gov/facts/bathyuses.html n.d.

[7] https://sealevel.jpl.nasa.gov/overview/oceansurfacetopo/ n.d.

[8] Duan Z, Bastiaanssen WGM. Estimating water volume variations in lakes and reservoirs from four operational satellite altimetry databases and satellite imagery data. Remote Sens Environ 2013;134:403–16.

[9] http://www2.csr.utexas.edu/grace/. n.d.

[10] https://swot.jpl.nasa.gov/home.htm. n.d.

[11] Calmant S, Seyler F, Cretaux JF. Monitoring Continental Surface Waters by Satellite Altimetry. Surv Geophys 2008;29:247–69. doi:10.1007/s10712-008-9051-1.

[12] Li C, Sun B, Jia K, Zhang S, Li W, Shi X, et al. Multi-band remote sensing based retrieval model and 3D analysis of water depth in Hulun Lake, China. Math Comput Model 2013;58:771–81.

[13] Jay S, Guillaume M. A novel maximum likelihood based method for mapping depth and water quality from hyperspectral remote-sensing data. Remote Sens Environ 2014;147:121–32.

[14] Bramante JF, Raju DK, Tsai Min S. Derivation of bathymetry from multispectral imagery in the highly turbid waters of Singapore’s south islands: A comparative study. Digit 8-Band Res Chall 2010 2013:701–10.

[15] Stumpf RP, Holderied K, Sinclair M. Determination of water depth with high-resolution satellite imagery over variable bottom types. Limnol Oceanogr 2003;48:547–56. doi:10.4319/lo.2003.48.1_part_2.0547.

[16] Clark RK, Fay TH, Walker CL. Bathymetry calculations with Landsat 4 TM imagery under a generalized ratio assumption. Appl Opt 1987;26:4036_1. doi:10.1364/AO.26.4036_1.

[17] Karimi N, Bagheri MH, Hooshyaripor F, Farokhnia A, Sheshangosht S. Deriving and Evaluating Bathymetry Maps and Stage Curves for Shallow Lakes Using Remote Sensing Data. Water Resour Manag 2016;30:5003–20. doi:10.1007/s11269-016-1465-9.

[18] Werdell PJ, McKinna LIW, Boss E, Ackleson SG, Craig SE, Gregg WW, et al. An overview of approaches and challenges for retrieving marine inherent optical properties from ocean color remote sensing. Prog Oceanogr 2018;160:186–212.

[19] Traganos D, Reinartz P. Mapping Mediterranean seagrasses with Sentinel-2 imagery. Mar Pollut Bull 2018;134:197–209.

[20] Zadeh LA. Fuzzy logic= computing with words.-IEEE Transactions on Fuzzy Systems, vol. 4,№ 2, pp. 103-111 n.d.

[21] Zhang FF, Zhang B, Li JS, Shen Q, Wu Y, Song Y. Comparative analysis of automatic water identification method based on multispectral remote sensing. Procedia Environ Sci 2011;11:1482–7.