الجامعة المصرية اليابانية للعلوم والتكنولوجيا

E-JUST

Egypt - Japan University of Science and Technology

エジプト日本科学技術大学

ASSESSMENT OF LAND USE LAND COVER CLASSIFICATION AND BATHYMETRY DETECTION OVER SHALLOW COASTAL AND LAKE ZONES USING MACHINE LEARNING ALGORITHMS

by

Hassan Mohammed Hassan Hussein

A Thesis Submitted to the
Graduate School of Environment, Energy, and Chemical &
Petrochemical Engineering,

Egypt-Japan University of Science and Technology (E-JUST)

In Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

in

Environmental Engineering

September 2017

DECLARATION AND CERTIFICATE OF ORIGINALITY

I certify that in the preparation of this thesis, I have observed the provisions of E-JUST Code of Ethics dated 8 September 2013. Further; I certify that this work is free of plagiarism and all materials appearing in this thesis have been properly quoted and attributed.

I certify that all copyrighted material incorporated into this thesis is in compliance with the international copyright law and that I have received written permission from the copyright owners for my use of their work, which is beyond the scope of the law.

I agree to indemnify and save harmless E-JUST from any and all claims that may be asserted or that may arise from any copyright violation.

I hereby certify that the research work in this thesis is my original work and it does not include any copied parts without the appropriate citation.

Hassan Mohamed Hassan Hussein

September 2017

SUMMARY

The thesis wider objective aimed to develop an assessment and monitoring computerized system that uses the satellite images to Land-Use Land-Cover classification (LULC) and bathymetry detection in coastal/Lakes water bodies.

So, the first part of this research assessed a methodology for LULC using multispectral satellite images. The study was executed in a heterogeneous coastal area separated to five classes: sand, building, water, vegetation, and grass-lake-type. This methodology used the Bagging Ensemble (BE) technique with Random Forest (RF) as a base classifier for improving classification performance through decreasing errors and prediction variance. A supervised pixel-based classification method with Principle Component Analysis (PCA) for feature selection from available attributes using a Landsat 8 satellite image was proposed. The used attributes were coastal, visible, near-infrared, short-wave infrared and thermal bands as well as Normalized Difference Vegetation Index (NDVI) besides Normalized Difference Water Index (NDWI). To assess the classification accuracy of BE with RF, it was compared to BE with Support Vector Machine (SVM) and Neural Network (NN) as base classifiers. The results were assessed according to omission, commission errors, and overall accuracy. The revealed results of the proposed methodology using BE with RF outperforms NN and SVM classifiers with 93.3% overall accuracy. The BE with SVM and NN as base classifiers produced 92.6% and 92.1% overall accuracy, respectively. It was confirmed that using BE with RF as a base classifier outperforms SVM and NN as base classifiers.

On the other hand, the second part of this study assessed the performance of three proposed empirical approaches the ensemble regression trees fitting algorithm using bagging (BAG), ensemble regression trees fitting algorithm of least squares boosting (LSB), and support vector regression algorithm (SVR) for bathymetry calculations in four various areas: the shallow coastal area of El-Burullus Inlet, Egypt, which is a turbid sandy bottom area with depths to 6 m; the Alexandria harbor shallow coastal area, Egypt, as an example of a low-turbidity, silt-sand bottom water area with depths ranging from 4 m to 10.5 m; the Lake Nubia entrance zone, Sudan, which is considered a high-turbidity, unsteady, clay bottom area with a depth of 6 m; and Shiraho, Ishigaki Island, Japan, a coral reef area with a depth of 14 m. Landsat 8 and Spot 6 satellite images were used to assess the performance of the proposed models. These proposed models were used to obtain bathymetric maps using the reflectance of green, red bands, blue/red besides green/red band ratios. The bathymetry results of the proposed models were compared with the corresponding results produced from two conventional

empirical methods: the neural network (NN) model and the Lyzenga generalized linear model (GLM). Compared with echosounder field points, BAG, LSB, and SVR results revealed higher accuracy ranges from 0.04 to 0.35 m more than GLM. The BAG algorithm, produced the most correct results.

In addition, the third part of the study suggested RF and Multi-Adaptive Regression Spline (MARS) approaches for bathymetry mapping. Data from Landsat 7, Landsat 8, and Spot 6 satellite images were used to assess the performance of these models. These models were used to obtain bathymetric maps using the same abovementioned inputs. The algorithms were tested over the abovementioned study areas except El-Burullus areas which was replaced with El Nubia entrance zone using Landsat 7 image. The results were compared with the same two conventional empirical methods NN and GLM. When compared with field points, the RF and MARS results outperformed Lyzenga GLM results. Furthermore, the RF approach produced more accurate results with average 0.25 m RMSE enhancements range than the NN model.

Finally, the fourth part of this study proposed BE as a hybrid based approach for bathymetry detection. This approach was applied in two diverse areas with different number of available field points: Alexandria port, Egypt, and a part of Shiraho Island, Japan, with 12.5 m water depths. For NN and RF methods the green and red band logarithms corrected from atmospheric and sun-glint systematic errors of Landsat 8 and Quickbird satellite images were set as input data and water depths as output. The proposed approach ensemble the outputs from NN and RF approaches. To validate the improvement of BE proposed approach, it was compared with NN and RF results. Achieved results were also compared with echosounder water depths field data. From the produced results, around 20 cm and 10 cm improvements in the accuracy of detecting depths over two studied areas, respectively. As a result, it can be concluded that BE ensemble produced more accurate results than using single NN or RF approaches for bathymetry mapping over diverse areas.

ACKNOWLEDGMENT

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS & NOMENCLATURE

LULC	Land Use Land Cover	*Lλ*	Top of atmosphere spectral radiance
BAG	Bagging	*DN*	Digital Number
LSB	Least Square Boosting	*M1*	Multiplicative rescaling factor for radiances
SVR	Support Vector Regression		
RF	Random Forest		
MARS	Multi Adaptive Regression Spline	*A1*	Additive rescaling factor for radiances
BE	Bagging Ensemble		
NN	Neural Network	*Ri'*	De-glinted pixel reflectance value
GLM	Generalized Linear Model		
SVM	Support Vector Machines	*Ri*	Atmospherically corrected reflectance value
PCA	Principal Component Analysis		
NDVI	Normalized Difference Vegetation Indices	*bi*	Regression line slope
		RNIR	Corresponding pixel value in NIR band
NDWI	Normalized Difference Water Indices		
		MinNIR	Min NIR value present in the sample
SWIR	Short Wave InfraRed		
TIR	Thermal InfraRed	*LG*	Computed logarithm of the corrected Green band values
EC	Ensemble Classifiers		
MCF	Multiple Classifiers Fusion	*LR*	Computed logarithm of the corrected Red band values
DS	Dempster-Shafer		
GAB	Gentle AdaBoosting	*B/R*	Blue band divided by red band
LIDAR	Light Detection and Ranging		
SMO	Sequential Minimum Optimisation	*G/R*	Green band divided by red band
		ε	Precision
		C	Compromise between the flatness and the tolerated deviation larger than ε
		ω	Kernel parameter
		σ	Kernel parameter
		ζi	Slack variables

CHAPTER 1

INTRODUCTION

1.1. Land Cover Classification

Land Use Land Cover (LULC) mapping using multispectral images has numerous benefits, with covering large areas, fast acquisition to huge amounts of data, and cheaper costs compared to field methods [1][2]. For correct LULC classification, a proper algorithm is required. Therefore, several researchers have put forth great effort to increase classification accuracy by developing several classification approaches [3]. In recent times, Ensemble Classifiers (EC), or effective Multiple Classifiers Fusion (MCF), have been found to overtake single classifier systems [4]. By exploiting the benefits of diverse classification methods and decreasing their uncorrelated errors by fusing them, the overall accuracy can be enhanced [5].

Dara [5] and Benediktsson et al. [6] demonstrated the various approaches for multiple classification fusion systems. The ensemble classifier can be applied using many techniques, such as bagging, boosting, random forest, majority voting, and the weighted sum of base classifiers. Du et al. [7] and used various combinations of approaches including parallel bagging and sequential boosting classifier systems for classifying hyperspectral data. Salah et al. [1] used Fuzzy Majority Voting and Dempster-Shafer (DS) techniques for combining classification results of three different classifiers using Lidar and aerial images. Chu and Ge [3] used Feature Selection (FS) methods with Genetic Algorithm (GA) and multiple classifiers combination based on Dempster-Shafer Theory of Evidence for classifying land cover features using integration of SAR and satellite imagery. Guan et al. [8] applied RF to automatically select the optimal and uncorrelated features for land-use classification using a combination of Lidar data and ortho-imagery. Samia et al. [4] proposed a better ensemble algorithm depending on the margin theory as a fundamental for the new bagging technique to reduce both the required training data set and the complexity of ensemble approach, thereby enhancing the accuracy.

Zheng [9] used boosting and bagging ensemble techniques with NN as base classifiers and compared it against SVM and logistic regression models for binary prediction with financial time series data. The results show the bagging of NN was superior to SVM and logistics regression models, with a reduction of prediction variance.

Akar and Güngör [10] compared the RF ensemble technique to SVM and gentle adaboosting (GAB) using two different satellite images, Ikonos and Quickbird. The results show RF outperforms SVM and GAB.

Kulkarni and Kelkar [11] applied bagging, boosting, and adaboosting ensemble techniques with backpropagation neural networks with different numbers of hidden neurons for classifying Landsat satellite imagery. These ensembles were compared with single backpropagation neural network and radial basis function network. The achieved results demonstrated the outperforming of ensemble techniques compared to single classifiers. Further, the three ensemble methods gave almost equal results.

1.2. Bathymetry Detection

Coastal and lake shallow areas bathymetry is important for various applications, for instance, spatial monitoring of lakes, sustainable management of natural, and resources coastal engineering sciences [12-13-14-15]. Additionally, sediments erosion and deposition in these shallow areas are rapid because of wave propagation, tidal currents, as well as severe human activities. Consequently, accurate measurements and updated monitoring of these areas, particularly bathymetry, need to be accomplished.

At present, single and multibeam echosounders as well as Lidar represent the conventional approaches for bathymetry measurements. The multibeam echosounder is particularly useful for deep waters with depths up to 500 m due to its higher accuracy and complete bottom coverage. On the other hand, the single-beam echosounder can perform sea bottom maps with suitable vertical accuracy at cheaper cost than the multibeam echosounder [16]. Nevertheless, despite the high level of depth accuracy they can produce, these systems are expensive and difficult to use, particularly in shallow zones where coral reefs, rocks, and general shallowness are an obstacle to the navigation of surveying ships [17]. In recent times, airborne Lidar technology has been developed for bathymetry applications. Still, both systems are expensive, time consuming, laborious, and have relatively low coverage capability.

Optical satellite images represent a time-effective, wide-coverage, and cheaper alternative to conventional techniques for bathymetry applications [16].

In line with the previous literature review, two methods have been used for bathymetry detection: analytical and empirical. The analytical techniques using spectral look- up tables in interpretation of remote sensing data especially in the water depths determination [18]-[19].

2

These methods need the spectral data about the bottom surface reflectance, suspended and dissolved substances and applied with hyperspectral satellite or airborne images [20]. On the other hand, empirical methods rely on the relation between the reflectance of the water bottom surface and water depths in sample positions, which makes these methods simpler.

Lyzenga and Stumpf methods are counted the most broadly used empirical methods for bathymetric data detection [20].

Lyzenga [21] presented the log-linear empirical methodology using single band for determining water depths from satellite or airborne images. The log-linear theory was dependent on get rid of all other reflected values affecting water bottom surface signals. Lately, the log-linear approach was developed to create a correlation between several bands and bathymetric values by Lyzenga et al. [22]. The later approach was performed successively by further studies using diverse satellite images: Quickbird [23], Worldview 2 [24], Spot 4 [16], in addition to Landsat 8 images [15].

Stumpf et al. [25] developed another method dependent on band ratios, as the difference in attenuation between two different spectral bands can be used for determining bathymetry. Lately, this approach has been improved by other scientists for instance Su et al. [26] and Bramante et al. [27].

In addition to these two empirical approaches, a novel alternative approach was created by Ceyhun and Yalçın [14] for bathymetry mapping using the Neural Network (NN) technique. This approach implement a nonlinear relation between satellite image spectral bands and bathymetric values, cope the drawbacks of regressive approaches. Continuously, numerous scientists have confirmed the precedency of NN method to traditional approaches using diverse satellite images, such as, Landsat images [28], IRS P6-LISS III images [29], and Quickbird images [30]. More details about bathymetry detection methods from satellite images can be found in [31].

1.3. General Research Gaps

First for LULC classification, Bagging in other words (bootstrap aggregating) is one of the most prevalent ensemble approaches, designed initially for improving machine learning methods. Decreasing the variance of unstable methods for instance NN, SVM, and Decision trees by averaging diverse algorithms is the main improvement of bagging method.

Subsequently, the results will be better than fitting a single base classifier. Also, this technique decreases the probabilities of over-fitting [33].

Bagging and Random Forest techniques have been broadly used for LCLU classification. This study is probably, to the best of authors' knowledge, the first study for integrating these techniques for classifying multispectral satellite imagery.

Second for Bathymetry detection, analytical approaches have three major demerits. First, the hyperspectral satellite images used by these approaches are not accessible for enormous areas with coarse spatial resolution. Furthermore, the alternative airborne techniques are costly especially for wide coverage areas. Second, the analysis of the hyperspectral images is computationally difficult. Finally, these methods are comparatively complex. As a result, the empirical methods with multispectral imageries considered a valuable alternative [32].

However, Lyzenga, Stumpf, and NN methods have numerous demerits. The Lyzenga approach suppose that the bottom cover surface is entirely homogenous and that the water column is similar in the entire bottom cover area. The Stumpf approach overcomes this drawback; still, it has no physical origin moreover its parameters are calculated by a trial procedure [28]. Finally, the NN method has numerous drawbacks for example it's complex black-box nature; sensitive to slightly minor alteration in the input data values, resulting in high variances in output results; and its weakness to overfit the input bathymetry data.

1.4. Research Objectives

The thesis wider objective is to develop an assessment and monitoring computerized system that uses the satellite images for LULC classification over coastal/Lakes water bodies. In addition, to detect bathymetry over the same coastal/Lakes areas.

In the first part of this study a LULC classification methodology was proposed using BE of RF as a base classifier. This proposed approach reduces the limitations of previous approaches, such as prediction variances, and thus improves the overall accuracy. The methodology was evaluated using Landsat 8 imagery of the El-Burullus Lake in Egypt, and compared with two other previous methods. The criteria used to evaluate the results include commission, omission errors, and the overall accuracy of each classifier.

In addition, this study proposed various empirical approaches for bathymetry detection in shallow coastal or lake areas and endeavors to overcome the disadvantages of the NN and Lyzenga generalized linear model (GLM) methods. These approaches are Bagging (BAG),

least squares boosting (LSB), support vector regression algorithm (SVR), random forest (RF), the multi-adaptive regression spline (MARS), and Bagging Ensemble (BE) of two supervised algorithms. All the proposed algorithms are more stable and more invincible to overfitting than NN, simpler than analytical approaches, and less affected by other environmental factors than Lyzenga GLM. The proposed bathymetry methodologies were applied using various satellite images. The achieved bathymetry results are then evaluated and compared with echosounder bathymetry data over different study areas.

The study aimed to:

1) Proposes improved method of LULC classification for assessment of water bodies detection.

2) Proposes improved methods for bathymetry determination using Multispectral images.

The proposed improved new methods can be used for bathymetry detection on the shallow water bodies including coral reefs areas.

1.5. Thesis Layout

The remainder of this thesis is structured as follows:

Chapter 2: Study Areas and Available Data.

Chapter 3: Methodology.

Chapter 4: Results.

Chapter 5: This chapter contains discussions, summarizing the main findings of the research with comparison to related researches. Additionally, it contains concluding remarks of research.

1.6. List of Publications

From the thesis, the following papers were published:

1. Hassan Mohamed, Abdelazim Negm, Mahmoud Salah, Kazuo Nadaoka and Mohamed Zahran, "Assessment of proposed approaches for bathymetry calculations using multispectral satellite images in shallow coastal/lake areas: a comparison of five models," Arabian Journal of Geosciences, Vol. 10, No. 42, (2017) [33].

2. Hassan Mohamed, Abdelazim Negm, Mohamed Zahran and Oliver C. Saavedra, "Assessment of Ensemble Classifiers using Bagging Technique for Improved Land Cover Classification of

Multispectral Satellite Images,". The International Arab Journal of Information Technology (IAJIT), Vol. 15, No. 3, (online 2017) [34].

3. Hassan Mohamed, Abdelazim Negm, Kazuo Nadaoka, Tarek Abdelaziz and Mohamed Elsahabi, "Comparative study of approaches to bathymetry detection in Nasser/Nubia Lake using multispectral SPOT-6 satellite imagery," Hydrological Research Letters Journal, Vol. 10 (2016) No. 1 p. 45-50 [35].

4. Hassan Mohamed, Abdelazim Negm, Mohamed Zahran and Sommer Abdel-Fattah, "Estimation of Bathymetry Using High-resolution Satellite Imagery: Case Study El-Burullus Lake, Northen Nile Delta," Book Chapter in The Handbook of Environmental Chemistry, Springer Berlin Heidelberg, p.p. 1-30, October 2016 [31].

5. Hassan Mohamed, Abdelazim Negm, Mohamed Zahran and Oliver C. Saavedra, "Bathymetry Determination from High Resolution Satellite Imagery Using Ensemble Learning Algorithms in Shallow Lakes: Case Study El-Burullus Lake," International Journal of Environmental Science and Development, Vol. 7, No. 4, April 2016 [36].

CHAPTER 2

STUDY AREAS AND AVAILABLE DATA

This chapter presents the study areas and available data used for LULC classification and bathymetry detection.

2.1. Study Areas

The study area located at the El-Burullus Lake and its surroundings. It is a coastal heterogeneous area with a diversity of main features as land, buildings, water, vegetation, and lake plants [37]. Hence, it can serve as a suitable test area for LCLU classification. Figure 2-1 illustrates the study area.
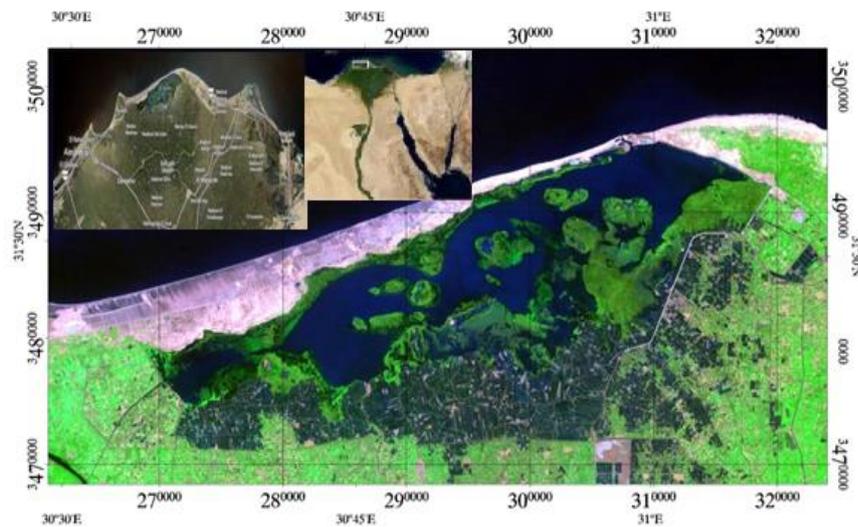
Figure 2-1. El-Burullus Lake, Nile-Delta, Egypt.

In this research, the bathymetry detection was divided to three parts and tested over six various areas with different satellite images. The first part of bathymetry study included four study areas as follow:

The first study area was the El-Burullus Inlet coastal zone, which extends along the northern part of the El-Burullus Lake of Egypt, and has dimensions of 9 km in the east–west direction and 2.5 km in the north–south direction (see Figure 2-2). It is a nearly uniform shallow, turbid water area with depths up to 6 m and high rates of sediment movement and coastal change. Most of the sea bottom is covered by sand [37].

The second study area was Alexandria harbor, Egypt (see Figure 2-3a). It is a properly deep, low turbidity, quiet water area, due to its coastal barriers, and has a depth range of 10.5 m. The port bottom cover surface is silt-sand.

The third study area was the entrance area of Lake Nubia, which is located in Sudan (see Figure 2-3b). It is a properly irregular, shallow, very turbid water area with depths up to 6 m with high amounts of sediment alterations in addition to annual flood changes. The lake has a clay bottom cover surface.

The forth study area was Shiraho which is a subtropical territory, located in the southeastern part of Ishigaki Island, Japan (see Figure 2-4). It is rough shallow and little turbid with depths up to 14 m. Shiraho is varied area with a rich marine biodiversity that consist of numerous ecosystems as seagrass, mangroves, and coral reefs.
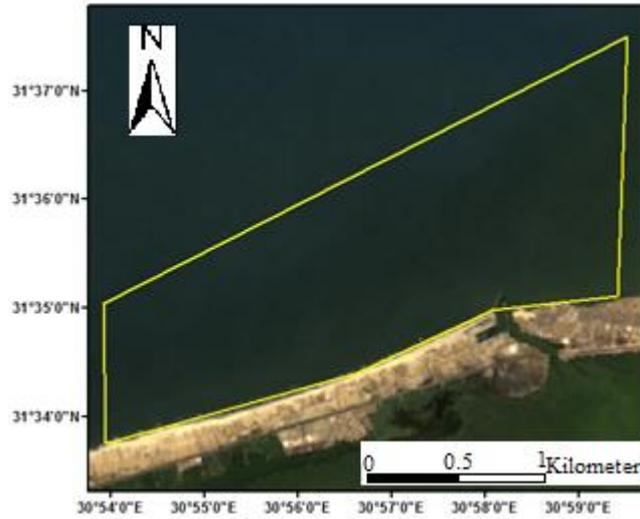
Figure 2-2. The 1st study area of El-Burullus coastal strait area, Egypt.

(a)   (b)
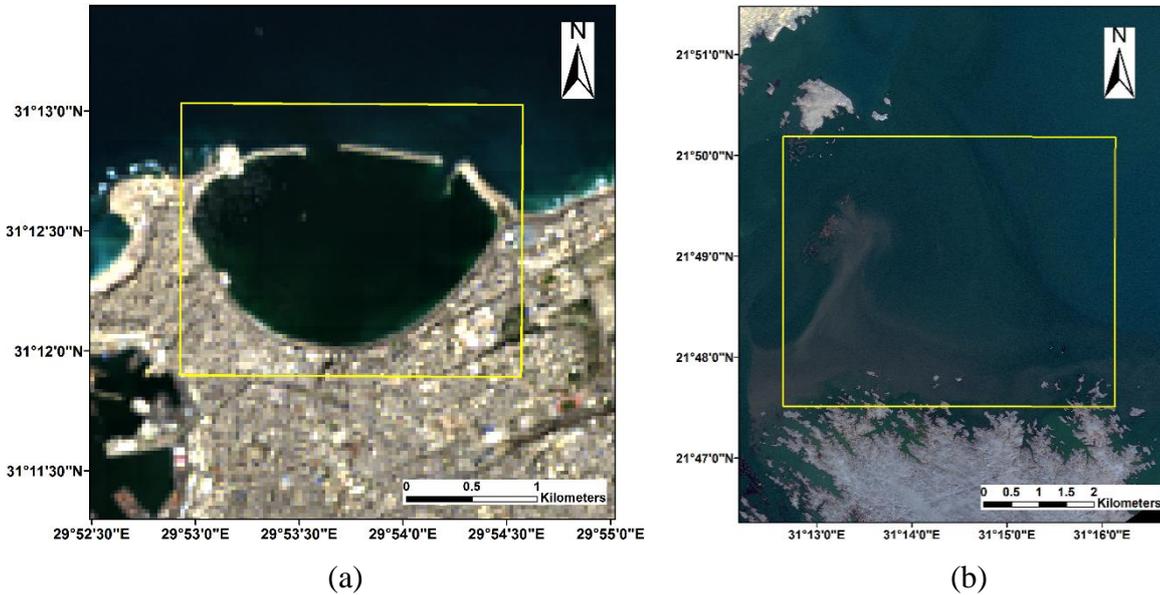
Figure 2-3. (a) The 2nd study area of Alexandria port coastal area, Egypt (b) The 3rd study area of Nubia Lake entrance part, Sudan.

Figure 2-4. The 4th study area of Shiraho, Ishigaki Island, Japan.

The second part of bathymetry study included the last abovementioned three areas of Alexandria harbor, Egypt, the Lake Nubia entrance zone, Sudan, and Shiraho, Ishigaki Island, Japan. In addition, another part from the south entrance of Nubia Lake was added for this study using Landsat 7 image as shown in figure 2-5.
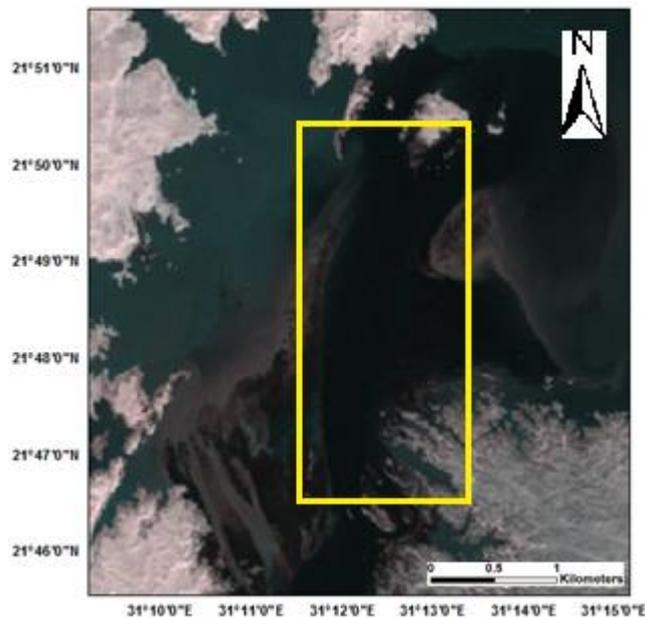
Figure 2-5. The added study area of Nubia Lake entrance zone, Sudan.

Finally, the third part of bathymetry study was performed over two study areas. In adittion to the abovementioned Alexandria harbor area, Egypt, a part of Shiraho, Ishigaki Island, Japan, was added for this study using Quickbird image as shown in figure 2-6.

Figure 2-6. The added study area which is a part of Shiraho, Ishigaki Island, Japan.

2.1.2. Available Satellite Images

A Landsat 8 satellite image with eleven multispectral bands was used for LCLU classification of Lake El Burullus study area. The image was picked up on 14 August 2014 (see Figure 2-7).
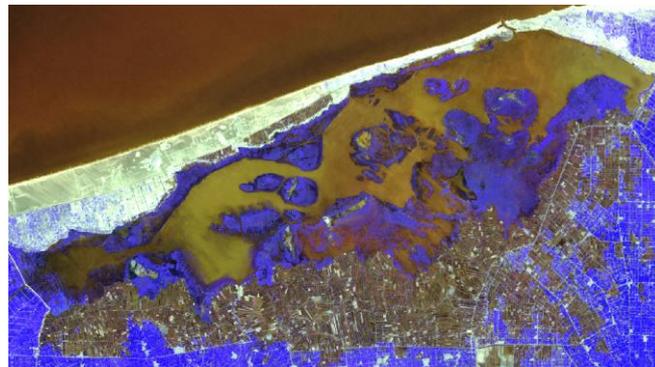
Figure 2-7. Landsat-8 satellite image over the selected study area.

For bathymetry determination Landsat 8 satellite images were used for the first, second and the forth study areas. Also, Spot 6 image with a spatial resolution of 1.5 m was used for the third study area. The necessary parameters for performing radiometric images corrections were included in images metadata files. The first Landsat 8 image was acquired during quiet weather surroundings on 22 March 2014. The second Landsat 8 image was acquired during calm weather environments on 3 August 2014, and the third

10

Landsat 8 image was collected during windy surroundings on 5 June 2013. The Spot 6 image was acquired during quiet weather conditions on 12 January 2014. These four images were selected so as to be synchronised with echosounder field observation times for all the study areas.

Moreover a Landsat 7 satellite image was added for the second part of bathymetry over the additional area of Lake Nubia entrance zone. The Landsat 7 image was collected during windy conditions on 15 December 2009.

Lastly, a Quickbird satellite image was added for the third part of the bathymetry study. Quickbird image have a spatial resolution of 0.6 m and was collected during windy conditions on 20 July 2007.

2.1.3. Echosounder Data

The field reference observed water depths of the first and second study areas used for validating the algorithms were acquired by a NaviSound model 210 Hydrographic Systems echosounder instrument with attached Trimble 2000 GPS. The maximum accessible depth level of this echosounder was 400 m, and its vertical accuracy was 0.01 m at 210 kHz (see Figures 2-8 and 2-9). The third study area observed field water depths were picked up by an Odom Hydrographic Systems Echotrac model DF 3200 MKII echosounder device with built-in DGPS. The depth range of the echosounder was 200 m and its vertical accuracy was 0.01 m ± 0.1% of depth (see Figure 2-10). Finally, the reference water depths of the forth study area were observed by a single beam Lowrance LCX-15MT dual frequency (50/200 kHz) transducer and 12-channel GPS antenna. The vertical accuracy was ± 0.03 m [38] (see Figure 2-11).

Approximately 500 field points were observed for the first study area, 2500 field points were observed for the second study area, 12500 for the third study area, and 14500 for the fourth study area. These points were used for validation and evaluation of all the bathymetric approaches.
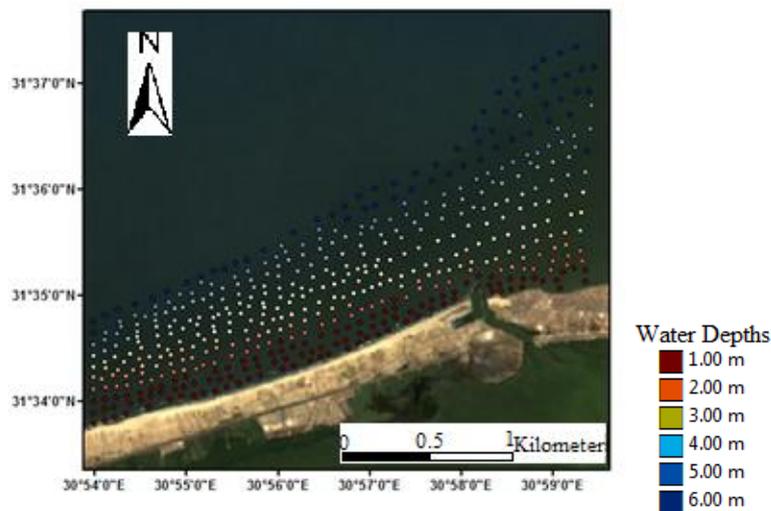
Figure 2-8. Field bathymetry Reference points of 1st study area from echo-sounder.
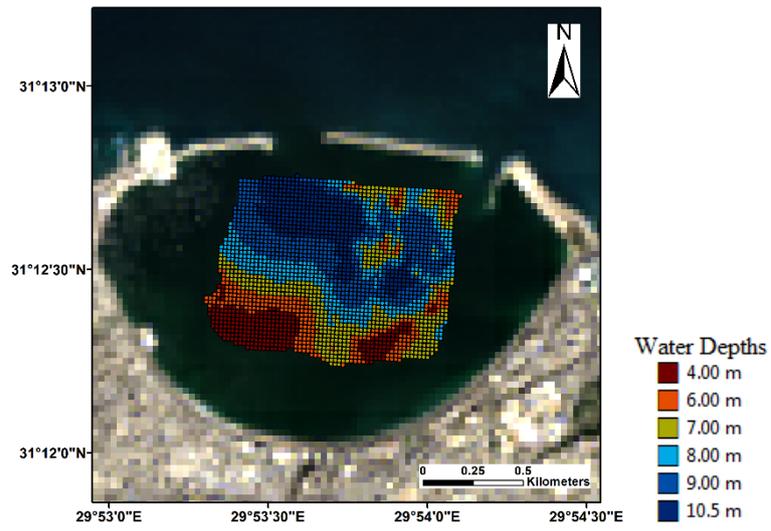
Figure 2-9. Field bathymetry Reference points of the 2nd study area from echo-sounder.
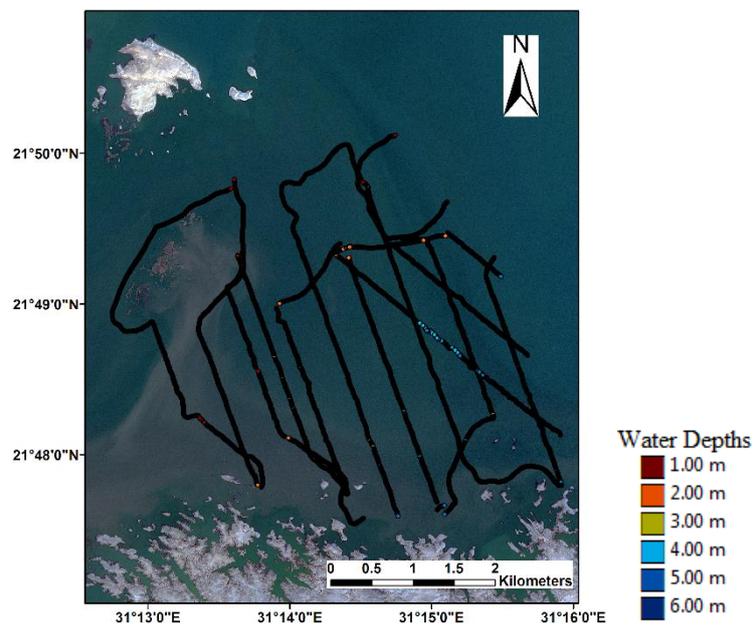
Figure 2-10. In-situ bathymetry Reference points of the 3rd study area from echo-sounder.

Figure 2-11. Field bathymetry Reference points of the 4th study area from echo-sounder.

The additional area of the second part of bathymetry study over Lake Nubia entrance zone have about 4500 observed field water depths. These points were picked up by an Odom Hydrographic Systems Echotrac model DF 3200 MKII echosounder device with built-in DGPS as shown in figure 2-12.
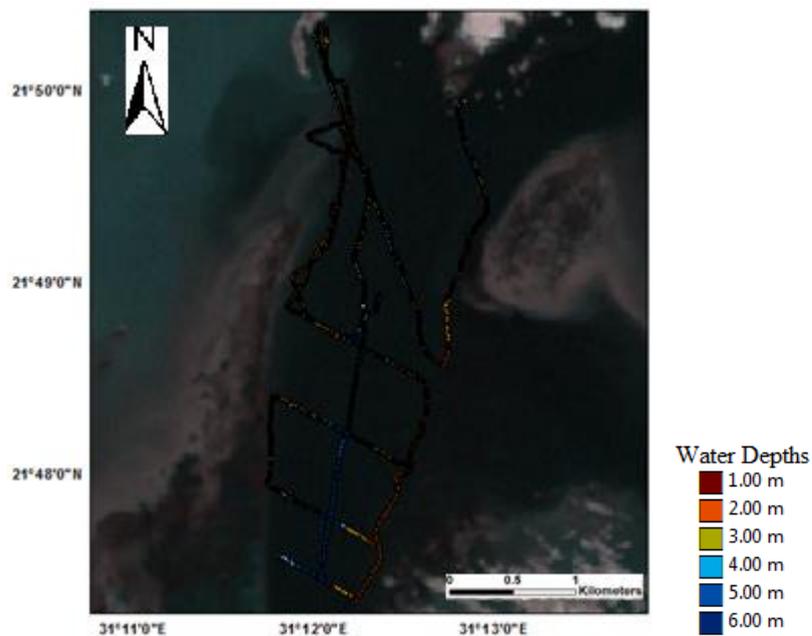
Figure 2-12. Field bathymetry Reference points of the Lake Nubia added study area from echo-sounder.

Finally, the third part of bathymetry study a part of Shiraho, Ishigaki, Japan reef area was added. About 8106 observed field water depths were picked up over this area with a single beam Lowrance LCX-15MT dual frequency (50/200 kHz) transducer with 12-channel GPS antenna (see Figures 2-13).
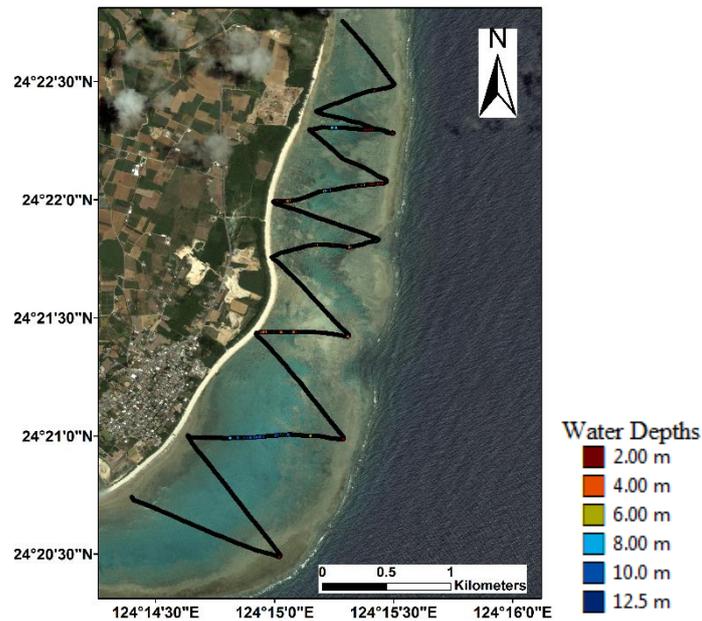
Figure 2-13. Field bathymetry Reference points of added part of Shiraho study area from echo-sounder.

The field data for Shiraho area were collected on 25-31 January 2013. Although there was a time difference between the Quickbird imagery collection and the field data observation but it is worth noting. As Shiraho area did not have any tsunami or big currents during these years so the bathymetry has no significant change [38].

CHAPTER 3

METHODOLOGY

3.1. Proposed Methodology for LULC Classification

The workflow processing steps for classifying Landsat-8 satellite image over El Burullus studied area were as follow:

3.1.1. Imagery Data Pre-Processing

First, calculating the spectral top of atmosphere reflectance of each pixel value from DN values using equation (1).

Second, computing At-Satellite Brightness Temperature for each pixel values of thermal infrared bands from the computed radiances using the Thermal Atmospheric correction tool in Envi program. All the required values were available in the image metadata file.

3.1.2. Creation of Attributes

For increasing the classification accuracy two other attributes NDVI and NDWI were computed from visible and the near-infrared bands. The NDVI were calculated using red and near infrared bands also the NDWI using green and near infrared bands. All the abovementioned steps were performed in an ENVI environment.

3.1.3. Selecting Uncorrelated Attributes using Principal Component Analysis Approach

PCA approach was used for selecting uncorrelated attributes as it's considered the most widely used method for feature selection from enormous data sets.

3.1.4. Classification Algorithms

RF, SVM and, NN with the back-propagation (BP) algorithm base classifiers were applied to LCLU classification then the ensemble is performed using the bagging approach.

1. Artificial Neural Network [3] [45] [34]

2. Support Vector Machines [39] [40] [41] [34]

3. Random Forest [50] [27] [51] [34]

4. Bagging Ensemble [30] [41] [56] [34]

For more details, the abovementioned references can be checked.

The following Figure 3-1 illustrates the workflow processing steps for classifying Landsat-8 satellite image over El Burullus studied area.
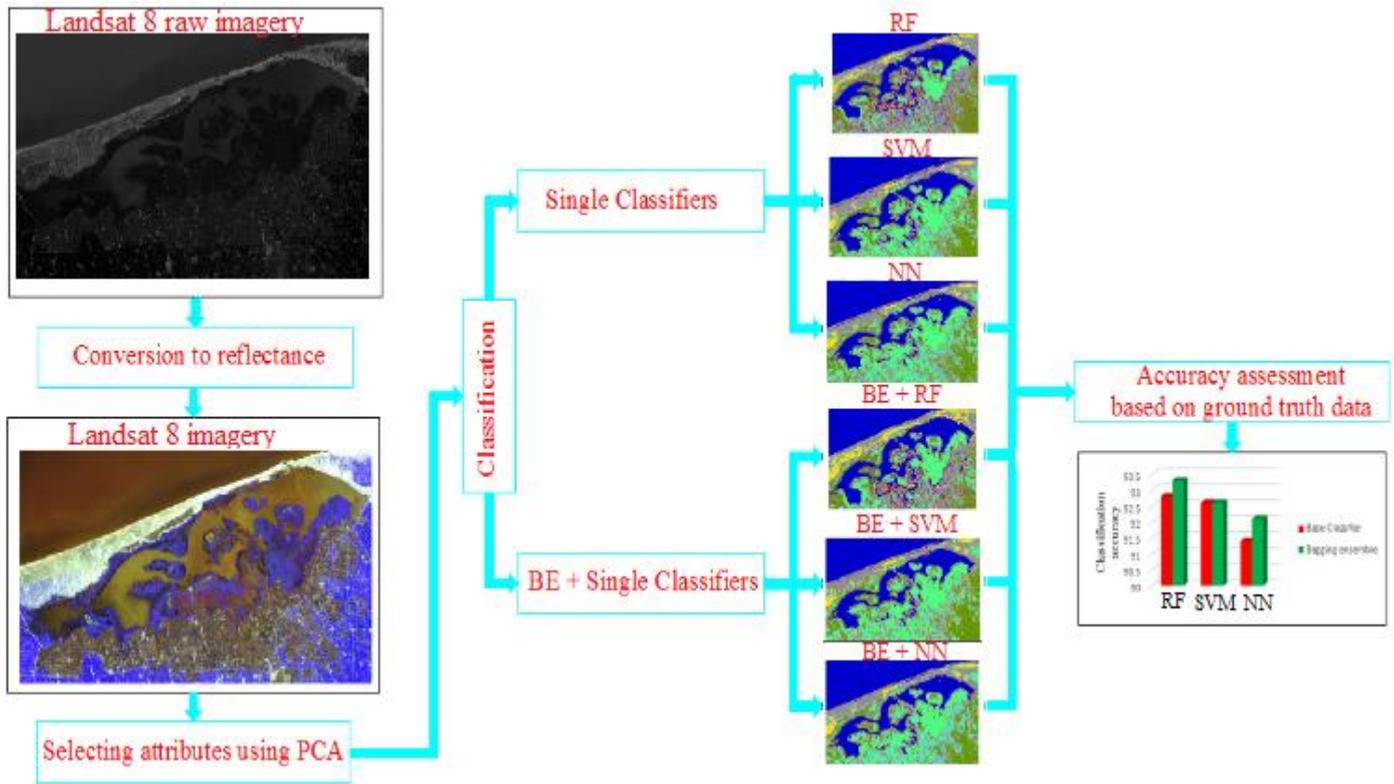
Figure 3-1. The processing steps for Landsat-8 images classification and classification accuracy comparison of all classifiers.

3.2. Proposed Methodologies for Bathymetry Determination (the first part)

All Landsat 8 and Spot 6 multispectral images of the studied areas were corrected for bathymetric mapping as follows: 1) Converting the image pixel values to radiance values using the images metadata file values according to equation 1.

2) Correcting atmospheric and sun glint errors for the image radiance values using the FLAASH tool and equation 2, respectively. These two steps were accomplished using ENVI 5.3 program.

3) Four inputs were extracted from the corrected reflectance images were used for training all approaches at the same location of sounding points. These values were red, green, blue/red, and green/red bands logarithms then the outputs were the detected water depths.

4) For all study areas, these values were randomly separated to independent 75% training and 25% testing points. For example, for El Burullus study area the field points were divided to 1875 and 625 points for training and testing, respectively.

16

5) Finally, the evaluation of all outputs from various approaches were done using the same independent testing points depending on RMSE and R^2 values.

The following Figure 3-2 illustrates the workflow of the bathymetry detection steps of the first part.
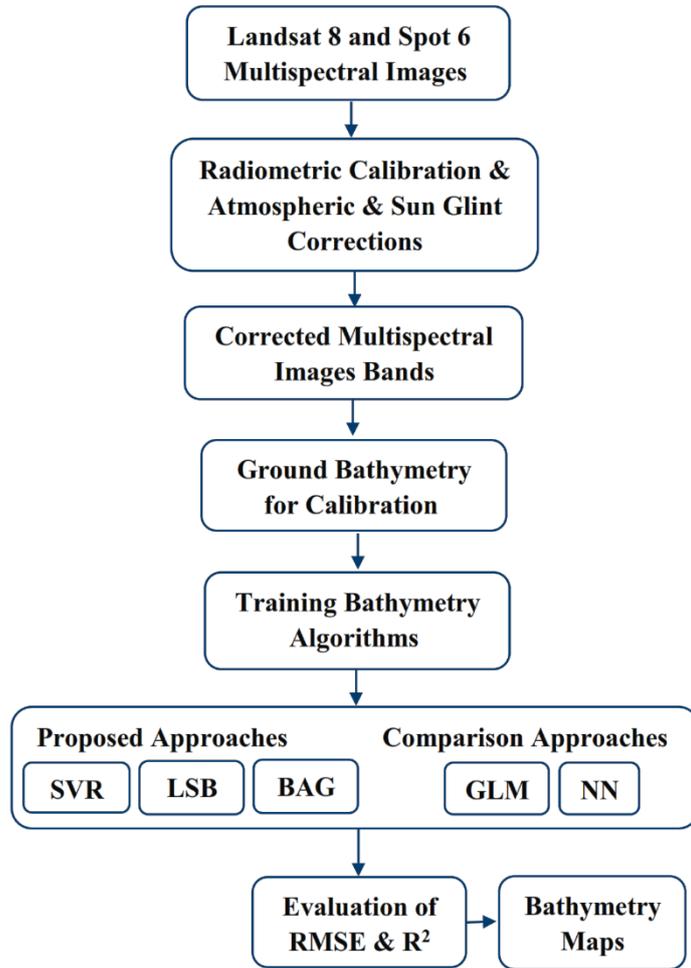
Figure 3-2. The processing steps of proposed methodologies for the first part of bathymetry mapping from satellite images by different approaches.

3.2.1. Imagery Data-Preprocessing

3.2.1.1. Spectral Top of Atmosphere Radiance

Spectral top of atmosphere radiance were calculated from the imagery pixel digital numbers (DN) values for each pixel as follows [42]:

$$L_\lambda = M_l * DN + A_l \tag{1}$$

Where L_λ = top of atmosphere spectral radiance, DN = digital number, M_l = multiplicative rescaling factor for radiances specified for each band, and A_l = additive rescaling factor for radiances specified for each band. The M_l and A_l values were presented in the images' metadata files (MTL files).

3.2.1.2. Atmospheric Correction

Atmospheric correction was performed to all images using the Fast Line-of-Sight Atmospheric Analysis of Hypercubes (FLAASH™) tool in the Envi 5.3 software package. FLAASH carry out radiative transfer based models founded on MODTRAN4 code [43] and has look-up tables for diverse categories of the atmosphere. Various categories of aerosols are included in the FLAASH tool, which explains the air particle properties as absorption and scattering. The calculated radiance images were used as input for FLAASH tool. Over all the study areas, the maritime category were selected as aerosol model category, the atmospheric model was tropical for hot areas, and two blue and infrared bands over water were selected as aerosol retrieval [26]. Finally, the results were the atmospherically corrected reflectance images.

3.2.1.3. Sun Glint Correction

The sun glint correction was applied after the atmospheric correction. The sun glint errors were corrected using the correlation between the bands used for bathymetry detection and the near-infrared band [44]-[16]. The de-glinted pixel value can be determined using Eq. 2:

$$R_i' = R_i * b_i (RNIR - MinNIR) \tag{2}$$

Where R_i' = de-glinted pixel reflectance value, R_i = atmospherically corrected reflectance value, b_i = regression line slope, RNIR = corresponding pixel value in NIR band, and MinNIR = min NIR value present in the sample.

The following approaches used for bathymetry detection.

1. Least Squares Boosting Fitting Ensemble [45] [46] [33]

2. Bagging Fitting Ensemble [47] [48] [33]

3. Support Vector Regression [49] [50] [33]

4. Lyzenga Generalized Linear Model Approach [22] [16] [33]

5. Artificial Neural Network Approach [14] [51] [33]

For more details, the abovementioned references can be checked.

3.3. Proposed Methodologies for Bathymetry Determination (the second part)

The Landsat 8, Landsat 7, and Spot 6 multispectral images of the four studied areas were corrected for bathymetric mapping using the same abovementioned processing steps.

Figure 3-3 illustrate the workflow of the bathymetry detection steps.
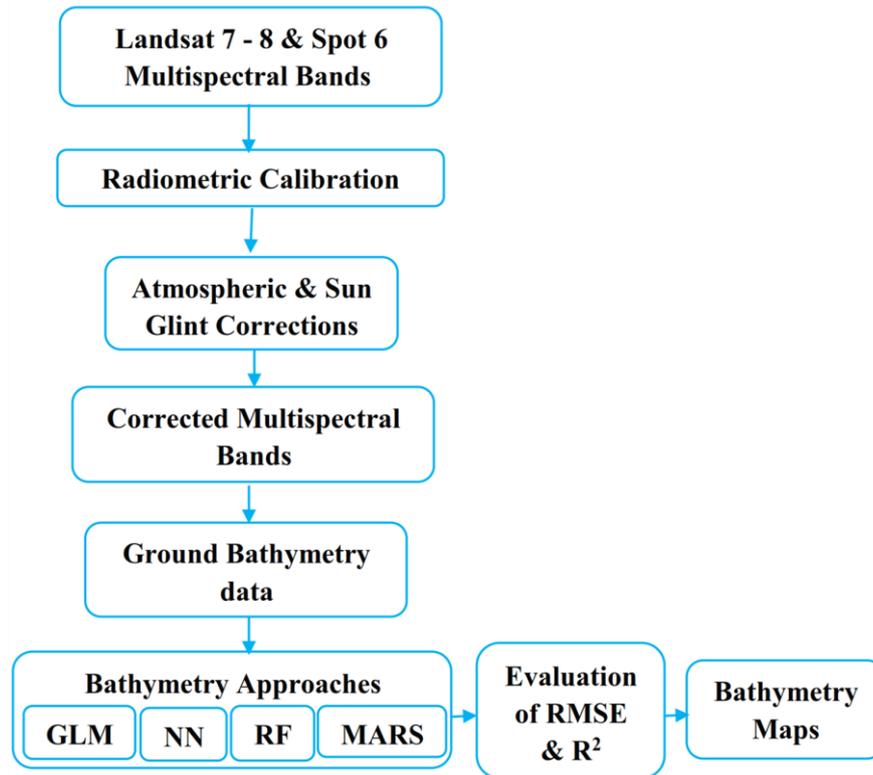
Figure 3-3. The processing steps of proposed methodologies for the second part of bathymetry mapping from satellite images by various approaches.

The following approaches proposed for bathymetry detection.

1. Random Forest [52] [8] [53]

2. Multi Adaptive Regression Spline [54] [55] [56]

For more details, the abovementioned references can be checked.

3.4. Proposed Methodologies for Bathymetry Determination (the third part)

For bathymetry mapping the NN and RF approaches were applied to the pre-processed Landsat 8 and Quickbird multispectral images and their predicted outputs were combined using BE algorithm.

The Landsat and Quickbird multispectral images of the study areas were corrected for bathymetric mapping as follows:

1) Converting the image pixel values to radiance values using the images metadata file values according to equation 1.

2) Correcting atmospheric and sun glint errors for the image radiance values using the FLAASH tool and equation 2, respectively. These two steps were accomplished using ENVI 5.3 program.

19

3) Four inputs were extracted from the corrected reflectance images are used for training all approaches at the same location of sounding points. These values were red, green, blue/red, and green/red bands logarithms then the outputs were the detected water depths.

4) For all study areas, these values were randomly separated to independent 65% training, 10% validation, and 25% testing points. For example, Shiraho study area field points were divided to 5472, 608 and 2026 points for training, validation, and testing, respectively.

5) BE approach was applied to ensemble the outputs from NN and RF methods. The 10% validation points were used for training the BE approach with NN and RF outputs as input values and predicted depths as outputs.

6) Finally, the evaluation of all outputs from various approaches were determined using the same independent testing points depending on RMSE and R^2 values.

The following figure illustrate the workflow of the bathymetry detection steps.
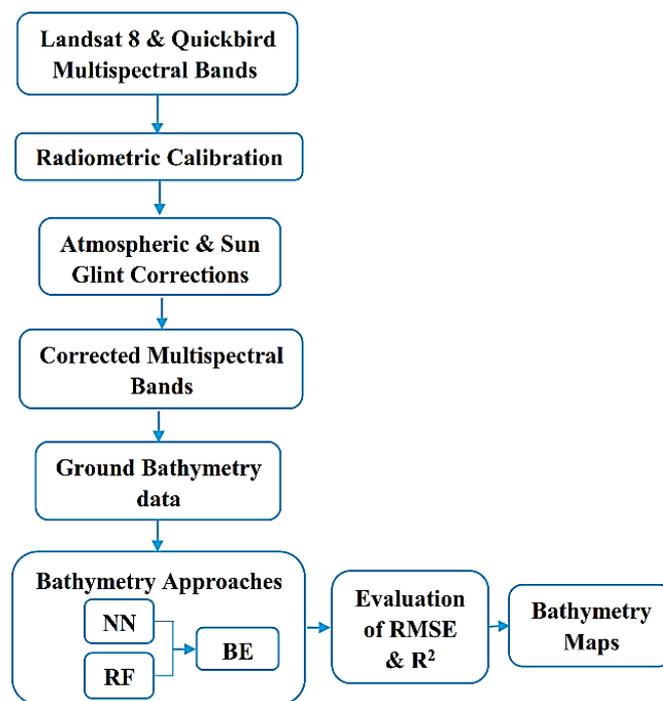
Figure 3-4. The Processing steps for the third part of bathymetry detection using BE approach.

The following approaches proposed for bathymetry detection.

1. Artificial Neural Network [14] [51] [33]

2. Random Forest [52] [8] [53]

3. Bagging Ensemble [11] [47] [57]

For more details, the abovementioned references can be checked.

CHAPTER 4

RESULTS

4.1. Results of LULC Proposed Methodology

The corrected input data was decreased to three principal components using 95% data variance. RF, SVM, and NN base classifiers were performed to Landsat 8 satellite image. At that point, bagging ensemble was used with these classifiers in a hierarchal structure. All these steps and the classification approaches were implemented in Matlab program environment.

For evaluating the accuracy of all used classifiers, reference field data was collected from Landsat 8 satellite image using field trip signature observations and old classified maps. The accuracy evaluation using omission, commission errors for all classes, and overall accuracy, with 1000 points as the test data uniformly spread over the study area. Table 4-1 presented evaluation results of the three single classifiers. Moreover, table 4-2 presented the related accuracy evaluation results of ensemble classifiers.

Figure 4-1 illustrates the classification resulted maps of RF, SVM, NN, BE with RF, BE with SVM, and BE with NN approaches.

(a) RF   (b) BE with RF

(c) SVM   (d) BE with SVM
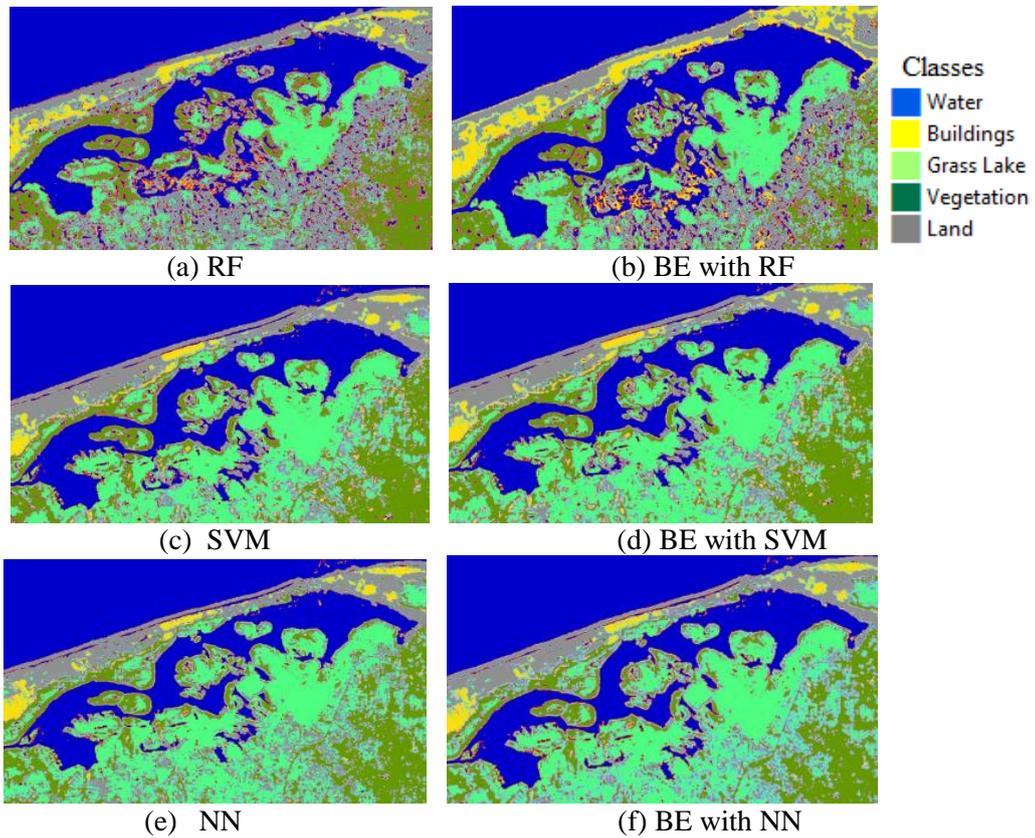
(e) NN   (f) BE with NN

Figure 4-1. Classification resulted maps for Landsat 8 satellite image. (a) RF (b) BE with RF (c) SVM (d) BE with SVM (e) NN (f) BE with NN.

Figure 4-2. Demonstrate classification accuracy enhancement using BE with RF, SVM, and NN with the three base classifiers.
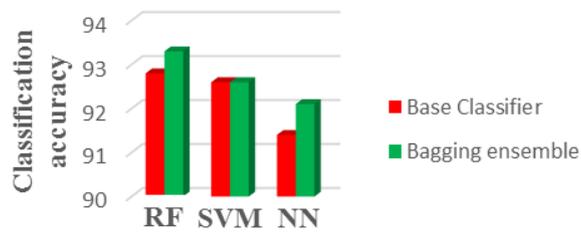
Figure 4-2. The classification accuracy enhancement using BE with the three base classifiers RF, SVM, and NN.

Table 4-1. The omission, commission errors for the five classes and overall accuracy of the three base classifiers RF, SVM, and NN.

Classifier	Class	Com. Err. (%)	Om. Err. (%)	Overall Acc. (%)
RF	Water	5.01	7.32	92.8
	Vegetation	4.15	6.73	
	Land	6.70	5.24	
	Buildings	12.44	10.66	
	Grass Lake	7.98	5.98	
SVM	Water	6.15	8.04	92.6
	Vegetation	3.69	6.28	
	Land	7.41	5.92	
	Buildings	11.37	8.78	
	Grass Lake	8.51	8.02	
NN	Water	7.5	10.63	91.4
	Vegetation	8.29	7.87	
	Land	9.79	2.78	
	Buildings	11.94	9.69	
	Grass Lake	5.32	11.44	

Table 4-2. The omission, commission errors for the five classes and overall accuracy of BE with the three base classifiers RF, SVM, and NN.

Classifier	Class	Com. Err. (%)	Om. Err. (%)	Overall Acc. (%)
BE with RF	Water	3.59	6.47	93.3
	Vegetation	4.61	2.82	
	Land	4.23	9.50	
	Buildings	14.69	7.22	
	Grass Lake	5.85	7.81	
BE with SVM	Water	4.62	7.00	92.6
	Vegetation	3.69	7.11	
	Land	6.35	7.81	
	Buildings	13.74	6.67	
	Grass Lake	8.51	8.51	
BE with NN	Water	5.13	8.87	92.1
	Vegetation	5.53	5.09	
	Land	7.41	6.42	
	Buildings	13.74	7.61	
	Grass Lake	7.45	11.68	

4.2. Results of First Part of Proposed Bathymetry Detection Methods

The proposed methods for bathymetry mapping SVR, LSB and BAG were tested using the corrected Landsat 8 and Spot 6 multispectral images and compared with NN and GLM algorithms. GLM produced the following equations for bathymetry mapping over the four study areas respectively:

$$Z_{El\ Burullus} = 3916.2 + 6231\ L_G - 6750\ L_R - 4088\ B/R - 3670\ G/R + 467.4\ L_G\ L_R - 4069\ L_G\ B/R + 1565\ L_R\ G/R - 216.4\ L_G\ G/R - 4463\ L_R\ B/R + 4618\ B/R\ G/R \tag{3}$$

$$Z_{Alex\ port} = 17.25 - 4.69\ L_G - 0.51\ L_R + 0.06\ B/R - 0.10\ G/R + 0.65\ L_G\ L_R - 0.03\ L_G\ B/R - 2.30\ L_R\ G/R + 0.06\ L_G\ G/R + 0.004\ B/R\ G/R \tag{4}$$

$$Z_{Nubia\ lake} = 2912.2 - 904.96\ L_G + 1219.7\ L_R - 3024.6\ B/R - 1900.7\ G/R + 19.35\ L_G\ L_R - 1.06\ L_G\ B/R + 1.07\ L_R\ G/R - 18.44\ L_G\ G/R - 1281.1\ L_R\ B/R + 2143.8\ B/R\ G/R \tag{5}$$

$$Z_{Shiraho} = -15.185 + 29.67\ L_G - 39.73\ L_R - 10.48\ B/R + 73.43\ G/R + 0.44\ L_G\ L_R + 28.63\ L_G\ B/R - 15.2\ L_R\ G/R + 5.09\ L_G\ G/R - 18.22\ L_R\ B/R + 3.36\ B/R\ G/R \tag{6}$$

Where L_G is the computed logarithm of the corrected green band values, L_R is the computed logarithm of the corrected red band values, B/R is blue divided by red and G/R is green divided by red logarithms band values.

The SVR was performed using the PUK kernel function and the sequential minimum optimisation SMO for solving the optimisation problem. The SVR were performed using the following set of parameters after many trials: C = 1, ε = 0.0, ζ = 0.001, and tolerance = 0.001. The parameters of PUK kernel were σ = 0.5 and ω = 0.5. Alternatively, the NN has been trained using Levenberg-Marquardt backpropagation training function with 10 hidden layers. The regression ensemble algorithms LSB and BAG were ensembles with 50 regression trees. These parameters for each algorithm were selected to achieve the least possible RMSE and maximum R^2 values. These algorithms and the statistical analysis were applied in MATLAB environment. The SVR code was originally established by Clark [58].

Figures 4-3, 4-5, 4-7, and 4-9 show the produced bathymetric maps resulted from the three models using Landsat 8 and Spot 6 satellite images over the four study areas. Figures 4-4, 4-6, 4-8, and 4-10 shows the assessment of the three models, and Tables 4-3, 4-4, 4-5, and 4-6 presents the resultant RMSE and R^2 values.

(a) GLM

(b) NN
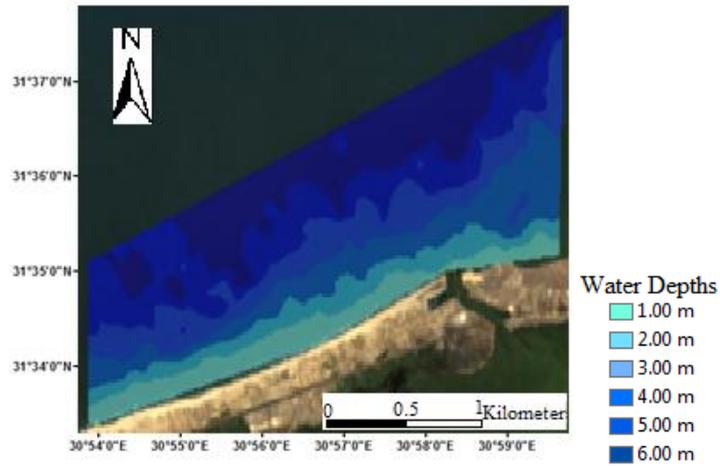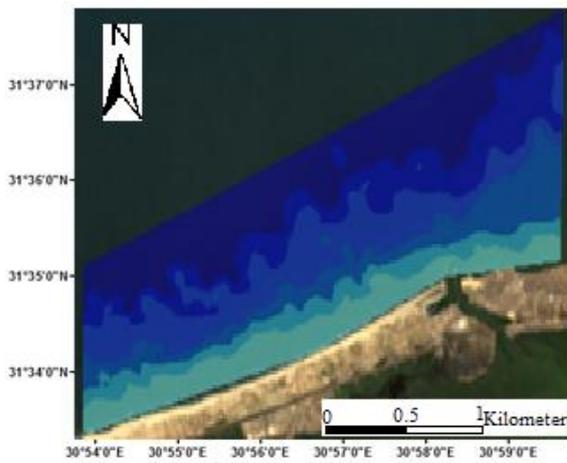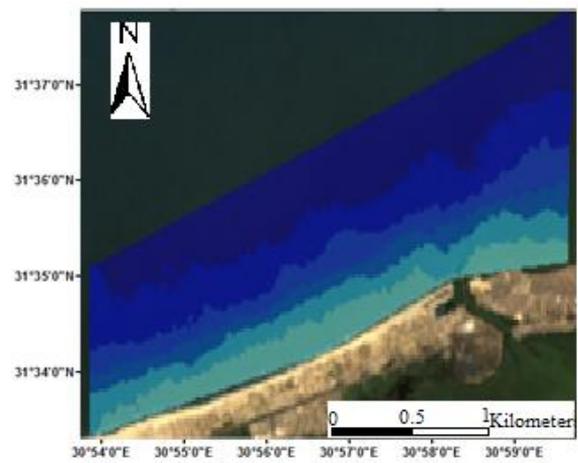
(c) SVR
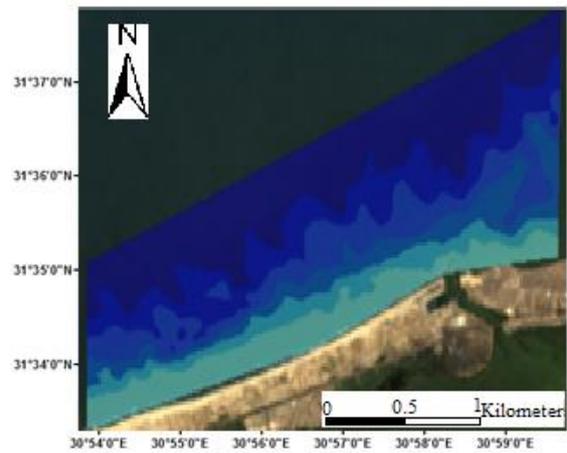
(d) LSB
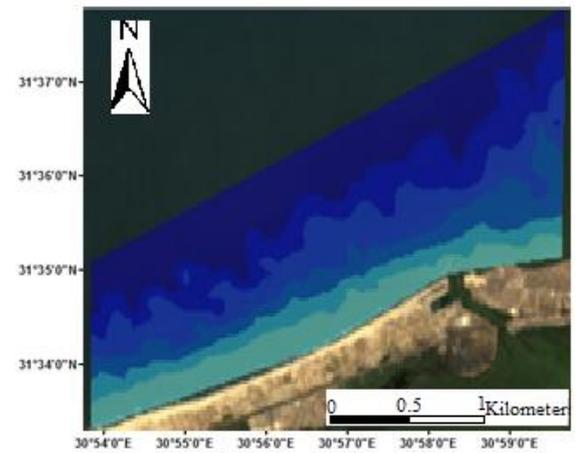
(e) BAG

Figure 4-3. Bathymetric maps produced from each algorithm using Landsat-8 imagery over El Burullus inlet area, Egypt. (a) GLM (b) NN (c) SVR (d) LSB (e) BAG.
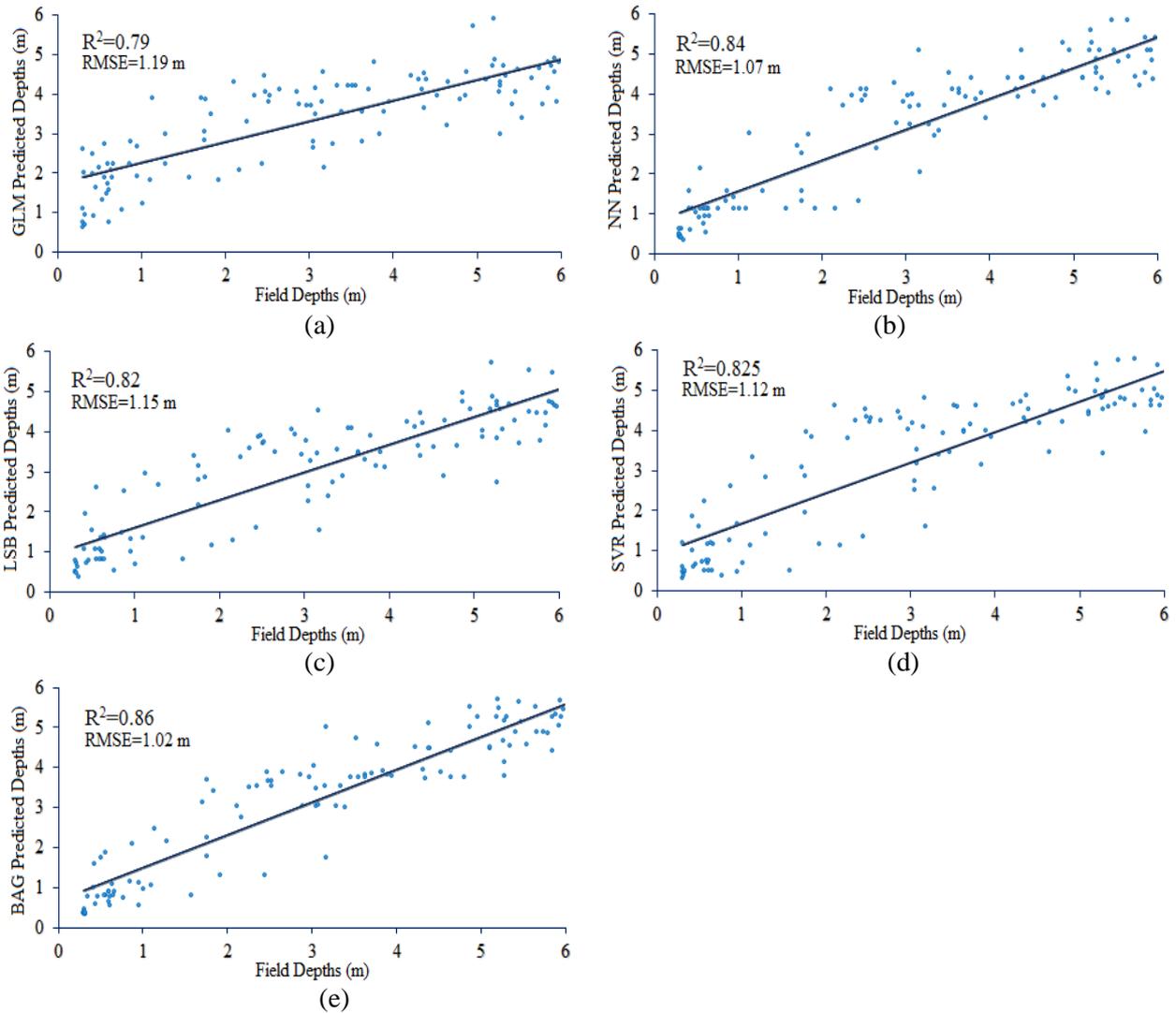
Figure 4-4. The resultant continuous fitted models for El Burullus Lake area, Egypt. Depths are represented as points, and the solid line represents the continuous fitted model (a) GLM (b) NN (c) SVR (d) LSB (e) BAG.

Table 4-3. The resultant RMSEs and R^2 of all methods for bathymetry mapping El Burullus inlet area, Egypt.

Methodology	GLM	NN	SVR	LSB	BAG
RMSE (m)	1.19	1.07	1.12	1.15	1.02
R^2	0.79	0.84	0.825	0.82	0.86
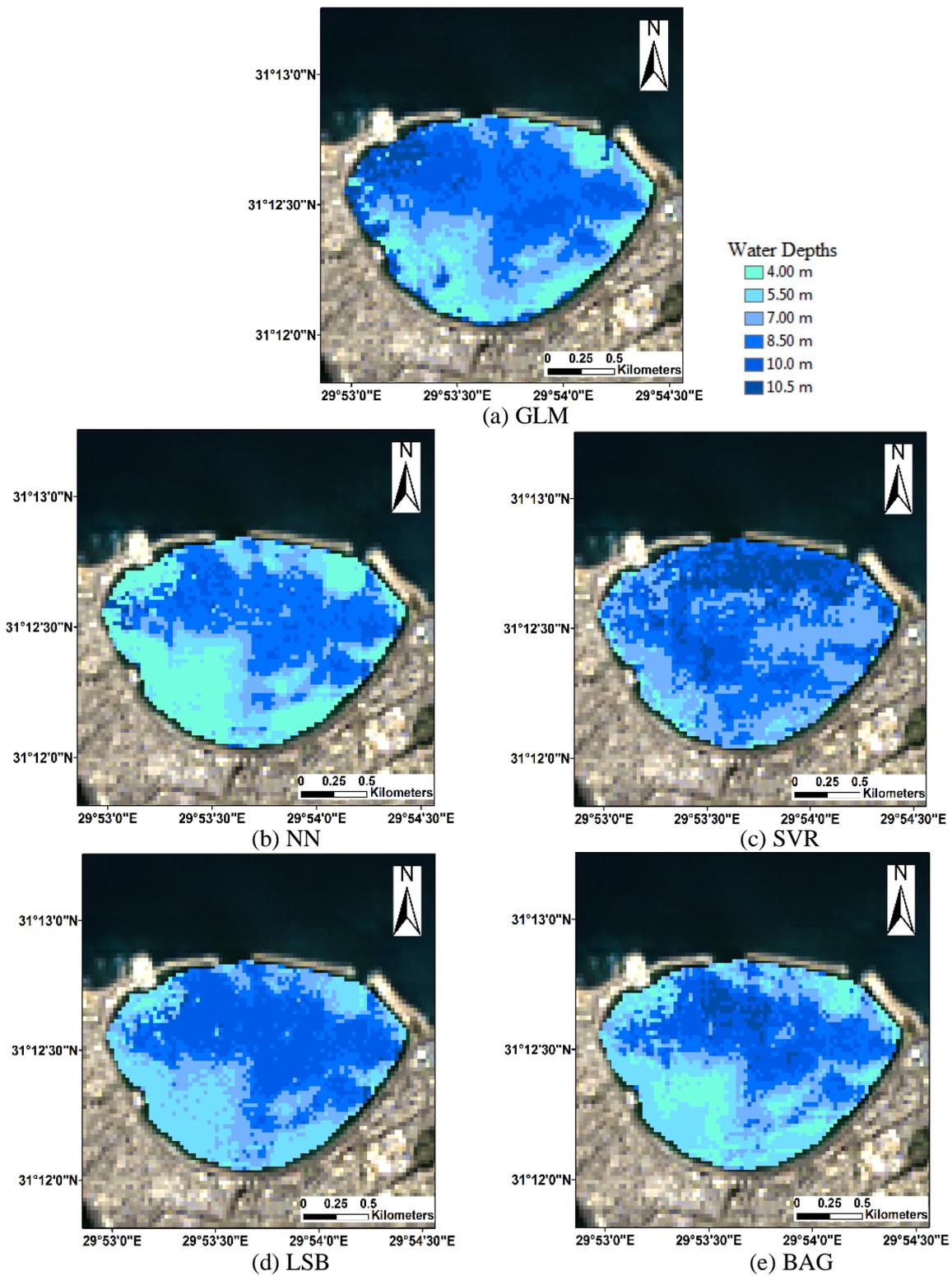
Figure 4-5. Bathymetric maps produced from each algorithm using Landsat-8 imagery over Alexandria port area, Egypt. (a) GLM (b) NN (c) SVR (d) LSB (e) BAG.
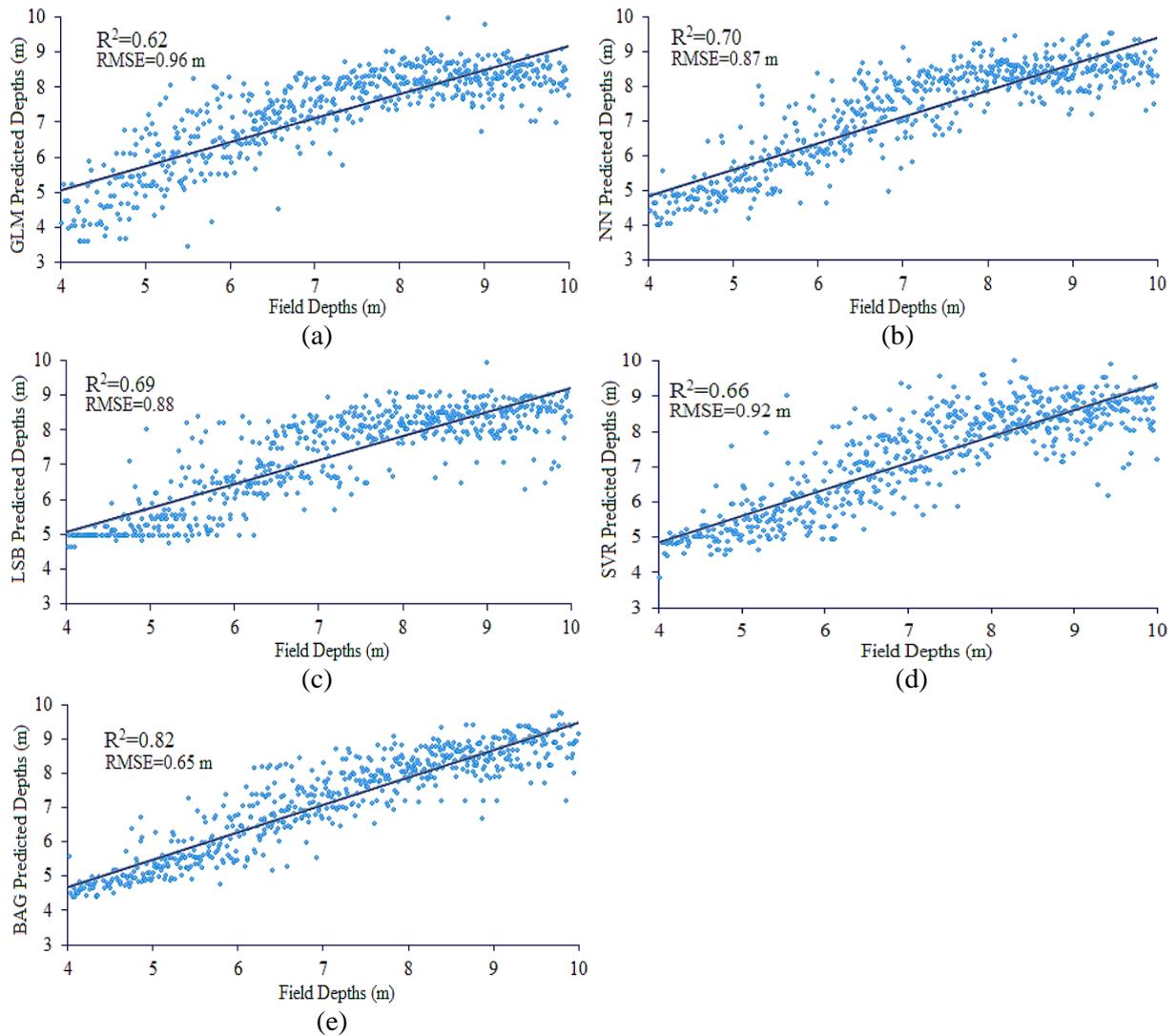
Figure 4-6. The resultant continuous fitted models for Alexandria harbor area, Egypt. Depths are represented as points, and the solid line represents the continuous fitted model (a) GLM (b) NN (c) SVR (d) LSB (e) BAG.

Table 4-4. The resultant RMSEs and R^2 of all methods for bathymetry mapping over Alexandria port area, Egypt.

Methodology	GLM	NN	SVR	LSB	BAG
RMSE (m)	0.96	0.87	0.92	0.88	0.65
R^2	0.62	0.70	0.66	0.69	0.82
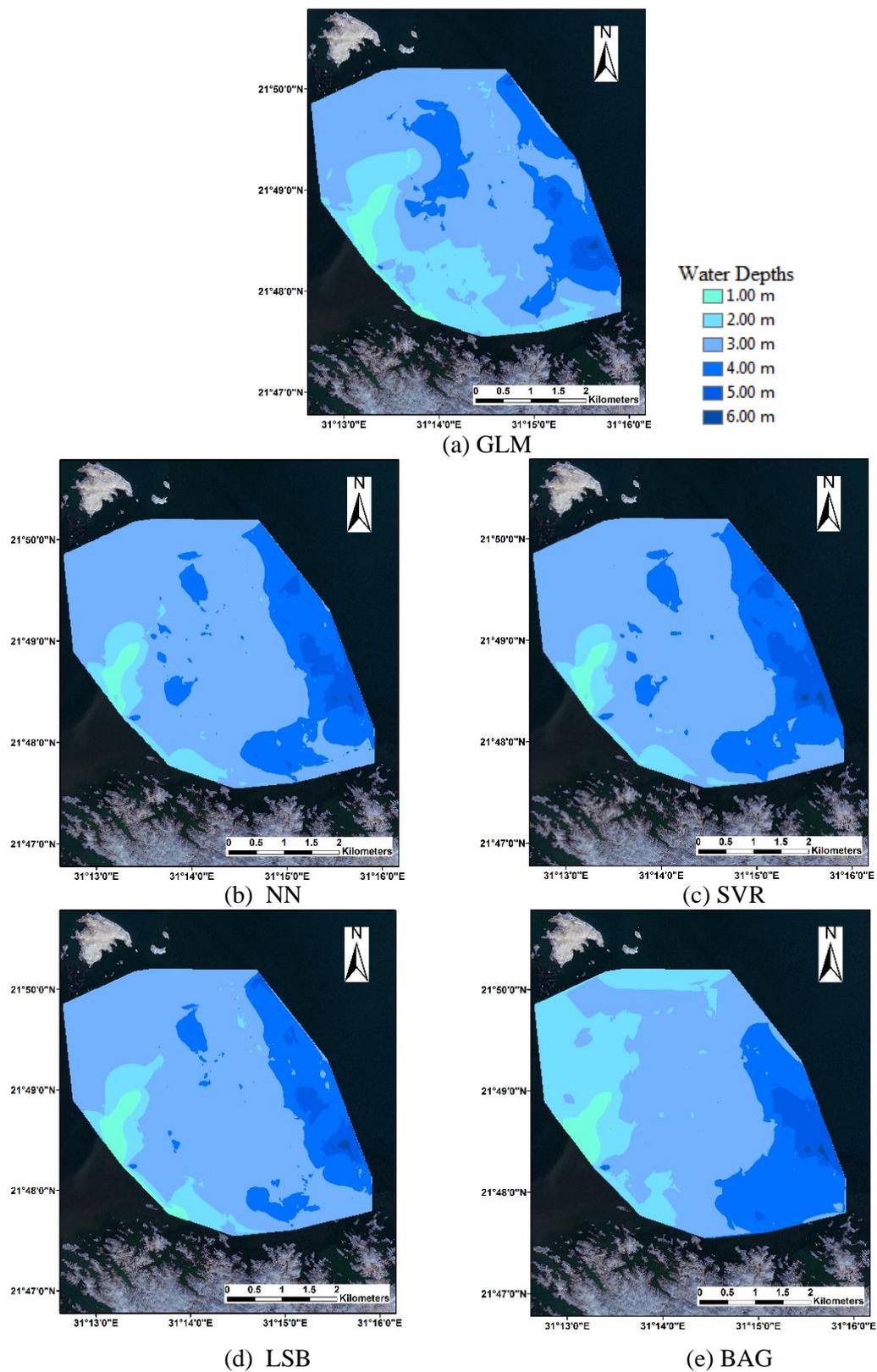
Figure 4-7. Bathymetric maps produced from each algorithm using Spot 6 imagery over Nubia Lake entrance zone, Sudan. (a) GLM (b) NN (c) SVR (d) LSB (e) BAG.
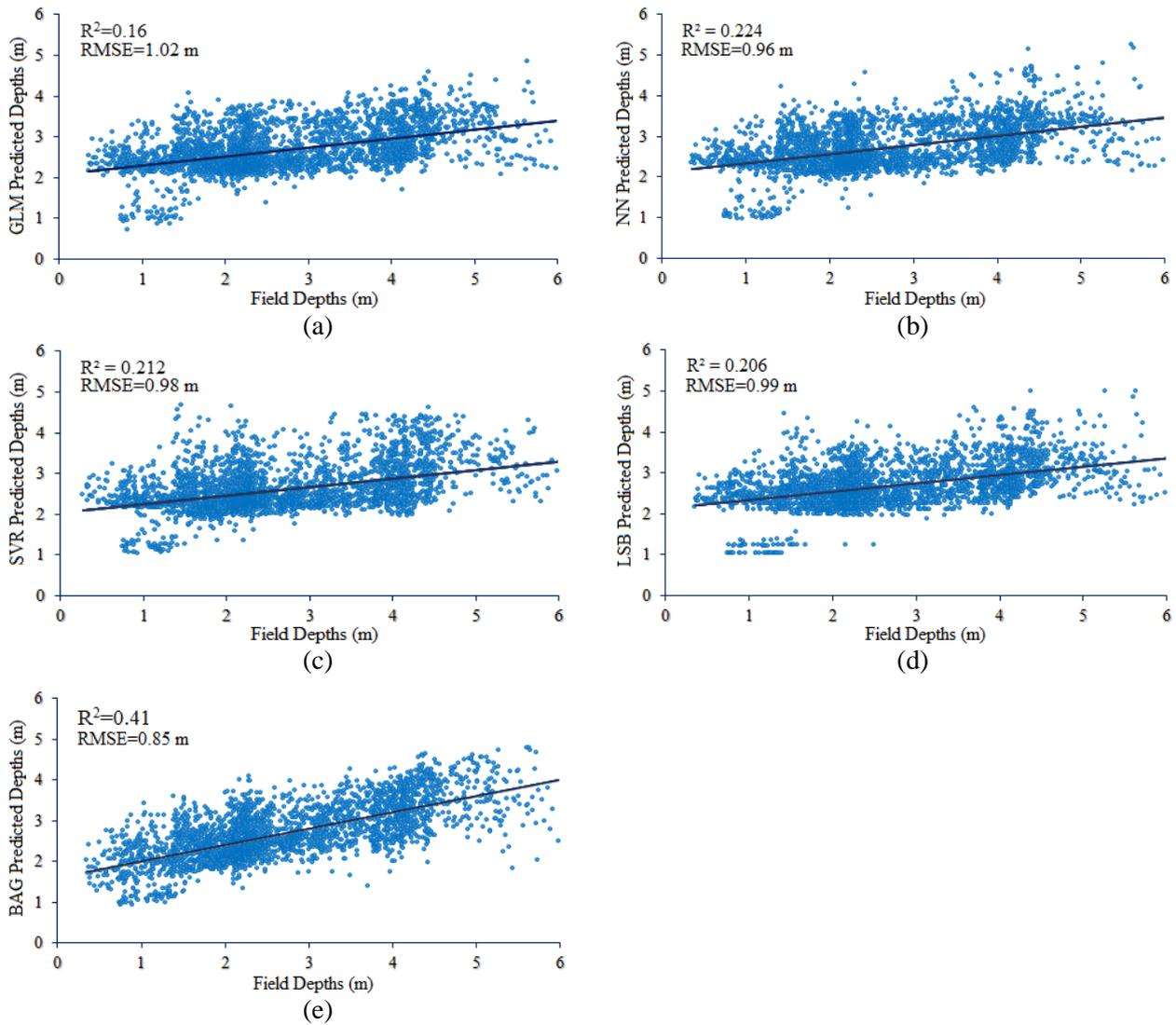
Figure 4-8. The resultant continuous fitted models for Nubia Lake entrance zone, Sudan. Depths are represented as points, and the solid line represents the continuous fitted model (a) GLM (b) NN (c) SVR (d) LSB (e) BAG.

Table 4-5. The resultant RMSEs and R^2 of all methods for bathymetry mapping for Nubia Lake entrance zone, Sudan.

Methodology	GLM	NN	SVR	LSB	BAG
RMSE (m)	1.02	0.96	0.98	0.99	0.85
R^2	0.16	0.224	0.212	0.206	0.41
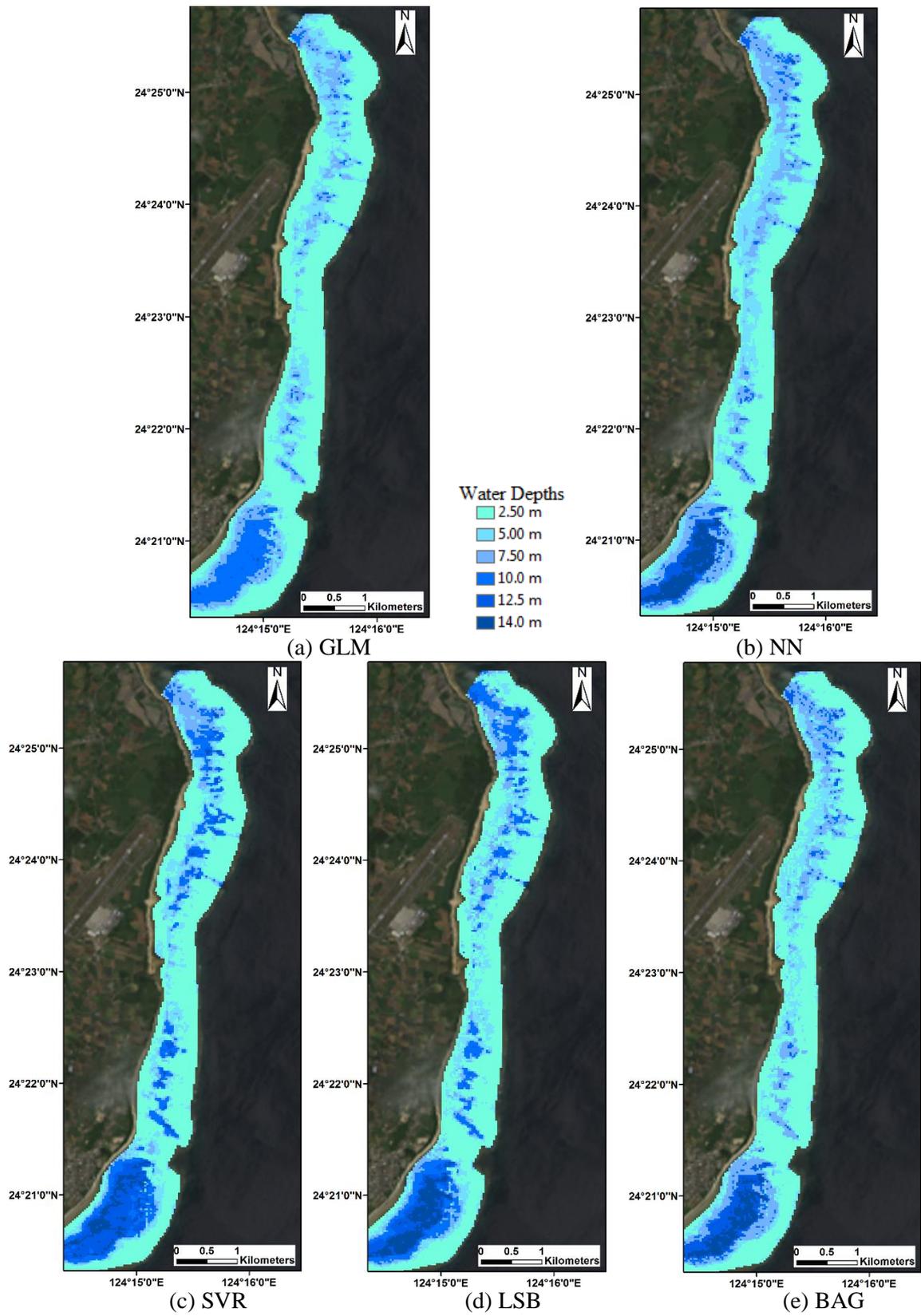
Figure 4-9. Bathymetric maps produced from each algorithm using Landsat-8 imagery over Shiraho Island area, Japan. (a) GLM (b) NN (c) SVR (d) LSB (e) BAG.
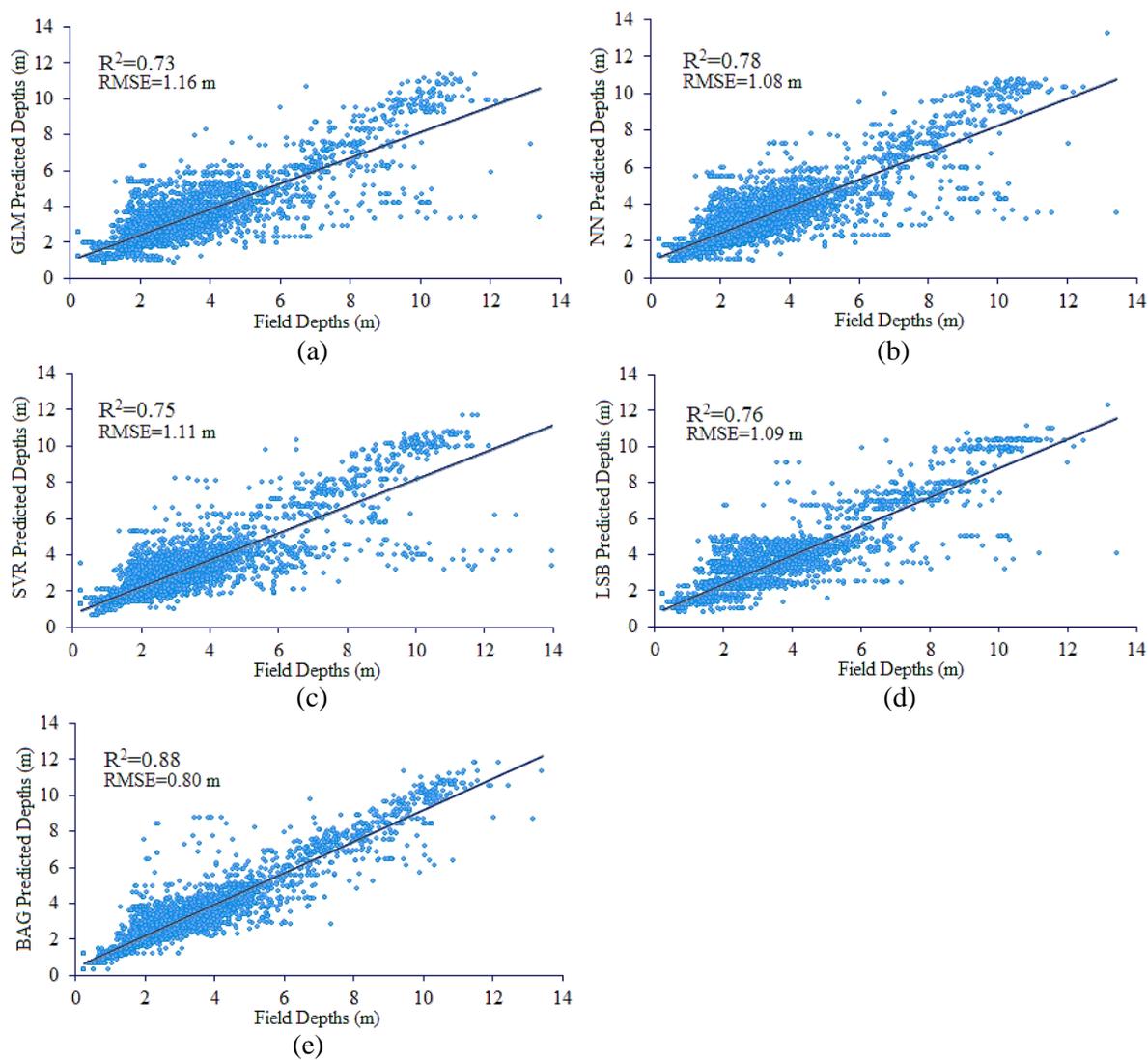
Figure 4-10. The resultant continuous fitted models for Shiraho Island, Japan. Depths are represented as points, and the solid line represents the continuous fitted model (a) GLM (b) NN (c) SVR (d) LSB (e) BAG.

Table 4-6. The resultant RMSEs and R^2 of all methods for bathymetry mapping for Shiraho Island, Japan.

Methodology	GLM	NN	SVR	LSB	BAG
RMSE (m)	1.16	1.08	1.11	1.09	0.80
R^2	0.73	0.78	0.75	0.76	0.88

4.3. Results of Second Part of Proposed Bathymetry Detection Methods

The proposed methods for bathymetry mapping RF and MARS were tested using the corrected Landsat 8, Landsat 7, and Spot 6 multispectral images and compared with NN and GLM algorithms. GLM produced the following equations for bathymetry mapping over Nubia study areas in addition to the three equations 4, 5, and 6.

$$Z_{\text{Nubia lake}} = -4676 + 11.33\text{LG} - 317.2\text{ LR} + 13.49\text{ B/R} + 18.25\text{ G/R} - 45.02\text{ LG LR} - 0.26\text{ LG B/R} + 0.03\text{ LR G/R} - 0.91\text{ LG G/R} + 10.48\text{ LR B/R} - 0.45\text{ B/R G/R} \tag{8}$$

The MARS was performed using the following parameters:

1) The maximum number of BFs in the forward phase before being pruned in the backward phase was set as 40.

2) The GCV value was set as 4

3) The linear piecewise modelling was used. Alternatively, the NN has been trained using Levenberg-Marquardt backpropagation training function with 10 hidden layers. The regression ensemble algorithms LSB and BAG were ensembles with 50 regression trees. These parameters for each algorithm were selected to have the least possible RMSE and maximum R^2 values. These algorithms and the statistical analysis were applied in MATLAB environment and the MARS model was originally established by Jekabson's toolbox [59].

Figures 4-11, 4-13, 4-15, and 4-17 show the produced bathymetric maps resulted from the two proposed models using Landsat 8, Landsat 7, and Spot 6 satellite images over the four study areas, figures 4-12, 4-14, 4-16, and 4-18 the assessment of the two models, and tables 4-7, 4-8, 4-9, and 4-10 presents the resultant RMSE and R^2 values.
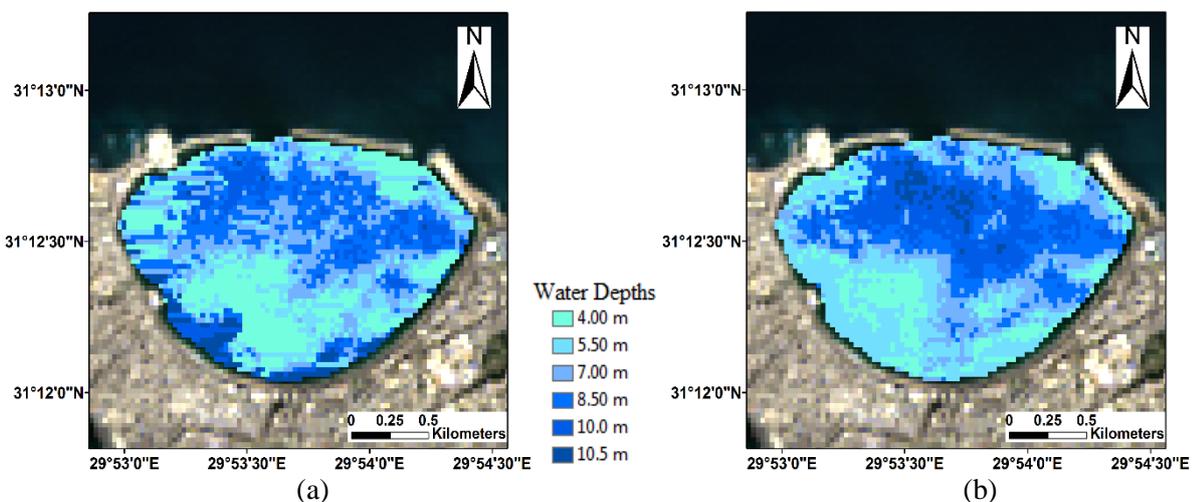
Figure 4-11: Bathymetric maps produced from Landsat-8 imagery over Alexandria harbor area, Egypt using (a) MARS (b) RF.
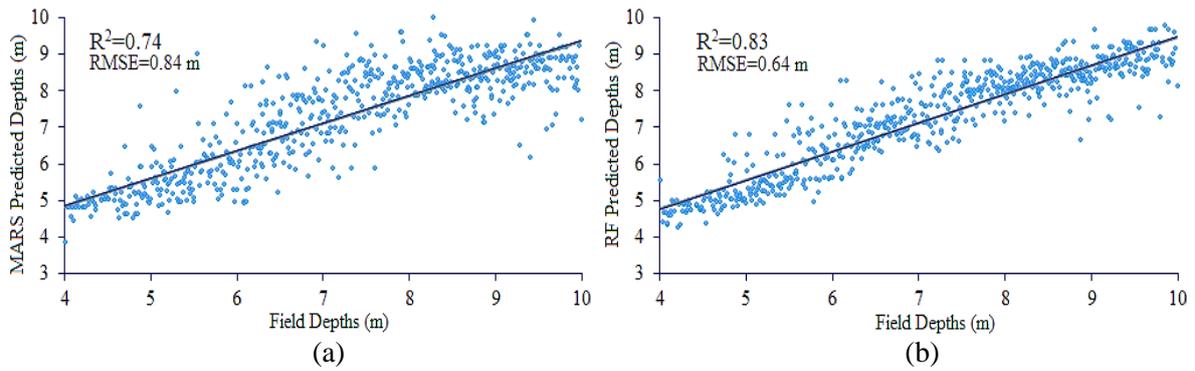
Figure 4-12: The resultant continuous fitted models for Alexandria port area, Egypt. Depths are appeared as points, and the solid line represents the continuous fitted model (a) MARS (b) RF.

Table 4-7: The resultant RMSEs and R^2 of all methods for bathymetry mapping over Alexandria port area, Egypt

Methodology	GLM	NN	MARS	RF
RMSE (m)	0.96	0.92	0.84	0.64
R^2	0.62	0.65	0.74	0.83
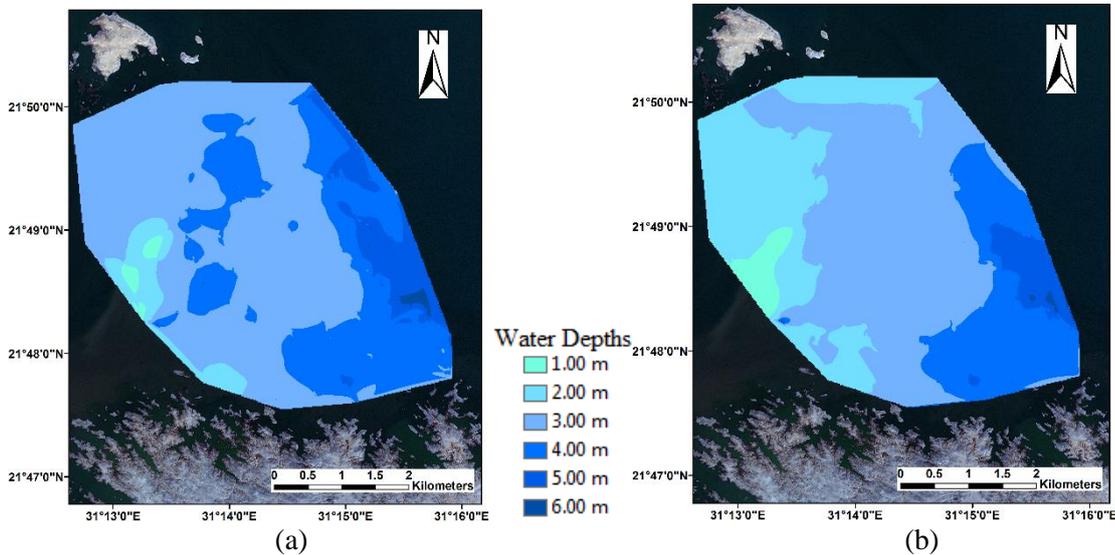
Figure 4-13: Bathymetric maps produced from Spot 6 imagery over Nubia Lake entrance zone, Sudan using (a) MARS (b) RF.
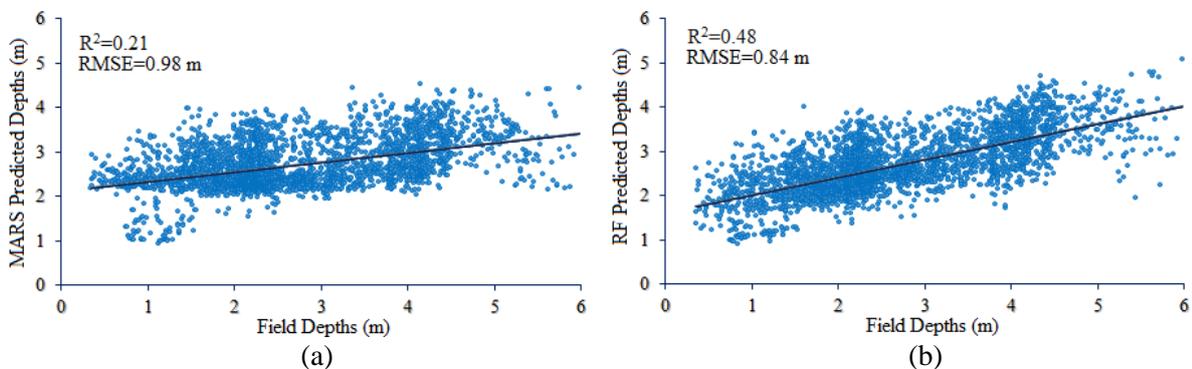
Figure 4-14: The resultant continuous fitted models for Nubia Lake entrance zone, Sudan. Depths are appeared as points, and the solid line represents the continuous fitted model (a) MARS (b) RF.

Table 4-8: The resultant RMSEs and R^2 of all methods for bathymetry mapping over Nubia Lake entrance zone, Sudan

Methodology	GLM	NN	MARS	RF
RMSE (m)	1.02	0.96	0.98	0.84
R^2	0.16	0.23	0.21	0.48
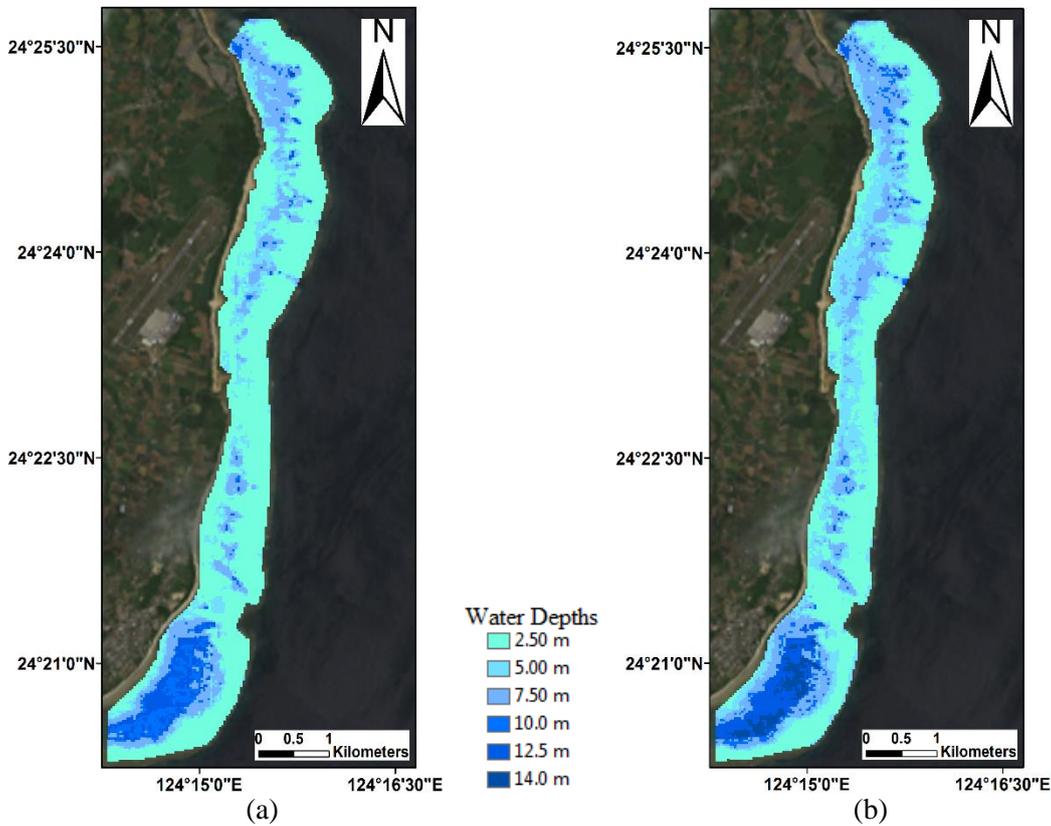
Figure 4-15. Bathymetric maps produced from Landsat-8 imagery over Shiraho Island area, Japan using (a) MARS (b) RF.
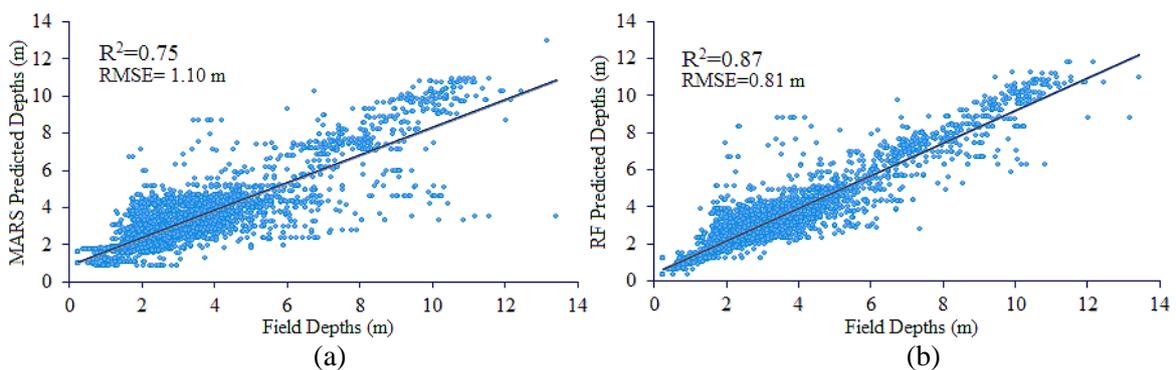
Figure 4-16: The resultant continuous fitted models for Shiraho Island, Japan. Depths are appeared as points, and the solid line represents the continuous fitted model (a) MARS (b) RF.

Table 4-9: The resultant RMSEs and R^2 of all methods for bathymetry mapping over Shiraho Island, Japan

Methodology	GLM	NN	MARS	RF
RMSE (m)	1.16	1.08	1.10	0.81
R^2	0.73	0.78	0.75	0.87
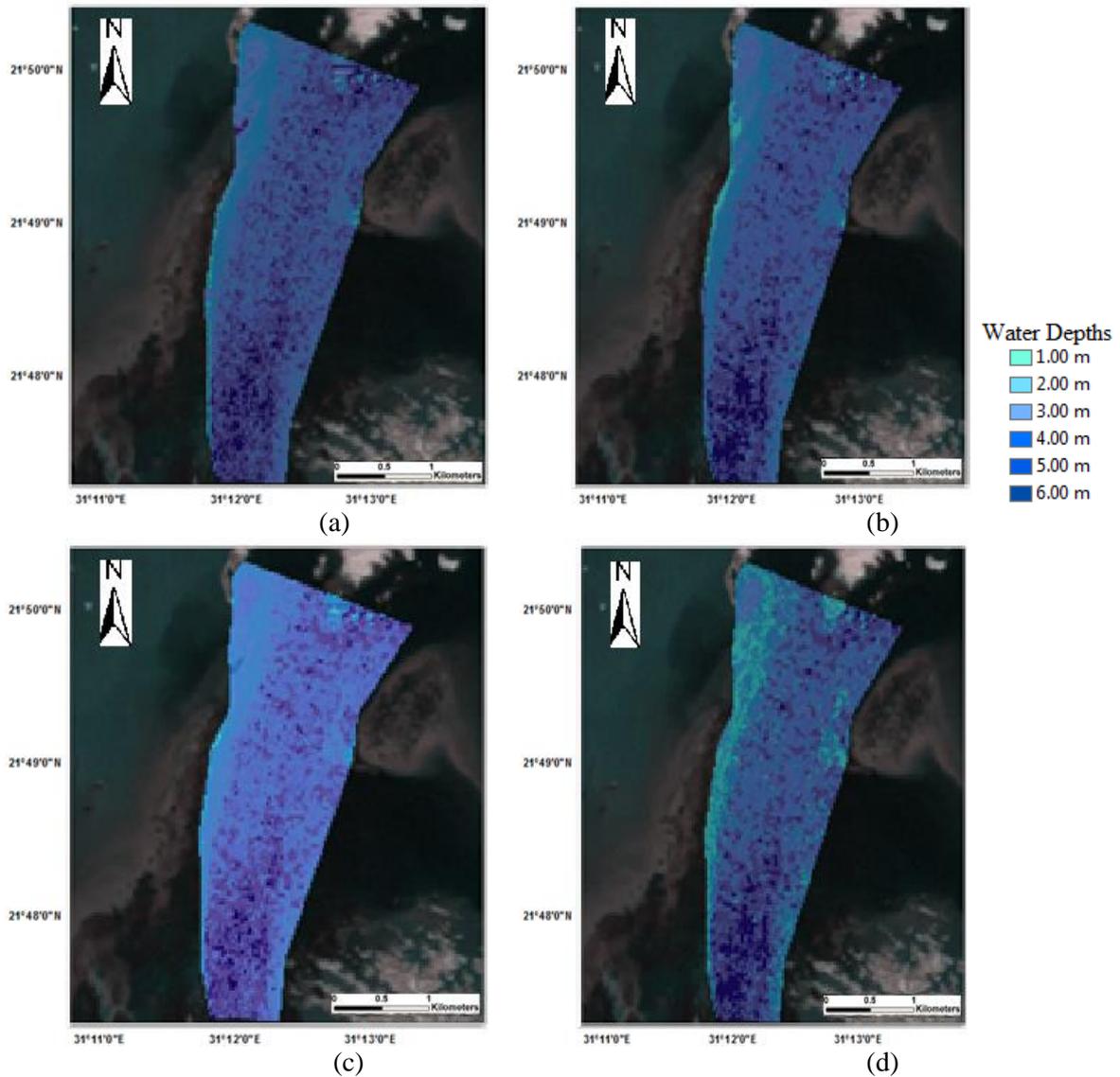
Figure 4-17. Bathymetric maps produced from Landsat-7 imagery over Nubia Lake entrance zone, Sudan using (a) GLM (b) NN (c) MARS (d) RF.
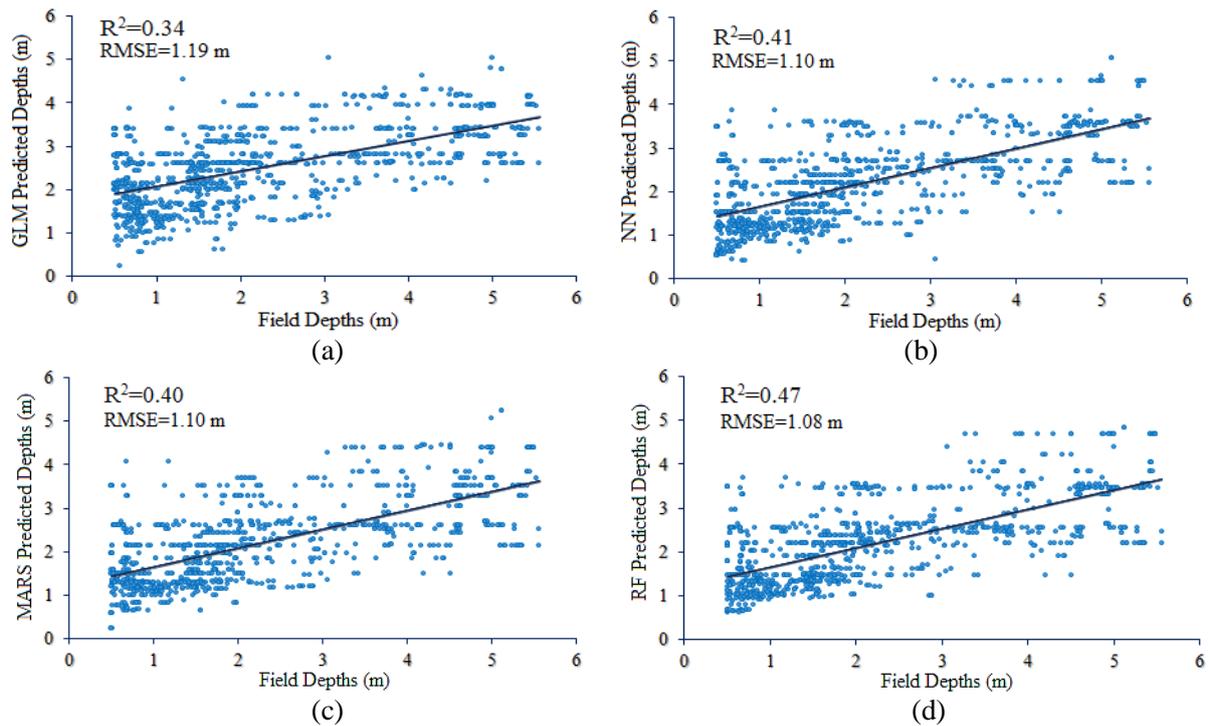
Figure 4-18: The resultant continuous fitted models for Nubia Lake entrance zone, Sudan. Depths are appeared as points, and the solid line represents the continuous fitted model (a) GLM (b) NN (c) MARS (d) RF

Table 4-10: The resultant RMSEs and R^2 of all methods for bathymetry mapping over Nubia Lake entrance zone, Sudan.

Methodology	GLM	NN	MARS	RF
RMSE (m)	1.19	1.10	1.10	1.08
R^2	0.34	0.41	0.40	0.47

4.4. Results of Third Part of Proposed Bathymetry Detection Methods

The parameters of the NN training function was Levenberg-Marquardt backpropagation with 10 hidden layers. RF model was constructed with 50 regression trees and 2/3 split percentage. Finally, BE algorithm ensemble the outputs with 50 trees. These parameters were selected for each algorithm based on the least possible RMSE and highest R^2 values. All these algorithms were implemented in Matlab environment.

Figures 4-19 and 4-21 shows the bathymetric maps computed by applying each model using the Landsat 8 and Quickbird satellite images for each study area, Figures 4-20 and 4-22 the evaluation of each model, and Tables 4-11 and 4-12 summarises the corresponding RMSE and R^2 values.
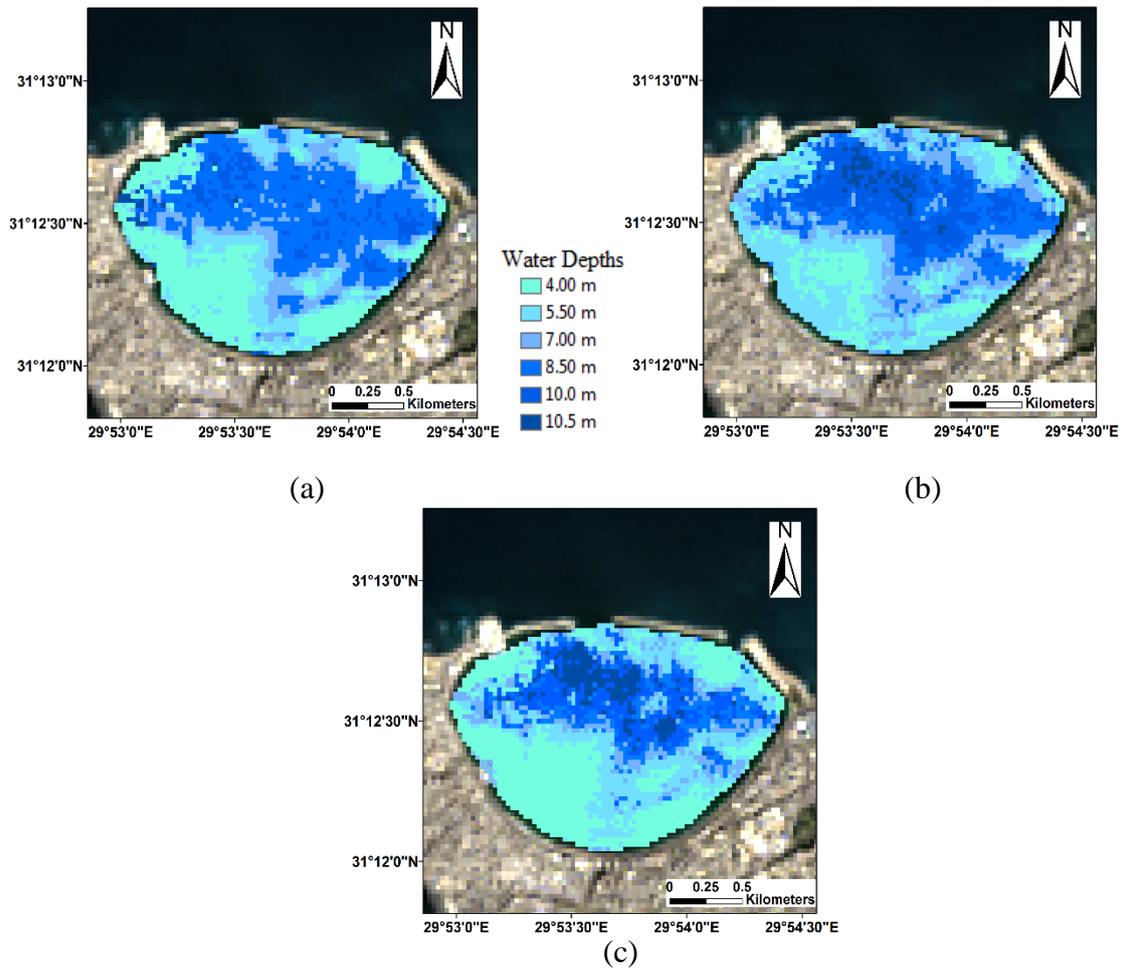
Figure 4-19. Bathymetric maps derived by applying each algorithm using Landsat-8 imagery over Alexandria harbor area, Egypt. (a) NN (b) RF (c) BE.
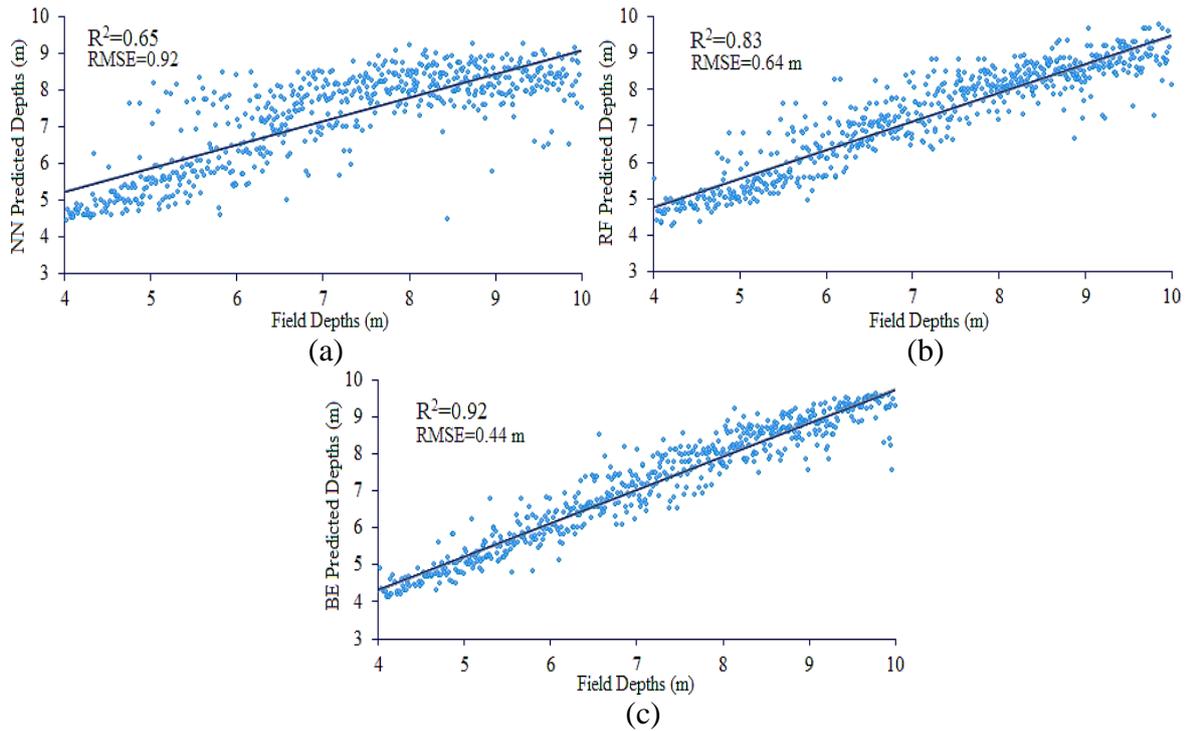
Figure 4-20. The continuous fitted models for Alexandria port area, Egypt. Depths are represented as points, and the continuous line represents the continuous fitted model (a) NN (b) RF (c) BE.

Table 4-11. The RMSEs and R^2 of all methods for bathymetry detection over Alexandria port area, Egypt.

Methodology	NN	RF	BE
RMSE (m)	0.92	0.64	0.44
R^2	0.65	0.83	0.92
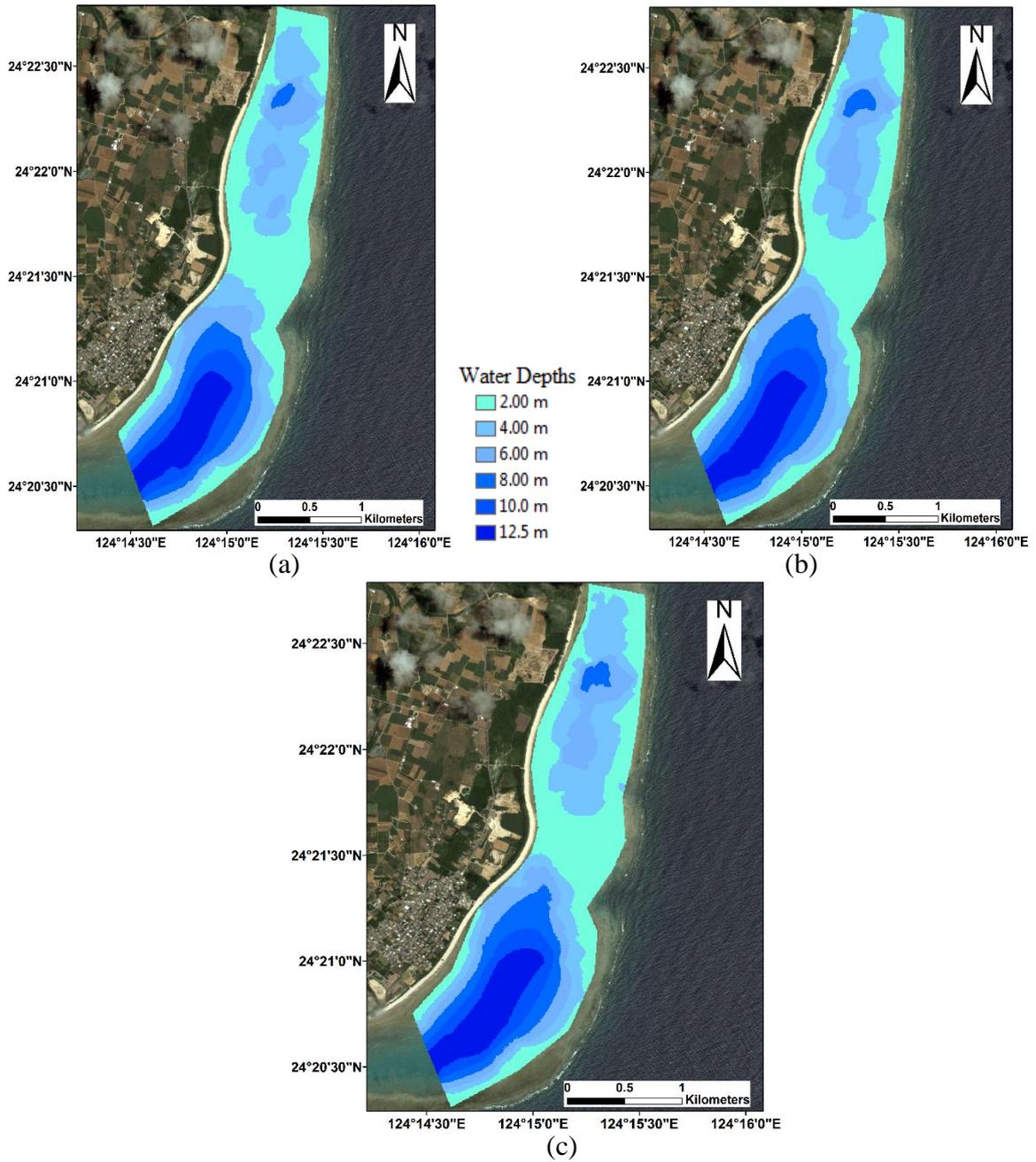
Figure 4-21. Bathymetric maps derived by applying each algorithm using Quickbird imagery over Shiraho Island area, Japan. (a) NN (b) RF (c) BE.
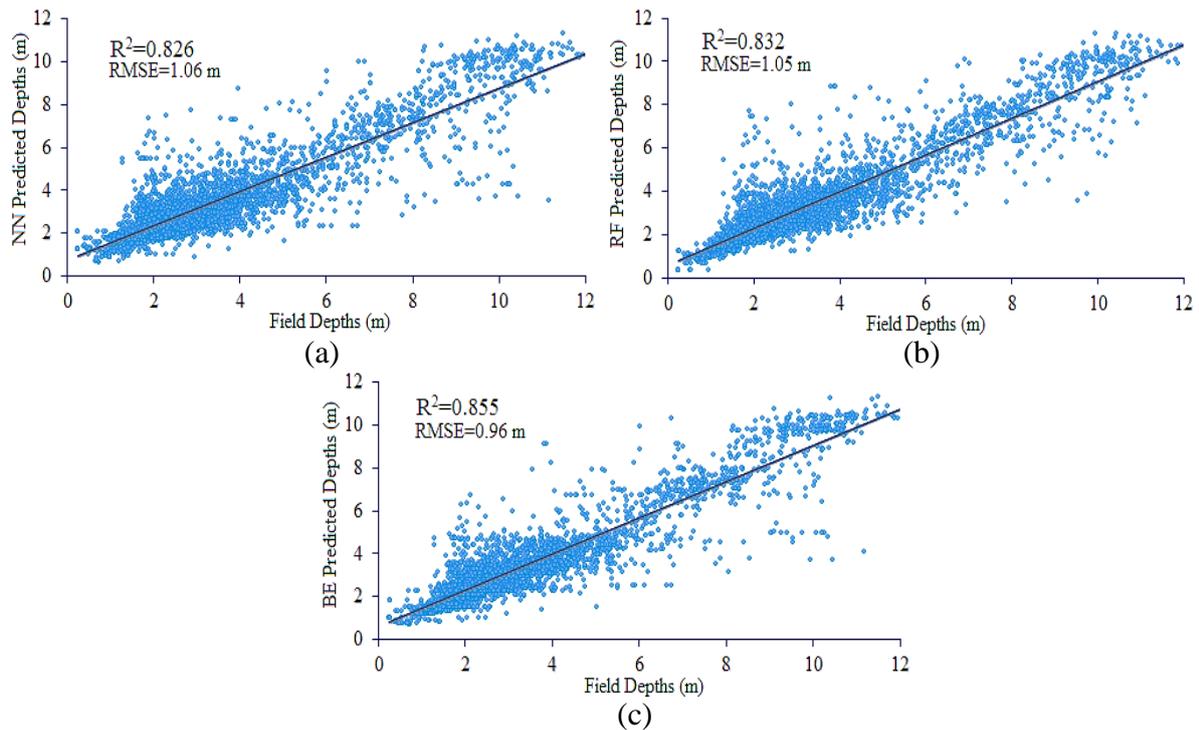
Figure 4-22. The continuous fitted models for Shiraho Island, Japan. Depths are represented as points, and the continuous line represents the continuous fitted model (a) NN (b) RF (c) BE.

Table 4-12. The RMSEs and R^2 of all methods for bathymetry detection for Shiraho Island, Japan.

Methodology	NN	RF	BE
RMSE (m)	1.06	1.05	0.96
R^2	0.826	0.832	0.855

4.5. Summary

This chapter outlined a pixel-based LULC classification methodology using the BE with RF in a hierarchal arrangement. This methodology was proposed and assessed using Landsat 8 satellite image over a coastal heterogeneous territory. To confirm the efficiency of the presented approach over SVM and NN base classifiers, classification was applied using a Landsat 8 satellite image over Egypt's Lake El-Burullus and its environments. All the necessary observed reference data were collected by manually on screen digitizing of the Landsat 8 image. The overall accuracy of the three base classifiers RF, SVM, and NN were 92.8%, 92.6%, and 91.4% in that order. The BE resulted in 92.6% with SVM and 92.1% with NN. BE with RF produced a 93.3% overall accuracy percentage.

In addition, three approaches for bathymetry mapping were presented. These approaches were applied over four diverse areas using Landsat 8 and Spot 6 satellite images. These areas

41

were selected to offer variation in studied surfaces, level of turbidity, and the water depths. The green, red, blue divided by red, and green divided by red band ratio logarithms corrected from atmospheric and sun-glint errors used as input data and water depth produced as output. To evaluate the presented methods, they were compared with the results of Lyzenga GLM and NN methods. All the produced results were also compared with field echosounder water depths values. The Lyzenga GLM correlation algorithm gave RMSE values of 0.96 m, 1.02 m, and 1.16 m, though the NN produced RMSE values of 0.87 m, 0.96 m, and 1.08 m in the three study areas, respectively. The proposed approaches, SVR, LSB, and BAG, yielded RMSE values of 0.92 m, 0.88 m, and 0.65 m for the first study area, 0.98 m, 0.99 m, and 0.85 m for the second study area, and 1.11 m, 1.09 m, and 0.80 m for the third study area, respectively.

Moreover, the second part of the study proposed RF and MARS approaches for bathymetry mapping. These approaches were applied over four diverse areas using Landsat 8, Landsat 7, and Spot 6 satellite images. These areas were selected to offer variation in water surfaces, level of turbidity, and the water depths. The green, red, blue divided by red, and green divided by red band logarithms corrected from atmospheric and sun-glint errors used as input data and water depth produced as output. To evaluate the presented methods, they were compared with the results of Lyzenga GLM and NN methods. All the produced results were also compared with field echosounder water depths values. Lyzenga GLM correlation algorithm gave RMSE values of 0.96 m and 1.16 m, though the NN produced RMSE values of 0.87 m and 1.08 m in the four study areas, respectively. The proposed approaches MARS and RF yielded RMSE values of 0.92 m and 0.65 m for the first study area, 0.98 m and 0.85 m for the second study area, and 1.11 m and 0.80 m for the fourth study area, in that order.

Finally, BE as a hybrid based approach for bathymetry detection were tested. This approach was applied in two diverse areas with different number of available field points: Alexandria port, Egypt and a part of Shiraho Island, Japan, with 12.5 m water depths. For NN and RF methods the green and red band logarithms corrected from atmospheric and sun-glint systematic errors of Landsat 8 and Quickbird satellite images were set as input data and water depths as output. The proposed approach ensemble the outputs from NN and RF approaches. To validate the improvement of BE proposed approach, it was compared with NN and RF results. Achieved results were also compared with echosounder water depths field data. NN yielded RMSE values of 0.92 m and 1.06 m, RF gave 0.64 m and 1.05 m, though the proposed BE approach produced RMSE values of 0.44 m and 0.96 m over the two

investigated areas, respectively. Approximately 20 cm and 10 cm increasing in the accuracy of detecting depths over a silt-sand area and coral reefs area, respectively.

CHAPTER 5
DISCUSSION AND CONCLUSIONS

5.1. Discussion for LULC classification

BE was applied with the three base classifiers RF, SVM, and NN in a hierarchal arrangement. The number of bagging trees after try and error was 10 trees, based on the maximum overall accuracy and minimum out-of-bag error. RF trees splitting was performed using Gini index diversity criterion. The optimum number of RF trees was calculated based on the random combinations of the three input variables and the maximum overall accuracy. The maximum overall accuracy was accomplished using 10 RF trees.

The commission and omission's errors presents the enhancements of the classification accuracy using BE with three base classifiers RF, SVM, and NN. All the classes commission and omission errors were decreased with the exception of the building class. The efficiency of BE in reducing the variance of unstable algorithms, especially RF and NN, is confirmed. BE increases the overall accuracy of the two base classifiers as RF with 0.5% and NN with 0.7% improvements.

For evaluating the complexity cost of proposed approach, two factors were tested its computational time and usage memory space. While BE improves the computational time and usage memory space with SVM and NN base classifiers to approximately 7 times in average, this problem can be solved. Dividing the study area into successive zones and increasing the memory reduce this demerit. Moreover, BE with RF base classifier had less computational time and usage space than BE with SVM and NN base classifiers.

5.2. Discussion for Bathymetry detection

To select the appropriate bands for bathymetry determination, a statistical analysis was performed to investigate the correlation between water depths and various imageries bands. This investigation demonstrated a strong correlation between the red and green bands with water depths [24]-[16]. Besides the red and green bands, the blue/red and green/red band ratios also demonstrated a strong correlation to the water depths.

The Lyzenga GLM model correlates the band combination directly to the water depth. In this research experiments, the best combination that achieved the lowest RMSE and highest R^2 values occurred between the green and red band logarithms as well as the blue/red and

green/red band ratios. Also, NNs were used to perform a correlation between the multilayers of the imagery bands as an input and water depth as an output through multidimensional nonlinear functions. The research results agree well with those of previous studies such as Ceyhun and Yalçın [14] and Gholamalifard et al. [28], which have argued that NNs outperform conventional models as the Lyzenga GLM model or the Stumpf ratio model. An NN suffers from one major issue in that it requires many trials to find the best weights for correlation as it is an unstable black-box approach having significant fluctuations in RMSE and R2 values from one trial to another.

The SVR algorithm, on the other hand, is a stable approach that uses minimum sequential optimisation to correlate the imagery bands with water depth. The optimum kernel function was selected, after several trials, from the radial basis function kernel, the polynomial kernel, and the Pearson universal kernel based on minimum RMSE and maximum R2. The latter outperformed the other kernel functions with the highest R2 and lowest RMSE. Also, the optimum SVR parameters, C, ε, ζ, ω, and σ, were selected based on the minimum RMSE criterion.

LSB and BAG are fitting ensembles of regression tree algorithms that have two different theories for collecting regression trees. LSB works sequentially by focusing on the missed regression values of the previous tree. On the contrary, the BAG ensemble averages regression trees built from a bootstrapped random selection from input data. For both ensembles, the optimum number of regression trees was selected after sequential trials of various numbers of trees (10, 20, 30... 100), and the best values were achieved with 50 trees. Both algorithms use the Gini diversity index for the splitting trees that are not pruned. The randomness of the regression trees and the splitting of the data into training and testing sets argues that the ensembles were not overfitting the input data. The results illustrate a preference of all proposed algorithms to Lyzenga GLM in addition to outperformance and greater stability of the BAG ensemble compared to the NN approach.

The MARS algorithm, on the other hand, is a stable approach that avoids overfitting and achieves comparable results to NN in the study areas. After several trials, the optimum number of BFs and GCVs were selected on the basis of the minimum RMSE and maximum R2.

The RF algorithm is a fitting ensemble of regression trees algorithm that averages regression trees built from a bootstrapped random selection from input data. The optimum number of regression trees was selected after sequential trials of various numbers of trees,

following which the best values were achieved with 50 decision trees. Also, the Gini diversity index was used for the splitting trees that were not pruned. Two RF regression tree ensembles were created and combined. The randomness of the regression trees and the splitting of the data into training and testing sets ensured that the ensembles did not overfit the input data.

Many researchers used low resolution satellite images for bathymetry detection especially Landsat images exploiting their free availability [28]-[16]-[15]-[60].

To compare achieved results with comparable studies many factors should be considered. These factors are images spatial resolutions, bottom type, water turbidity, availability of adequate number of field points in the study area, and depths range. For instance, Sánchez-Carnero et al. [16] confirmed the outperformance of Lyzenga GLM compared to principal component analysis (PCA) and green band correlation algorithms using Spot 4 imagery with 10 m resolution over turbid water in a shallow coastal area. The GLM yielded RMSE of 0.88 m in a depth range of 6 m. Pacheco et al. [15] tested the Landsat 8 coastal, blue, and green bands for bathymetry detection using Lyzenga GLM over clear waters in a shallow coastal area and achieved an RMSE of 1.01 m in a depth range of 12 m. Also, Gholamalifard et al. [28] supported the better performance of the NN approach compared to PCA and a red band correlation using Landsat 5 imagery over a deep water area. The research produced RMSE of 2.14 m in a depth range of 45 m. Kibele J. and Shears N. [61] proposed K-Nearest Neighbor (KNN) approach for detecting bathymetry over a clear coral reef area with large patches of sand and 20 m depths range. KNN method achieved RMSE of 0.8 m using Worldiew 2 satellite image better than Lyzenga linear method. However, a large number of field points, approximately 300000 points, were required for training the algorithm. Linda et al. [30] developed Neuro-Fuzzy approach for detecting bathymetry over a sandy bottom coastal area with clear water using Quickbird imagery. A RMSE of about 0.64 m has been achieved over a depths range of 14 m with small training samples and a slight sea conditions.

Our results are comparable to those of the studies for the NN and Lyzenga GLM approaches within the same depth ranges.

5.3. Conclusions

1) From the produced results, it can be concluded that BAG, LSB, and SVR approaches achieved more accurate results than Lyzenga GLM for bathymetry detection over four diverse study areas.

2) The outperformance of BAG approach compared to NN approach was confirmed over four diverse study areas.

3) From the produced results, it can be concluded that RF and MARS approaches achieved more accurate results than Lyzenga GLM for bathymetry detection over four diverse study areas.

4) The outperformance RF approach to NN approach was confirmed over four diverse study areas.

5) BE ensemble provide more accurate results than using single NN or RF approaches for bathymetry mapping over diverse areas.

6) From the presented classification results, BE enhances commission errors for the three base classifiers and decrease omission errors for RF and NN base classifiers also the precedence of BE with RF base classifier to other base classifiers, for instance SVM and NN was confirmed.

REFERENCES

[1] M. Salah, J. Trinder , A.Shaker, and M. Hamed, "Integrating Multiple Classifiers with Fuzzy Majority Voting for Improved Land Cover Classification," *ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. XXXVIII, no. WG III/2, pp. 7–12, 2010.

[2] J. Shlens, "A Tutorial on Principal Component Analysis," [Online]. Available: http://arxiv.org/abs/1404.1100, 2014, [Accessed: 20-Dec-2014].

[3] H. Chu and L. Ge, "Combination of Genetic Algorithm and Dempster-Shafer Theory of Evidence for Land Cover Classification using Integration of Sar and Optical Satellite Imagery," *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. Vol. XXXIX-B7*, vol. XXXIX, pp. 173–178, 2012.

[4] B. Samia, G. Li, and C. Nesrine, "Classification of Remote Sensing Data using Margin-Based Ensemble Methods," in *IEEE International Conference on Image Processing*, pp. 2601–2606, 2013.

[5] R. Dara, "Cooperative Training in Multiple Classifier Systems" , by Waterloo, Ontario, Canada, 2007.

[6] J. Benediktsson, J. Chanussot, and M. Fauvel, "Multiple Classifier Systems in Remote Sensing: From Basics to Recent Developments," in *Lecture Notes in Computer Science*, vol. 4472, and F. R. M. Haindl, J. Kittler, Ed. Springer-Verlag Berlin Heidelberg 2007, pp. 501–512.

[7] P. Du, J. Xia, W. Zhang, K. Tan, Y. Liu, and S. Liu, "Multiple Classifier System for Remote Sensing Image Classification: A review," *Sensors*, vol. 12, no. 4, pp. 4764–4792, 2012.

[8] H. Guan, J. Yu, J. Li, and L. Luo, "Random Forests-Based Feature Selection for Land-Use Classification Using Lidar Data and Orthoimagery," *XXII ISPRS Congr.*, vol. XXXIX, pp. 203–208, 2012.

[9] Z. Zheng, "Boosting and Bagging of Neural Networks with Applications to Financial Time Series," Working paper, Department of Statistics, University of Chicago, pp. 1–27, 2006.

[10] Ö. Akar and O. Güngör, "Classification of Multispectral Images using Random Forest Algorithm," *J. Geod. Geoinf.*, vol. 1, no. 2, pp. 105–112, 2012.

[11] S. Kulkarni and V. Kelkar, "Classification of Multispectral Satellite Images Using Ensemble Techniques of Bagging , Boosting and Ada- Boost," in *International Conference on Circuits, Systems, Communication and Information Technology Applications (CSCITA) Classification*, pp. 253–258, 2014.

[12] L. Leu and H. Chang, "Remotely Sensing in Detecting the Water Depths and Bed Load of Shallow Waters and Their Changes," *Ocean Eng.*, vol. 32, no. 10, pp. 1174–1198, Jul. 2005.

[13] J. Gao, "Bathymetric Mapping by Means of Remote Sensing : Methods , Accuracy and Limitations," *Prog. Phys. Geogr.*, vol. 33, no. 1, pp. 103–116, 2009.

[14] Ö. Ceyhun and A. Yalçın, "Remote Sensing of Water Depths in Shallow Waters via Artificial Neural Networks," *Estuar. Coast. Shelf Sci.*, vol. 89, no. 1, pp. 89–96, 2010.

[15] A. Pacheco, J. Horta, C. Loureiro, and Ó. Ferreira, "Retrieval of Nearshore Bathymetry from Landsat 8 Images: A tool for Coastal Monitoring in Shallow Waters," *Remote Sens. Environ.*, vol. 159, pp. 102–116, 2015.

[16] N. Sánchez-Carnero, J. Ojeda-Zujar, D. Rodríguez-Pérez, and J. Marquez-Perez, "Assessment of Different Models for Bathymetry Calculation using SPOT Multispectral Images in a High-Turbidity Area: the Mouth of the Guadiana Estuary," *Int. J. Remote Sens.*, vol. 35, no. 2, pp. 493–514, 2014.

[17] G. Chust, M. Grande, I. Galparsoro, A. Uriarte, and Á. Borja, "Capabilities of the Bathymetric Hawk Eye LiDAR for Coastal Habitat Mapping: A case study within a Basque estuary," *Estuar. Coast. Shelf Sci.*, vol. 89, no. 3, pp. 200–213, Oct. 2010.

[18] M. Lesser and C. Mobley, "Bathymetry, Water Optical Properties, and Benthic Classification of Coral Reefs using Hyperspectral Remote Sensing Imagery," *Coral Reefs*, vol. 26, no. 4, pp. 819–829, 2007.

[19] V. Brando, J. Anstee, M. Wettle, A. Dekker, S. Phinn, and C. Roelfsema, "A Physics Based Retrieval and Quality Assessment of Bathymetry from Suboptimal Hyperspectral Data," *Remote Sens. Environ.*, vol. 113, no. 4, pp. 755–770, 2009.

[20] E. Vahtmäe and T. Kutser, "Airborne Mapping of Shallow Water Bathymetry in the

Optically Complex Waters of the Baltic Sea," *J. Appl. Remote Sens.*, vol. 10, no. 2, pp. 1–16, 2016.

[21]   D. Lyzenga, "Shallow-Water Bathymetry using Combined Lidar and Passive Multispectral Scanner Data," *Int. J. Remote Sens.*, vol. 6, no. 1, pp. 115–125, 1985.

[22]   D. Lyzenga, N. Malinas, and F. Tanis, "Multispectral Bathymetry using a Simple Physically Based Algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 8, pp. 2251–2259, 2006.

[23]   A. Knudby, C. Roelfsema, M. Lyons, S. Phinn, and S. Jupiter, "Mapping Fish Community Variables by Integrating Field and Satellite Data, Object-Based Image Analysis and Modeling in a Traditional Fijian Fisheries Management Area," *Remote Sens.*, vol. 3, no. 3, pp. 460–483, 2011.

[24]   G. Doxani, M. Papadopoulou, P. Lafazani, C. Pikridas, and M. Tsakiri-Strati, "Shallow-Water Bathymetry Over Variable Bottom Types Using Multispectral Worldview-2 Image," *ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. XXXIX-B8, no. pp. 159–164, 2012.

[25]   R. Stumpf, K. Holderied, and M. Sinclair, "Determination of Water Depth with High-Resolution Satellite Imagery over Variable Bottom Types," *Limonology Oceanogr.*, vol. 48, no. 1_part_2, pp. 547–556, 2003.

[26]   H. Su, H. Liu, and W. Heyman, "Automated Derivation of Bathymetric Information from Multi-Spectral Satellite Imagery Using a Non-Linear Inversion Model," *Mar. Geod.*, vol. 31, no. May, pp. 281–298, 2008.

[27]   J. Bramante, D. Raju, and T. Sin, "Multispectral Derivation of Bathymetry in Singapore's Shallow, Turbid Waters," *Int. J. Remote Sens.*, vol. 34, no. 6, pp. 2070–2088, 2013.

[28]   M. Gholamalifard, T. Kutser, A. Esmaili-Sari, A. Abkar, and B. Naimi, "Remotely Sensed Empirical Modeling of Bathymetry in the Southeastern Caspian Sea," *Remote Sens.*, vol. 5, no. 6, pp. 2746–2762, 2013.

[29]   S. Moses, L. Janaki, S. Joseph, J. Gomathi, and J. Joseph, "Lake Bathymetry from Indian Remote Sensing (P6-LISS III) Satellite Imagery using Artificial Neural Network Model," *Lakes Reserv. Res. Manag.*, vol. 18, no. 2, pp. 145–153, 2013.

[30] L. Corucci, "Approaching Bathymetry Estimation from High Resolution Multispectral Satellite Images using a Neuro-Fuzzy Technique," *J. Appl. Remote Sens.*, vol. 5, no. 1, p. 53515, 2011.

[31] H. Mohamed, A. Negm, M. Zahran, and S. Abdel-Fattah, "Estimation of Bathymetry Using High-resolution Satellite Imagery: Case Study El-Burullus Lake, Northern Nile Delta," in *The Handbook of Environmental Chemistry*, Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1–30, 2016.

[32] J. Bramante, D. Raju, and T. Sin, "Multispectral Derivation of Bathymetry in Singapore's shallow, Turbid Waters," *Int. J. Remote Sens.*, vol. 34, no. 6, pp. 2070–2088, 2013.

[33] H. Mohamed, A. Negm, M. Salah, K. Nadaoka, and M. Zahran, "Assessment of Proposed Approaches for Bathymetry Calculations using Multispectral Satellite Images in Shallow Coastal/Lake Areas: a Comparison of Five Models," *Arab. J. Geosci.*, vol. 10, no. 2, p. 42, 2017.

[34] H. Mohamed, A. Negm, M. Zahran, and O. Saavedra, "Assessment of Ensemble Classifiers using Bagging technique for Improved Land Cover Classification of Multispectral Satellite Images," *Int. Arab J. Inf. Technol.*, vol. 15, no. 3, 2017.

[35] H. Mohamed, A. Negm, K. Nadaoka, T. Abdelaziz, and M. Elsahabi, "Comparative Study of Approaches to Bathymetry Detection in Nasser / Nubia Lake using Multispectral SPOT-6 Satellite Imagery," *Hydrol. Res. Lett.*, vol. 10, no. 1, pp. 45–50, 2016.

[36] H. Mohamed, A. Negm, M. Zahran, and O. Saavedra, "Bathymetry Determination from High Resolution Satellite Imagery Using Ensemble Learning Algorithms in Shallow Lakes: Case Study El-Burullus Lake," *Int. J. Environ. Sci. Dev.*, vol. 7, no. 4, pp. 295–301, 2016.

[37] M. Ali, "Impact of Drain Water on Water Quality and Eutrophication Status of Lake Burullus, Egypt, a Southern Mediterranean Lagoon," *African J. Aquat. Sci.*, vol. 36, no. 3, pp. 267–277, 2011.

[38] A. Collin, K. Nadaoka, and T. Nakamura, "Mapping VHR Water Depth, Seabed and Land Cover Using Google Earth Data," *ISPRS Int. J. Geo-Information*, vol. 3, no. 4, pp. 1157–1179, 2014.

[39] C. Cortes and V. Vapnik, "Support-Vector Networks," *Mach. Learn.*, vol. 20, pp. 273–297, 1995.

[40] T. Kavzoglu and I. Colkesen, "A Kernel Functions Analysis for Support Vector Machines for Land Cover Classification," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 11, no. 5, pp. 352–359, 2009.

[41] A. Elhadi, R. Ismail, and O. Mutanga, "A Comparison of Selected Machine Learning Classifiers in Mapping a South African Hetrogeneous Coastal Zone: Testing the Utility of An Object-Based Classification with WorldView-2 Imagery," *PROCEEDINGS-SPIE Int. Soc. Opt. Eng. Earth Resour. Environ. Remote Sensing/GIS Appl. III Vol 8538*, vol. 8538, p. 85380P–1 to 85380P–10, 2012.

[42] Landsat-8, "Using the USGS Landsat 8 Product," 2013. [Online]. Available: http://landsat.usgs.gov/Landsat8_Using_Product.php. [Accessed: 20-Jan-2016].

[43] A. Berk, L. Bernstein, G. Anderson, P. Acharya, D. Robertson, J. Chetwynd, and S. Adler-Golden, "MODTRAN Cloud and Multiple Scattering Upgrades with Application to AVIRIS," *Remote Sens. Environ.*, vol. 65, no. 3, pp. 367–375, Sep. 1998.

[44] D. Hedley, R. Harborne, and J. Mumby, "Technical note: Simple and robust removal of sun glint for mapping shallow-water benthos," *Int. J. Remote Sens.*, vol. 26, no. 10, pp. 2107–2112, 2005.

[45] P. Casale, O. Pujol, and P. Radeva, "Embedding Random Projections in Regularized Gradient Boosting Machines," in *Ensembles in Machine Learning Applications*, Berlin Heidelberg: Springer-Verlag Berlin Heidelberg, pp. 201–216, 2011.

[46] A. Ihler, "Machine Learning and Data Mining,Ensembles of Learners," California, 2012.

[47] B. Ghimire, J. Rogan, V. Rodriguez-Galiano, P. Panday, and N. Neeti, "An Evaluation of Bagging , Boosting , and Random Forests for Land-Cover Classiication in Cape Cod ,Massachusetts, USA," *GIScience Remote Sens.*, vol. 5, no. 5, pp. 623–643, 2012.

[48] M. Galar, A. Fern, E. Barrenechea, and H. Bustince, "A Review on Ensembles for the Class Imbalance Problem: Bagging-, Boosting-, and Hybrid-Based Approaches," *IEEE Trans. Syst. MAN, Cybern. C Appl. Rev.*, vol. 42, no. 4, pp. 463–484, 2012.

[49] J. Smola and B. Schölkopf, "A Tutorial on Support Vector Regression," *Stat. Comput.*, vol. 14, no. 3, pp. 199–222, 2004.

[50] A. Farag and R. Mohamed, "Regression Using Support Vector Machines : Basic Foundations," Louisville, pp. 1-17, 2004.

[51] B. Razavi, "Predicting the Trend of Land Use Changes Using Artificial Neural Network and Markov Chain Model (Case Study: Kermanshah City)," *Res. J. Environ. Earth …*, vol. 6, no. 4, pp. 215–226, 2014.

[52] L. Breiman, "Random Forests," *J. Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[53] H. Kim and G. Sohn, "Random Forests Based Multiple Classifier System for Power-Line Scene Classification," *Int. Arch. Photogramm.*, vol. XXXVIII, pp. 29–31, 2011.

[54] O. Kisi and K. Parmar, "Application of Least Square Support Vector Machine and Multivariate Adaptive Regression Spline Models in Long Term Prediction of River Water Pollution," *J. Hydrol.*, vol. 534, pp. 104–112, 2016.

[55] S. Emamgolizadeh, S. Bateni, D. Shahsavani, T. Ashrafi, and H. Ghorbani, "Estimation of Soil Cation Exchange Capacity using Genetic Expression Programming (GEP) and Multivariate Adaptive Regression Splines (MARS)," *J. Hydrol.*, vol. 529, pp. 1590–1600, 2015.

[56] C. Conoscenti, M. Ciaccio, N. Caraballo-Arias, Á. Gómez-Gutiérrez, E. Rotigliano, and V. Agnesi, "Assessment of Susceptibility to Earth-Flow Landslide using Logistic Regression and Multivariate Adaptive Regression Splines: A case of the Belice River basin (western Sicily, Italy)," *Geomorphology*, vol. 242, pp. 49–64, 2015.

[57] R. DeFries and J. Chan, "Multiple Criteria for Evaluating Machine Learning Algorithms for Land Cover Classification from Satellite Data," *Remote Sens. Environ.*, vol. 74, no. 3, pp. 503–515, 2000.

[58] R. Clark, "A MATLAB Implementation of Support Vector Regression (SVR)," *Support Vector Regression - File Exchange - MATLAB Central*, 2013. [Online]. Available: http://www.mathworks.com/matlabcentral/fileexchange/43429-support-vector-regression. [Accessed: 20-Feb-2015].

[59] G. Jekabsons, "ARESLab: Adaptive Regression Splines Toolbox for Matlab/Octave," 2016, [Online]. Available: http://www.cs.rtu.lv/jekabsons/. [Accessed: 15-May-2016].

[60] M. Salah, "Determination of Shallow Water Depths using Inverse Probability Weighted Interpolation: A Hybrid System- Based Method," *Int. J. Geoinformatics*, vol. 12, no. 1, pp. 45–55, 2016.

[61] J. Kibele and N. Shears, "Nonparametric Empirical Depth Regression for Bathymetric Mapping in Coastal Waters," *IEEE J. Sel. Top. Appl. EARTH Obs. Remote Sens.*, vol. 9, no. 11, pp. 5130–5138, 2016.

ملخص

الهدف الرئيسى من هذة الرسالة هو تطوير نظام لتقييم و متابعة تصنيف استخدام الاراضى و تقييم الاعماق للمناطق الساحلية و البحيرات باستخدام صور الاقمار الصناعية.

فالجزء الاول من الرسالة يقدم نموذج مقترح لتصنيف الاراضى باستخدام صور الاقمار الصناعية. تم عمل هذا الجزء من الدراسة على منطقة بحيرة البرلس و هى منطقة ساحلية متنوعة و غير متجانسة و تم تصنيفها لخمس انواع من الغطاء الارضى شملت المبانى و المسطحات المائية و الاراضى الزراعية و الاراضى الرملية و الحشائش. النموذج المقترح يستخدم المجمع Bagging Ensemble لل Random Forest كمصنف رئيسى لتقليل الاخطاء الناتجة عن التباين فى النتائج.وقد تم اقتراح أسلوب قائم على تصنيف كل خلية على حدة بطريقة التعلم المسبق و اختيار العناصر المناسبة للتصنيف من السمات المتاحة بواسطة طريقة Principal Component Analysis باستخدام مرئيات القمر الصناعي لاندسات 8. وكانت العناصر المستخدمة من الصورة هى الاطوال الموجية الساحلية و المرئية و تحت الحمراء و الحرارية بالاضافة الى مؤشر الفارق المعيارى للغطاء النباتى و الغطاء المائى لتقييم دقة طريقة التصنيف المقترحة تم مقارنتها مع طريقتين سابقتين يستخدمان مجمع ال Bagging Ensemble مع طريقتى ال Support Vector Machines و ال Neural Network. تم تقييم النتائج وفقا لاخطاء الاضافة و النقصان والدقة الشاملة. و بينت النتائج تفوق الطريقة المقترحة على الطرق السابقة بدقة كلية بلغت نسبتها 93.3 % . اما الطرق السابقة فكانت نسبة دقتها 92.6% و 92.1% على التوالى. هذة النسب تؤكد تفوق الطريقة المقترحة على الطرق السابقة.

ومن ناحية أخرى اشتمل الجزء الثاني من الدراسة تقييم أداء ثلاثة نماذج مقترحة هي خوارزمية Bagging, Least Square Boosting, and Support Vector Regression لحسابات قياس الأعماق في أربعة مناطق مختلفة كالتالى: المنطقة الساحلية الضحلة من مدخل البرلس، مصر، وهي منطقة ذات قاع رملي عكر مع أعماق تصل إلى 6 م؛ منطقة ميناء الإسكندرية الساحلية الضحلة، كمثال على العكارة المنخفضة وهى منطقة ذات قاع رملي مع أعماق تتراوح بين 4 م إلى 10.5 م؛ منطقة مدخل بحيرة النوبة، السودان، والتي تعتبر ذات عكارة عالية، غير مستقرة، منطقة ذات قاع من الطين بعمق 6 م؛ و شيراهو، جزيرة إشيغاكي، اليابان، و هى منطقة شعاب مرجانية مع عمق يصل الى 14 م. تم استخدام مرئيات الاقمار الصناعية لاندسات 8 و مرئيات سبوت 6 لتقييم أداء النماذج المقترحة. وقد استخدمت هذه النماذج المقترحة للحصول على خرائط قياس الأعماق باستخدام انعكاس الاطياف الاخضر والأحمر و النسبة بين الأزرق و الأحمر إلى جانب النسبة بين الطيف الاخضر و الاحمر. تمت مقارنة نتائج قياس الأعماق للنماذج المقترحة مع النتائج المقابلة من طريقتين سابقتين تقليديتين Neural Network and Lyzenga Linear approaches. وبالمقارنة مع نقاط القياس الارضى ، كشفت نتائج الطرق الثلاثة المقترحة عن دقة أعلى تتراوح من 0.04 إلى 0.35 متر أكثر من طريقة لايزنجا الخطية و كانت نتائج خوارزمية Bagging Ensemble الأكثر دقة.

وبالإضافة إلى ذلك، فإن الجزء الثالث من الدراسة اقترح نموذجى Random Forest and Multi-Adaptive Regression Spline (MARS) لانتاج خرائط قياس الأعماق من مرئيات الاقمار الصناعية. تم استخدام بيانات من لاندسات 7، لاندسات 8، و سبوت 6 لتقييم أداء هذه النماذج. وقد استخدمت هذه النماذج للحصول على خرائط الأعماق باستخدام نفس المدخلات المذكورة أعلاه. تم اختبار الخوارزميات على مناطق الدراسة المذكورة أعلاه باستثناء مناطق البرلس التي تم استبدالها بمنطقة مدخل النوبة باستخدام صورة لاندسات 7. تمت مقارنة النتائج مع نفس الطريقتين

السابقتين التقليديين Neural Network and Lyzenga Linear approaches. بالمقارنة مع النقاط الحقلية كانت نتائج الطريقتين المقترحتين افضل من طريقة لايزنجا الخطية. وعلاوة على ذلك، أسفرت طريقة ال Random Forest عن نتائج أكثر دقة بمتوسط قدره 0.25 متر من التحسينات افضل من طريقة ال Neural Network.

وأخيرا، فإن الجزء الرابع من هذه الدراسة اقترح استخدام طريقة هجين للدمج يستند إلى طريقة ال Bagging Ensemble لعمل خرائط قياس الأعماق. تم تطبيق هذه الطريقة في منطقتين متنوعتين مع عدد مختلف من النقاط الميدانية المتاحة: ميناء الإسكندرية، مصر، وجزء من جزيرة شيراهو باليابان، بعمق 12.5 م من المياه. وفيما يتعلق بالطرق Random Forest and Neural Network تم استخدام الاطياف الاخضر والأحمر اللتين تم تصحيحهما من الأخطاء المنهجية من الغلاف الجوى و الوميض الشمسي في صور القمر الصناعي لاندسات 8 و كويكبيرد كبيانات للإدخال وأعماق المياه كنواتج. وللتحقق من صحة تحسين النهج المقترح تمت مقارنته بنتائج نفس الطرق منفردة. كما تم مقارنة النتائج المحققة مع البيانات الحقلية لأعماق المياه. من النتائج المحققة كان نسبة التحسن حوالي 20 سم و 10 سم في دقة الكشف عن ألاعماق فى منطقتي الدراسة المذكورتين، على التوالي. ونتيجة لذلك، يمكن استنتاج أن طريقة دمج النتائج المقترحة قد أنتجت خرائط أكثر دقة من استخدام طريقة منفردة سواء ل Random Forest او Neural Network .

تقييم تصنيف استخدام الاراضي و الغطاء الارضي و تحديد الاعماق الضحلة للمناطق الساحلية و البحيرات باستخدام خوارزميات التعلم الالي

رسالة علمية

مقدمة الى كلية الدرسات العليا

فى مصادر الطاقة والبيئة والهندسة الكيميائية والبتروكيميائية

الجامعة المصرية اليابانية للعلوم والتكنولوجيا

كاستيفاء جزئي لمتطلبات الحصول على درجة دكتوراه فى الفلسفة

فى

الهندسة البيئية

مقدمة من

حسن محمد حسن حسين

سبتمبر 2017