International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys.13, No.1 (Mar-23)

http://dx.doi.org/10.12785/ijcds/130136

Calculating Spectral Index Based on Linear SVM Methods for Landsat OLI: Baiji Sand Dunes a Case Study, Iraq

Ehsan Ali Al-Zubaidi[1,2], Furkan Rabee[1] and Ahmed H. Al-Sulttani[2]

[1]*Department of Computer Science,Faculty of Computer Science and Mathematics, University of Kufa, Najaf, Iraq*
[2]*Department of Environmental Planning,Faculty of Physical Planning, University of Kufa, Najaf, Iraq*

Received 27 Jan. 2022, Revised 15 Sep. 2022, Accepted 7 Mar. 2023, Published 16 Mar. 2023

Abstract: Machine learning and remote sensing technologies can effectively investigate and monitor the dynamic features of dunes, sand accumulations, and environmental changes. However, the present article examines the performance of models constructed using various Linear SVM model implementations during the binary classification task. In R software, three linear SVM libraries (LibSVM, LibLINEAR, and GLM) are used to find the optimal and accurate spectral sand index. The data of Landsat-8 (OLI) satellite is used and calibrated to reflectance value, and then 15 sequences of Normalized Difference (ND) are generated. The most important and weighted NDs among the 15 trained have been chosen (ND34, ND47, and ND57). Nine Linear SVM methods have been calculated. LibLINEAR library contains several classification types (LibLINEAR1 to LibLINEAR7). The accuracy of the result images is done by assigning random points to six levels 500 to 1000 points according to the reference image created by supervised classification. The average Kappa and overall accuracy for all levels of random points show that the optimal three methods are LibL1, LibL7, and LibL3; with Kappa values of 77.20%, 76.96%, and 76.83%, respectively, and overall accuracy values of 88.60%, 88.48%, and 88.41%, respectively. In contrast, the widely used LibSVM shows less accuracy with more execution time than LibLINEARs.

Keywords: SVM, LibSVM, LibLINEAR, GLM, Remote Sensing, Sand Dunes Index

1. INTRODUCTION

Machine learning applications such as genetics and image processing have grown more popular. Modeling and extracting characteristics from remote sensing have benefited from machine learning approaches [1][2][3]. Statistical machine learning (ML) techniques can be used to analyze remote sensing (RS) data [4]. The Random Forest RF and Support Vector Machine SVM classifiers have enhanced remote sensing imagery classifying during the previous two decades[5]. The classification accuracy of SVM and RF classifiers is comparable to that of Convolution Neural Network (CNN) and other machine learning classifiers often used in remote sensing data analysis [6][7]. SVM and RF efficiently process and model multi-dimensional data with satisfactory classification results [8].

Many studies have been implemented to specify the best classifier by machine learning for mapping the Land use and Land cover (LULC) [9][10][11][12]. According to several research, SVM and RF are better than other machine learning methods for classifying LULC [13][14]. The classification accuracy of LULC is dependent on several factors, such as spatial and temporal resolution and the processing of imagery data [15]. The medium-resolution sensors are better and suitable for detecting most human-nature features. In contrast, the low to medium-resolution sensors are insufficient for regional and large-scale studies [16]. Landsat 8 Operational Land Imager (OLI) and Sentinel-2 Multispectral Instrument (S2/MSI) are industry-standard satellites for medium-resolution land imagery [17]. They constitute cost-effective methods for describing large-scale landscape processes [18].

The increase in spatial resolution, open access policies, and systematic global coverage have ushered in a new era of LULC and Land Use and Land Cover Classification (LULCC) application development [19][20]. This factor has facilitated the handling and analysis of large Earth spatiotemporal datasets comprised of multi-source data with Landsat-like resolution [21]. Finding an optimal LULC classifier for Earth observation applications required testing the accuracy of several techniques [22]. Various spectral indices have been created to monitor and study desertification in arid and semi-arid areas. Desertification can be mapped based on remotely sensed imagery using standard classification techniques, spectral indices, or spectral mix-

E-mail address: ihsana.kareem@uokufa.edu.iq, furqan.rabee@uokufa.edu.iq, ahmedh.alsulttani@uokufa.edu.iq

ture analysis [23][24][25].

Normalized Difference Index (NDI) is a multi-spectral satellite image classification technique to detect and identify a specific phenomenon and earth feature. NDI is generally designed as the ratio of the difference between reflectance values in two bands or more and the sum of the same bands' values:

$$NDI = \frac{(p1 - p2)}{(p1 + p2)} \quad (1)$$

p1 and *p2* are the two specified reflectance bands. For example, *p1* and *p2* represent the reflectance of Near-infrared NIR and Red bands, respectively, in the case of the Normalized Difference Vegetation Index (NDVI). At the same time, *p1* and *p2* refer to Green and Near-Infrared in Normalized Difference Water Index NDWI. Besides, several indices for the phenomenon of dunes were developed [26][27][28][29]. Many applications identify a specific phenomenon by applying a threshold to a calculated NDI. For example, in NDVI, the threshold value represents the separator value of the pixels, representing the presence of vegetation cover. If the values of the pixels are greater than the threshold value, it means the presence of vegetation and vice versa [30].

Earth observation, land-use monitoring, and following up on the changes are very important. Therefore, it is necessary to provide an easy, fast, and high-accuracy method for classifying satellite images [31][32]. The main objective of this study is to evaluate and choose the optimum classification technique for sand dune data by comparing various SVM classification approaches. Baiji sand dunes field was chosen as a case study. Baiji field is one of the main sand dunes fields in Iraq. The proposed spectral index will help monitor the sand dunes encroachment and expansion in such arid and semi-arid regions.

2. RELATED WORKS

SVM is a kernel-based method for determining the best hyperplane for splitting data. SVM is used in various applications for classifying remote sensing data. This supervised learning approach solves binary classification problems. Many remote sensing research studies used the library (LibSVM) to classify LULC [33][34][35][36][37]. A set of previous studies using the SVM algorithm with different kinds of the dataset will be presented and discussed.

Horvath et al. [38] explain the reason and function of a Unix script-driven program for evolutionary searching of optimum SVM model parameters as calculated by the LibSVM library, resulting in support vector machine models with the highest predictive value and stability.

Wang et al. [33], 2018 Developed a water index was developed to classify images from the Sentinel-2 satellite with a resolution of 10 meters. The proposed index works on high-resolution images with small cut sizes, and the results were compared with other water indices. The SVM algorithm was used to extract the new index equation. The results were tested on three small study areas. The total number of pixels selected for training is (48,821) pixels, and the size of the reference data site is 300 m-by-300 m.

Ni et al. [39], The support vector machine (SVM) approach was used in this study. The eigenvectors were used in the conversion of the original data. MATLAB Library for support vector machines was used to identify three fault types (leakage of shift clutch hydraulic cylinder, obstruction in oil flow, and obstruction in proportional valve spool). The classification accuracy of test samples was 90%.

Zhang et al. [40], Electronic tongue (e-tongue) and electronic nose (e-nose) were designed and used in the study to assess the different ages of rice wines. E-tongue, E-nose, and Direct Fusion Data sets were used to identify rice wines of different ages. Sets of Weighted Fusion and Optimized Direct Fusion Data were used with partial least squares regression, extreme learning machine, and support vector machines (LibSVM). All algorithms performed well with solid prediction results, but LibSVM had the best correlation coefficient.

Gonzalez-Lima and Ludeña [37], presented two algorithms: theoretical support and numerical tests to show their performance on real-world issues by using LibSVM data. The optimization issue that develops in the training stage of support vector machines for huge data sets, possibly in high dimensions, may be solved using a new approach based on Locality-Sensitive Hashing (LSH).

Chou et al. [41], suggested Heuristic intelligence, decomposition methodologies, the one-against-one (OAO) method, and the least-squares support vector machine (LSSVM). Improved Firefly Algorithm FA was used to tune the LSSVM hyperparameters automatically, resulting in an optimum classification model. The updated optimization approach is tested against a set of ten test functions. Geotechnical engineering datasets related to seismic bumps and soil liquefaction are then utilized to demonstrate how the suggested method may be used to solve binary issues in the geotechnical engineering field.

3. STUDY AREA

The study area is located in north-central Iraq and north of Tikrit province. It is a natural area bounded by Hamrin hills from the north and north-east, Tigris River from the west and south, and the Al-Udhaim River from the east and southeast. Geographically, the study area is bounded by lat. 35.00° 33.50°N and long. 43.40°E 44.30°E Figure 1.

Iraq is suffering from the sequence of climate changes because of its fragile arid and semi-arid environment. Due to desertification and climate changes, Iraq is losing Oss 100 square kilometers of its arable lands annually [42]. This year, Iraq is the most country in the region affected by the dust storm frequency increase [43]. The source of dust storm in the region are the specific areas inside and

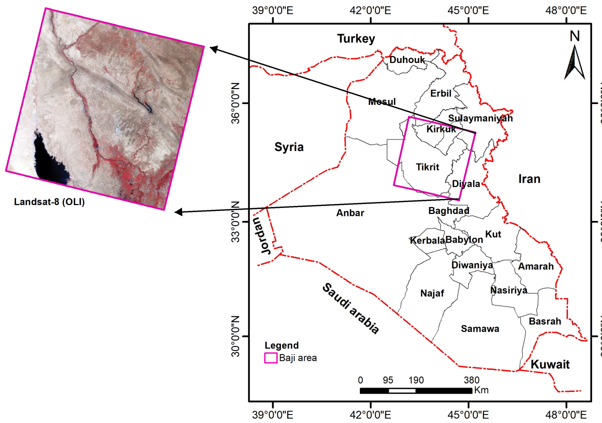

Figure 1. Study area and the Landsat-8 scene location

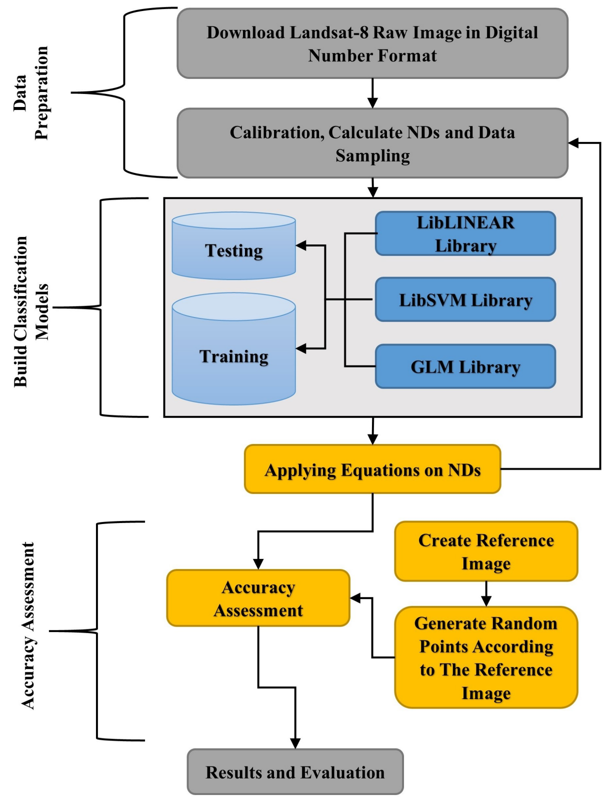

Figure 2. Methodological workflow

outside Iraq called hotspots, and Baiji sand dunes area is one of the 6 hotspots in Iraq [42][44][45]. Hence, there is an urgent need to conduct studies and find techniques that facilitate the continuous monitoring and assessment of sand dune areas.

4. MATERIALS AND METHODS

The methodology workflow consists of three main stages: data preparation, model construction, and accuracy assessments: Image calibration, normalized differences (NDs) calculation, and data sampling were essential tasks throughout the data preparation phase. The workflow in Figure 2 illustrates the main stages used in this paper.

Nine linear classification algorithm techniques are implemented in the second step. The reason for using a linear SVM is because it separates a dataset linearly into two categories by a hyperplane. In addition, the linear SVM calculates weights for each feature, which is necessary for determining which characteristics best represent the occurrence of phenomena [33].

After classifying the data, each classifier's linear equations are extracted and applied to the NDs. In the final step, the accuracy score of each technique is computed using three metrics: Overall Accuracy (OA), (Kappa), and F-score. In the end, the results are compared, the nine procedures are evaluated based on the results, and the most accurate linear equation is selected.

A. Materials

For this study, a Landsat-8 image was collected and downloaded from the United States Geological Survey's (USGS) website (https://earthexplorer.usgs.gov). Landsat 8 raw imagery acquisition Date (2018-08-16), Path (169), Row (36), Width (7881 pixels), and Height (7881 pixels). Six Landsat-8 bands are selected in total, Blue, Green, Red, NIR, SWIR1 and SWAR2.

Landsat-8 OLI raw imagery was acquired in the GeoTag Image file format (GeoTIFF), and the values of pixels in Digital Numbers DN. DN values need to be radiometrically calibrated to convert them to reflectance for all bands using ENVI software. The total pixels in each Landsat-8 scene are about 62 (180 x 180 km) with a spatial resolution of 30m. Landsat-8 scene must be radiometrically calibrated. The first enhancement of the satellite image is by converting it to reflectance.

Due to the huge data, sampling of data has been suggested. Data sampling is done by collecting samples from each phenomenon in the image LULC (Vegetation, Bare land, Sand Dunes, Water, and Built-up) through the selection tool (polygon) in ArcGIS software. Fifteen NDs have been calculated to extract the NDs value according to the sampling polygons. The NDs were calculated among six selected bands Table I.

Labeling process, all sampled pixels are combined into a single table. The dependent variable Class Y is created by giving a (+1) label to the pixels of the dune dunes and Class X (-1) to the pixels that do not represent the feature of the dune Table II. Finally, labeled pixels with NDs values are exported in CSV file format for processing. The six bands will be represented by b1, b2, b3, b4, b5, and b6. The data of studied pixels are divided into training pixels (83733) and testing pixels (41866). The total number of pixels used in the study is (125599).

TABLE I. NORMALIZED DIFFERENCES BETWEEN BANDS

Feature	Equations
ND23	(Float("b2"-"b3"))/(Float("b2"+"b3"))
ND24	(Float("b2"-"b4"))/(Float("b2"+"b4"))
ND25	(Float("b2"-"b5"))/(Float("b2"+"b5"))
ND26	(Float("b2"-"b6"))/(Float("b2"+"b6"))
ND27	(Float("b2"-"b7"))/(Float("b2"+"b7"))
ND34	(Float("b3"-"b4"))/(Float("b3"+"b4"))
ND35	(Float("b3"-"b5"))/(Float("b3"+"b5"))
ND36	(Float("b3"-"b6"))/(Float("b3"+"b6"))
ND37	(Float("b3"-"b7"))/(Float("b3"+"b7"))
ND45	(Float("b4"-"b5"))/(Float("b4"+"b5"))
ND46	(Float("b4"-"b6"))/(Float("b4"+"b6"))
ND47	(Float("b4"-"b7"))/(Float("b4"+"b7"))
ND56	(Float("b5"-"b6"))/(Float("b5"+"b6"))
ND57	(Float("b5"-"b7"))/(Float("b5"+"b7"))
ND67	(Float("b6"-"b7"))/(Float("b6"+"b7"))

TABLE II. TRAINING PIXELS NUMBERS

LULC	Pixels number	Class Label
Sand	60664	+1
Bare land	25184	-1
Built-up	5114	-1
Vegetation	14637	-1
Water	20000	-1
Total	125599	+1and-1

B. Methods

SVM was initially created as a supervised binary classifier. A classifier uses a model to determine whether a sample belongs to Class (label +1) or Class (label -1). The basic principle of SVM classifier training is to identify the maximal-margin hyperplane that separates Class +1, and Class -1, which referred known as the functional margin. The support vectors are the samples that help define such hyperplanes; see the circled samples lying on the solid hyperplanes displayed in Figure 3 [46][47].

Three libraries (LibSVM, LibLINEAR, and GLM) have been used to construct classification models for extracting linear equations. The LibSVM model was trained using the e1071 package [48][49], and the seven LibLINEAR models were trained using the LibLINEAR package [50]. The GLM model was trained in GLM package. All models were created using the R programming language. Training an SVM classifier is a simple case of finding the hyperplane with the largest possible functional margin, separating Class (+1) and Class (-1) data.Figure 3 illustrates a support vector where circular samples are placed on solid hyperplanes to define them better.

Before beginning the training process, it is necessary to partition the data into three sections: two for training and one for testing. Testing ensures that the algorithms

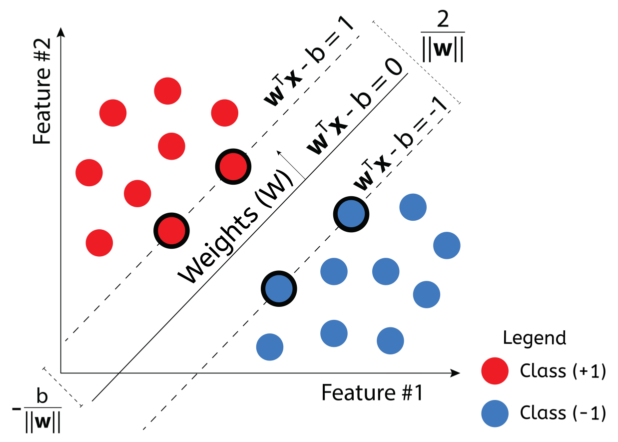

Figure 3. Hard margin SVM

are implemented correctly and have comparable results. It is essential to consider modifying the values of specific parameters to provide a fair comparison. As mentioned, (125599) pixels were used in this study, (83733) pixels for training, and (41866) pixels for testing.

The ND numbers are reduced by focusing on the NDs with the highest weight Figure 3. The most weighted NDs among the 15 trained NDs are (ND34, ND47, and ND57). Once the NDs were reduced, the remaining classification methods were applied, and the linear equation for the sand dune index was calculated and compared. The processing was done using three SVM libraries: LibLINEAR, the logistic regression technique (GLM) library, and (LibSVM). LibLINEAR is currently supports (1) L1-regularized L2-loss SVC, (2) L1-regularized L2-loss SVC, (3) L1-loss SVC, (4) L2-loss SVC, (5) L2-regularized LR, (6) L1-regularized LR (7) L2-regularized L2-loss SVR, (8) L1-loss SVR, and (9) one-class support vector machine. Each library algorithm is discussed in detail as follows:

1) LibSVM Library

The first library is (LibSVM), which is widely used in remote sensing studies [33][51][52]. This package implements SVM algorithms of several varieties (Linear, polynomial, radial basis function, and sigmoid). In the current paper, linearity is relied on because the equation for the dune index is linear. LibSVM has been a support vector machine library since 2000 and is one of the most extensively used SVM applications today. Between 2000 and 2010, the bundle received over 250,000 downloads. It uses several SVM formulations, and it is possible to use LibSVM for classification, regression, and distribution estimation [49]. Training and testing set of data are often separated in a classification process. Attributes and a single objective value (i.e., the class labels) are connected with each instance in this training set (i.e., the features of observed variables). The C-SVC classifier in LibSVM is used, and this classifier is solved by Cortes and Vapnik [53].

2) Generalized Linear Models (GLM)

The second library (GLM) refers to a statistical models ordered and random components [47][48][49]. For these models, log-likelihood is used to generalize the analysis of variance. Logistic regression can be used to derive maximum likelihood estimates of parameters using observations distributed according to an exponential family and linearly transformable systematic effects.

3) LibLINEAR Library

The third library used is (LibLINEAR) which contains a collection of linear support vector machines for primal and dual classification. Using multipliers makes it possible to describe primal and dual solutions differently. If an item is accurately classified but is located far from the decision plan, that object is not included in any further calculations. Burges [54] and Smola and Schölkopf [55] provide excellent tutorials on the mathematical details. Open-source LibLINEAR is a library for large-scale linear types. It includes support vector machines and logistic regression in a linear form [50][56][57].

4) Accuracy Assessment

The accuracy assessment procedure consists of two stages: the first stage is assessing the accuracy of the models by dividing the input data by three-thirds, two for training, and one for testing and calculating each method's execution time. The second stage was calculating the accuracy by distributing different scales of random points.

In more detail, the first step is to determine the accuracy of all algorithms using R software (caret package). The second step is to assess the accuracy of the results by assigning random points to six levels according to the reference image and results, beginning with 500 points to1000 points. The reference image is generated by supervised classification in ENVI 5.3 software depending on field observation and Google Earth.

The steps of classification accuracy and error assessment were done using ArcGIS as follows; (1) distribute a set of random points (500-1000) with an increasing 100 points in each set, (2) Using the reference image and the linear model image, determine the class type for each random point, and (3) computing confusion matrix. The confusion matrix structure example is shown in Table III.

TABLE III. EXAMPLE OF CONFUSION MATRIX STRUCTURE

Actual Class		Prediction Class
	+1	-1
+1	True Positive	False Negative
-1	False Positive	True Negative

Accurately classified pixels are denoted by the abbreviations ”true positives” (TP) and ”true negatives” (TN). While false positive (FP) and false negative (FN) pixels are both examples of inaccurate classification [51][58][59]. The accuracy assessment parameters are calculated using the following formula.

$$OA = \frac{(TP + TN)}{(Total)} \tag{2}$$

$$Kappa = \frac{(Po - Pe)}{(1 - Pe)} \tag{3}$$

$$Po = \frac{(TP + TN)}{(Total)} \tag{4}$$

$$Pe = \frac{P(+1)}{P(-1)} \tag{5}$$

$$P(+1) = \frac{(TP + FN)}{(Total)} * \frac{(TP + FP)}{(Total)} \tag{6}$$

$$P(-1) = \frac{(FP + TN)}{(Total)} * \frac{(FN + TN)}{(Total)} \tag{7}$$

The F-score is used which calculated depended on the producer's accuracy PA and the user's accuracy UA and all required parameters can be calculated using the following formulas [58][59].

$$OM.E(sand) = \frac{(TP)}{(Total)} * 100 \tag{8}$$

$$OM.E(non) = \frac{(FN)}{(Total)} * 100 \tag{9}$$

$$CO.E(sand) = \frac{(TN)}{(Total)} * 100 \tag{10}$$

$$CO.E(non) = \frac{(FP)}{(Total)} * 100 \tag{11}$$

$$PA = \frac{TP}{(TP + FN)} \tag{12}$$

$$UA = \frac{TP}{(TP + FP)} \tag{13}$$

$$F - Score = \frac{(2 * PA * PU)}{(PA + P)} \tag{14}$$

5. RESULTS

All linear equations generated from the models have been applied to NDs bands (ND34, ND47, and ND57). The results are illustrated by two types of analysis, qualitative and quantitative, as follows in Figure 4 and 5:

A. Qualitative Analysis

This section includes the linear equations that nine classification models calculated. Seven of these models are members of the LibLINEAR (LibL1, LibL2, LibL3, LibL4, LibL5, LibL6, and LibL7), LibSVM, and GLM librar Table IV. The linear equations were applied on NDs

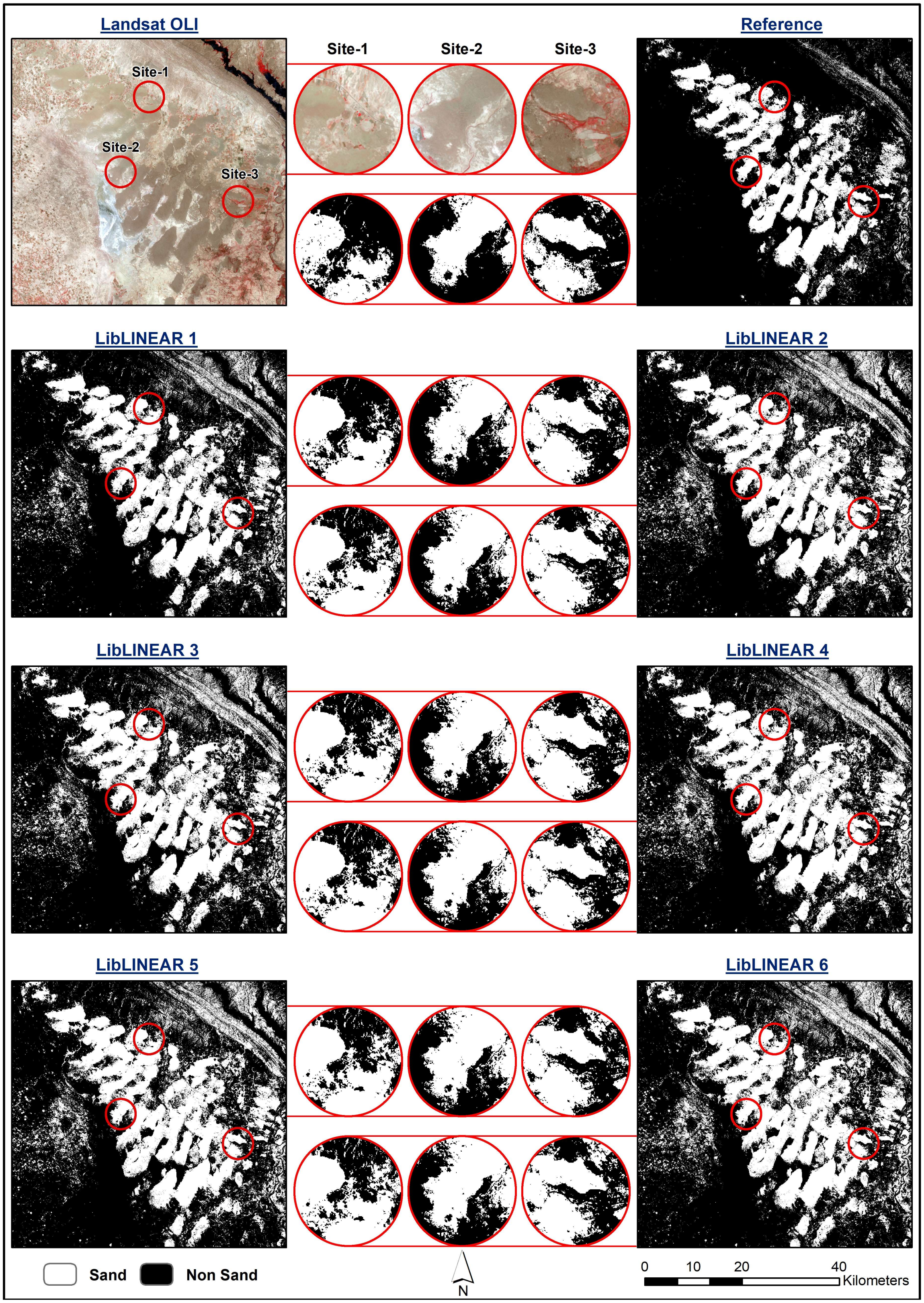

Figure 4. Result images of Linear SVM formulae

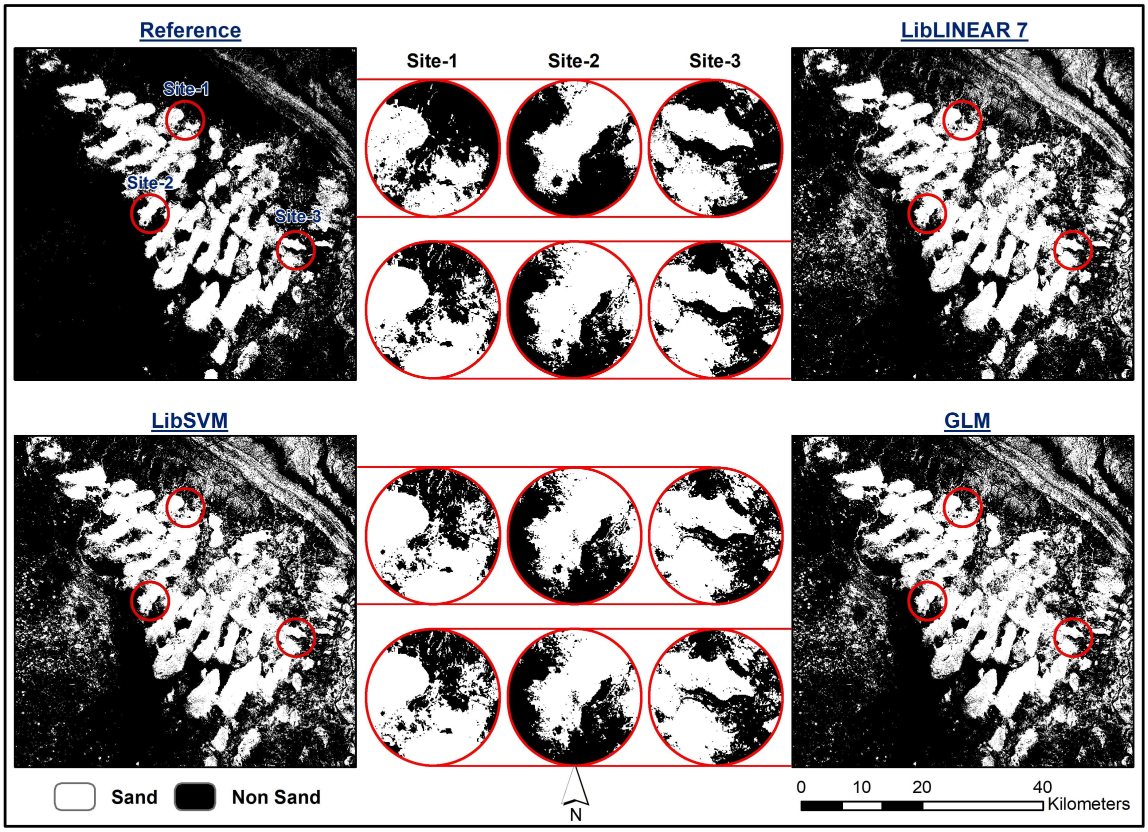

Figure 5. Result images of Linear SVM formulae

for mapping sand dunes and drifting sand Figure 4 and 5. Each classifier's formula is considered a spectral index and applied on chosen highly weighted NDs. In the first step, the suggested equation consists of 15 terms (NDs) since the index was built based on all training features. In the second step, unnecessary features with the least weight are eliminated, and a short spectral index formula that is more easily applicable is developed. The threshold value must be set in order to map sand dunes. The threshold is specified according to the bias number. Table IV shows the threshold values for each classifier.

Comparing the classification result to the reference image. The white region represents sand dunes and drifting sands, while the black area represents all other natural phenomena. Three sample locations have been selected, which are represented by red circles to enlarge and clarify the differences in Figure 4 and 5.

TABLE IV. LINEAR EQUATIONS

Index	Linear Equation	Threshold
LibL1	11*ND34-21*ND47-6*ND57	0.5
LibL2	10*ND34-21*ND47-6*ND57	0.5
LibL3	22*ND34-44*ND47-20*ND57	0.7
LibL4	12*ND34-23*ND47-10*ND57	0.3
LibL5	10*ND34-20*ND47-7*ND57	0.4
LibL6	14*ND34-60*ND47-59*ND57	1.1
LibL7	39*ND34-84*ND47-40*ND57	1.6
LibSVM	6*ND34"-59*"ND47"-10*ND57	4
GLM	0.86*ND34-135*ND47–15*ND57	11

The reference map is the standard and accurate map of sand dunes to check and test all Linear SVM results. The thresholds are optimized according to the bias, and then visually evaluated in comparison to the reference. The threshold values for each classifier model are listed in Table IV. The visual interpretation of results in Figure 4 and 5 gives an impression that the best three models for classifying sand dunes in Baiji are (LibL1, LibL7, and LibL3) respectively. However, the visual interpretation is insufficient, so the results must be compared and analyzed statistically.

B. Quantitative Analysis

Statistical results will be compared using the OA, K, and F-score criteria. In subsection *B.1*, the confusion matrix results and execution time for each model using the R software. In subsection *B.2*, the accuracy assessment of all classifiers' results. After executing all classifiers on data consisting of (83733) pixels for training and (41866) pixels for testing. Table V contains the OA, Kappa, and F-score criteria results with the execution time for each classifier.

Kappa results indicate that (LibL1, LibL7, and LibL3) have the highest accuracy because of their use of dual-mode Figure 6. The execution time for the training process and model construction is also calculated in Figure 7. It shows that the Lib2 requires the minimum processing time while LibSVM requires the maximum.

TABLE V. ACCURACY AND TRAINING TIME OF ALL CLASSIFIERS

Classifiers	OA	Kappa	F-Score	Training Time
LibL1	**99.50**	**99.00**	**99.48**	**3.962898 secs**
LibL2	98.81	97.63	98.78	0.022938 secs
LibL3	**99.17**	**98.34**	**99.14**	**0.471736 secs**
LibL4	99.16	98.33	99.13	1.018278 secs
LibL5	98.83	97.66	98.79	0.059839 secs
LibL6	99.09	98.19	99.06	0.092747 secs
LibL7	**99.46**	**98.93**	**99.45**	**1.081107 secs**
LibSVM	99.04	98.09	99.01	17.85497 secs
GLM	98.85	97.71	98.82	0.619344 secs

Figure 6. OA, Kappa, and F-score of all classifiers

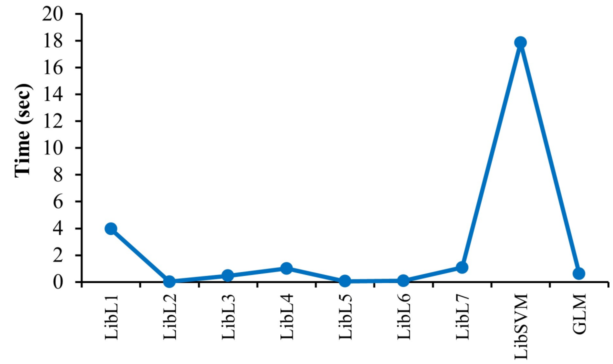

Figure 7. Execution time for all classifiers

In general, the libraries (LibLINEAR) and (GLM) need less training time than the library (LibSVM), so it is possible to benefit from these two libraries in dealing with large data that has high dimensional data. The library (LibLINEAR) was explicitly designed to work with data that comprises millions of rows. Seven different types of this library were chosen and compared in this study. The results indicated that type (LibL2) is almost the fastest training time. Calculating the execution time showed that the (LibSVM) library takes more time than other techniques; this is because this library is dedicated to classifying linear and non-linear data using multiple kernels. Besides linear kernels, LibSVM supports radial basis functions (RBFs), sigmoid, and polynomial kernels.

1) Accuracy Assessment and Execution Time of Models

This section will analyze the accuracy results for all classifiers and the execution time consumption during the training processing. The processing (training and testing) was implemented on hardware components with the following characteristics: (1) Intel(R) Core (TM) i7-9750H CPU @ 2.60GHz, (2) Memory 32.0 GB with Speed: 2667 MHz. The data was divided into two parts, where the first part includes 70% of input data for training and the second 30% for testing and calculating the accuracy. Before starting the training, some important parameters must be set, the cost parameter C = 10 and Bias = 1. The following Table VI summarizes the number of classified pixels.

TABLE VI. NUMBER OF CLASSIFIED PIXELS

Classifiers	TP	FN	FP	TN	Total
LibL1	**20099**	**204**	**3**	**21560**	**41866**
LibL2	20047	495	0	21324	41866
LibL3	**20047**	**347**	**0**	**21472**	**41866**
LibL4	20047	348	0	21471	41866
LibL5	20047	488	0	21331	41866
LibL6	20044	375	3	21444	41866
LibL7	**20218**	**221**	**1**	**21426**	**41866**
LibSVM	20047	399	0	21420	41866
GLM	20047	478	0	21341	41866

2) Accuracy Assessment of Results

The average of accuracy parameters is calculated in Table VII and Figure 8 shows the OA, Kappa, and F-score.

The accuracy calculation results indicate that (LibL1, LibL7, and LibL3) have the highest accuracy because of their use of dual-mode. Dual-mode is recommended for data that has large dimensions. In addition, the three methods that have the least amount of accuracy in comparison to the others are as follows: (LibSVM, GLM, and LibL6) Figure 8.

TABLE VII. THE AVERAGE ACCURACY OF RANDOM POINTS

Classifiers	OA	Kappa	F-Score
LibL1	**88.60**	**77.20**	**88.55**
LibL2	88.06	76.13	88.47
LibL3	**88.41**	**76.83**	**88.77**
LibL4	88.23	76.47	88.68
LibL5	88.20	76.41	88.54
LibL6	85.98	71.96	86.99
LibL7	**88.48**	**76.96**	**88.81**
LibSVM	87.46	74.92	87.85
GLM	87.35	74.70	87.59

Source: Appendix A

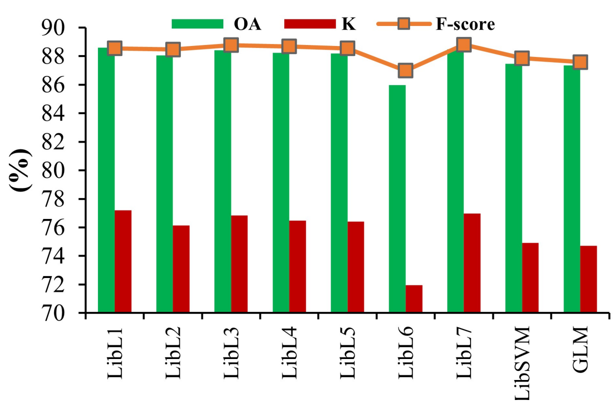

Figure 8. OA, Kappa, and F-score of all classifiers

6. DISCUSSION

In many previous studies, the library (LibSVM) has been used in various applications for classifying remote sensing data. This supervised learning approach solves binary classification problems using classification and regression. Several studies conducted using the library (LibSVM) in various remote sensing applications are discussed below.

Wang et al. [33], uuse a library (LibSVM) to classify Sentinel 2 satellite images based on an accurate water index termed "MuWI-R." The accuracy results from comparison with other water indices, such as NDWI, MNDWI, AWEI-nsh, and AWEI-sh, show that MuWI-R is more efficient in identifying and detecting water bodies. The index has been used in various locations, including Northern India, Venice, the Gulf of Mexico, and Colombia. Classification accuracy results for all indices: MuWI-R, NDWI, MNDWI, AWEI-nsh, and AWEI-shis, are 95.94%, 88.81%, 91.44%, 91.30%, and 90.94%, respectively.

Kranjčić et al. [51], used (LibSVM) to accurately extract green areas in the urban to accelerate urban planning and improve town environments. The study area is comprised of two Croatian towns: Varadin and Osijek. They use the OpenStreetMap classification system. Overall, kappa indices of 0.87 and 0.89 show high classification accuracy for green urban area extraction.

Jiang et al. [52], used LibSVM to distinguish between sea ice and seawater In Antarctica. The research area is the level 1B (L1B) data from the HY-2A/B altimeter from November 2018. The altimeter waveforms from the polar areas are analyzed. The LibSVM with a CPU separation accuracy rate is less than 40% for all bands except band Ku from HY-2B ALT.

Ajay et al. [60], studied aerial image classification using GURLS and LibSVM. This effort aims to classify a large collection of aerial images effectively. They use the Grand Unified Regularized Least Squares (GURLS) algorithm and a support vector machine library (LibSVM). LibSVM compares the various kernel techniques used in GURLS and LibSVM. The experiment uses three sets of aerial image data from the electrical engineering department at Banja Luka University's DSP laboratory, financed by the European Union's WUSAUSTRIA project. Based on the experiment results, it can be determined that the GURLS library exceeds LibSVM in prediction accuracy. The review of previous studies shows that (LibSVM) library is widely used in various remote sensing applications. The current study uses seven different LibLINEAR and GLM libraries (LibSVM).

In total, nine linear SVM classifiers have been used in the current study. The primary purpose is to compare the results in accuracy and execution time. The result shows that LibSVM is ranked seventh in average OA and Kappa, after (LibL1, LibL7, LibL3, LibL4, LibL5, and LibL2), respectively. LibLINEAR shows promising results with an OA of 88.60% and a Kappa coefficient of 77.20%. The article obtains accuracy values (Overall User, Producer, and F-score) of more than 88.5% for sand and non-sand pixels.

LibLINEAR is designed for large-scale data analysis. One of the primary issues with LibSVM is that it takes excessive time to process remote sensing data with large attributes. It is discovered by training the library (LibSVM) on the data of the Baiji area, which contains around 83733 pixels and three columns indicating NDs (ND34, ND47, and ND57), that LibSVM takes a longer time than LibLINEAR significantly and has less accuracy. In contrast, LibSVM accuracy results are OA of 87.46% and the Kappa coefficient of 74.92%. After sorting through the training time results for each method, LibSVM comes in last place, with a time of around 17.85497 seconds. This means that LibSVM is inappropriate for big data sets in remote sensing. Therefore, some studies implement LibSVM using a GPU processor, which is known for its processing speed compared to CPU processors.

The proposed work's limitation is the impossibility of dealing with a complete Landsat8 dataset which requires fast and efficient processing hardware resources. The Landsat-8 image contains approximately 62 million pixels. It is difficult to process this number of pixels. Therefore, it has been suggested that sampling the data with a specific number of pixels with sufficient representation of the studied phenomenon.

7. CONCLUSIONS

The use of machine learning in remote sensing and GIS technologies is the most effective way to construct a spectral index. The current study has applied different linear SVM methods to remote sensing Landsat-8 OLI data. The result of the accuracy assessment using different scales of random points shows that LibL1, LibL7, and LibL3 are the optimal three linear SVM methods. In contrast, the widely used method LibSVM has shown very low accuracy and is one of the three least accurate methods besides GLM and LibL6. This study concludes that simple linear equations can be utilized in a shorter time and with less effort to quickly identify and estimate the area of sand dunes.

The average K and OA for all levels of random points show that the optimal three methods are LibL1, LibL7, and LibL3; they record K values of 77.20%, and 76.96%, and 76.83%, respectively, and OA of 88.60%, 88.48%, and 88.41%, respectively. On the other hand, the widely used, LibSVM has less accuracy with more execution time than LibLINEARs. LibLINEARs can be used to solve substantial linear problems. For big datasets, the LibLINEAR implementation of SVM is preferred because of its speed and reliability. Execution time shows that LibSVM with a linear kernel requires more training and testing time than LibLINEARs and GLM. Finally, we recommend LibLINEAR because it is faster to train and has competitive accuracy.

APPENDIX A. ACCURACY ASSESSMENT DATA OF RANDOM POINTS

Method	Metrics%	Random Points 500		600		700		800		900		1000	
		Sand	Non*	Sand	Non	Sand	Non	Sand	Non	Sand	Non	Sand	Non
LibL1	Co.E*	10.44	10.75	12.95	12.70	10.94	13.29	9.69	11.27	11.45	10.76	10.95	11.85
	Om.E*	10.8	10.4	12.66	13	13.71	10.57	11.5	9.5	10.66	11.55	12	10.8
	OA	89.4		87.14		87.98		89.5		88.88		88.68	
	Kappa	78.8		74.29		75.96		79		77.77		77.37	
	F-Score	89.37		87.18		87.79		89.39		88.93		88.62	
LibL2	Co.E	15.12	8.73	16.40	10.50	14.59	10	13.03	8.73	14.72	8.02	13.04	8.49
	Om.E	8	16.4	9.66	17.66	9.42	15.42	8.25	13.75	7.33	16	8	13.8
	OA	87.8		86.47		87.69		89		88.33		89.08	
	Kappa	75.6		72.95		75.39		78		76.66		78.17	
	F-Score	88.29		86.99		88.05		89.29		88.81		89.40	
LibL3	Co.E	14.49	8.65	15.83	5	13.18	10.11	12.23	8.87	14.31	7.54	12.96	8.07
	Om.E	8	16.4	9.66	17.66	9.42	15.42	8.25	13.75	7.33	16	8	13.8
	OA	88.2		86.47		88.26		89.37		88.77		89.38	
	Kappa	76.4		72.95		76.53		78.75		77.55		78.77	
	F-Score	88.63		86.95		88.51		89.59		89.24		89.72	
LibL4	Co.E	15.01	7.92	16.92	10.94	14.28	9.45	12.82	8.70	14.63	7.35	13.627	6.79
	Om.E	7.2	16.4	10	18.33	8.85	15.14	8.25	13.5	6.66	16	6.2	14.8
	OA	88.2		85.80		88.12		89.12		88.66		89.48	
	Kappa	76.4		71.61		76.24		78.25		77.33		78.97	
	F-Score	88.71		86.39		88.48		89.40		89.17		89.93	
LibL5	Co.E	14.17	8.62	16.51	11.15	13.29	10.65	12.64	8.92	14.40	8.21	12.97	9.05
	Om.E	8	15.2	10.33	17.66	10.28	13.71	8.5	13.25	7.55	15.55	8.6	13.6
	OA	88.4		86.14		88.12		89.12		88.44		88.98	
	Kappa	76.8		72.28		76.25		78.25		76.88		77.97	
	F-Score	88.80		86.63		88.32		89.37		88.88		89.25	
LibL6	Co.E	17.54	6.97	21.09	10.27	19.10	8.10	16.62	6.87	20.56	6.84	17.85	6.16
	Om.E	6	20	8.66	24.33	6.85	22	6	18.75	5.55	24.44	5.2	20.6
	OA	87		83.63		85.55		87.62		85		87.08	
	Kappa	74		67.27		71.09		75.25		70		74.17	
	F-Score	87.85		84.82		86.58		88.36		86.29		88.02	
LibL7	Co.E	13.96	9.36	15.62	10.75	13.18	10.14	11.83	9.06	14.31	7.54	12.80	7.90
	Om.E	8.8	14.8	10	16.66	9.71	13.71	8.75	12.25	6.88	15.55	7.4	13.6
	OA	88.2		86.64		88.26		89.5		88.77		89.48	
	Kappa	76.4		73.28		76.53		79		77.55		78.97	
	F-Score	88.54		87.09		88.51		89.68		89.24		89.81	
LibSVM	Co.E	15.67	10.34	16.45	11.19	14.63	10.60	13.09	9.21	14.87	9.13	13.85	9.72
	Om.E	9.6	16.8	10.33	17.72	10	15.47	8.75	13.75	8.44	16	9.2	14.6
	OA	86.8		85.97		87.26		88.75		87.77		88.18	
	Kappa	73.6		71.94		74.53		77.5		75.55		76.37	
	F-Score	87.25		86.49		87.62		89.02		88.22		88.49	
GLM	Co.E	15	12.08	16.45	12.36	12.88	11.11	12.95	11.25	14.28	9.90	12.67	10.69
	Om.E	11.6	15.6	11.66	17.33	10.85	13.14	11	13.25	9.33	15.11	10.4	13
	OA	86.4		85.64		88.12		87.87		87.77		88.28	
	Kappa	72.8		71.28		76.25		75.75		75.55		76.57	
	F-Score	86.66		86.03		88.26		88		88.12		88.45	

Om.E = Omission error, Co.E = Commission error, and Non = No sand

References

[1] G. An, M. Xing, B. He, C. Liao, X. Huang, J. Shang, and H. Kang, "Using machine learning for estimating rice chlorophyll content from in situ hyperspectral data," *Remote Sensing*, vol. 12, no. 18, p. 3104, sep 2020. [Online]. Available: https://doi.org/10.3390%2Frs12183104

[2] B. Schauberger, J. Jägermeyr, and C. Gornott, "A systematic review of local to regional yield forecasting approaches and frequently used data resources," *European Journal of Agronomy*, vol. 120, p. 126153, oct 2020. [Online]. Available: https://doi.org/10.1016%2Fj.eja.2020.126153

[3] C. Zhang, S. Denka, H. Cooper, and D. R. Mishra, "Quantification of sawgrass marsh aboveground biomass in the coastal everglades using object-based ensemble analysis and landsat data," *Remote Sensing of Environment*, vol. 204, pp. 366–379, jan 2018. [Online]. Available: https://doi.org/10.1016%2Fj.rse.2017.10.018

[4] J. Holloway and K. Mengersen, "Statistical machine learning methods and remote sensing for sustainable development goals: A review," *Remote Sensing*, vol. 10, no. 9, p. 1365, aug 2018. [Online]. Available: https://doi.org/10.3390%2Frs10091365

[5] C. Novillo, P. Arrogante-Funes, and R. Romero-Calcerrada, "Improving land cover classifications with multiangular data: MISR data in mainland spain," *Remote Sensing*, vol. 10, no. 11, p. 1717, oct 2018. [Online]. Available: https://doi.org/10.3390%2Frs10111717

[6] M. Sheykhmousa, M. Mahdianpari, H. Ghanbari, F. Mohammadimanesh, P. Ghamisi, and S. Homayouni, "Support vector machine versus random forest for remote sensing image classification: A meta-analysis and systematic review," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 6308–6325, 2020. [Online]. Available: https://doi.org/10.1109%2Fjstars.2020.3026724

[7] P. Leinenkugel, R. Deck, J. Huth, M. Ottinger, and B. Mack, "The potential of open geodata for automated large-scale land use and land cover classification," *Remote Sensing*, vol. 11, no. 19, p. 2249, sep 2019. [Online]. Available: https://doi.org/10.3390%2Frs11192249

[8] A. Mellor and S. Boukir, "Exploring diversity in ensemble classification: Applications in large area land cover mapping," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 129, pp. 151–161, jul 2017. [Online]. Available: https://doi.org/10.1016%2Fj.isprsjprs.2017.04.017

[9] F. F. Camargo, E. E. Sano, C. M. Almeida, J. C. Mura, and T. Almeida, "A comparative assessment of machine-learning techniques for land use and land cover classification of the brazilian tropical savanna using ALOS-2/PALSAR-2 polarimetric images," *Remote Sensing*, vol. 11, no. 13, p. 1600, jul 2019. [Online]. Available: https://doi.org/10.3390%2Frs11131600

[10] X. Li, W. Chen, X. Cheng, and L. Wang, "A comparison of machine learning algorithms for mapping of complex surface-mined and agricultural landscapes using ZiYuan-3 stereo satellite imagery," *Remote Sensing*, vol. 8, no. 6, p. 514, jun 2016. [Online]. Available: https://doi.org/10.3390%2Frs8060514

[11] A. Jamali, "Evaluation and comparison of eight machine learning models in land use/land cover mapping using landsat 8 OLI: a case study of the northern region of iran," *SN Applied Sciences*, vol. 1, no. 11, oct 2019. [Online]. Available: https://doi.org/10.1007%2Fs42452-019-1527-8

[12] M. Carranza-García, J. García-Gutiérrez, and J. Riquelme, "A framework for evaluating land use and land cover classification using convolutional neural networks," *Remote Sensing*, vol. 11, no. 3, p. 274, jan 2019. [Online]. Available: https://doi.org/10.3390%2Frs11030274

[13] L. Ma, M. Li, X. Ma, L. Cheng, P. Du, and Y. Liu, "A review of supervised object-based land-cover image classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 130, pp. 277–293, aug 2017. [Online]. Available: https://doi.org/10.1016%2Fj.isprsjprs.2017.06.001

[14] G. Mountrakis, J. Im, and C. Ogole, "Support vector machines in remote sensing: A review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 3, pp. 247–259, may 2011. [Online]. Available: https://doi.org/10.1016%2Fj.isprsjprs.2010.11.001

[15] J. S. Deng, K. Wang, Y. H. Deng, and G. J. Qi, "PCA-based land-use change detection and analysis using multitemporal and multisensor satellite data," *International Journal of Remote Sensing*, vol. 29, no. 16, pp. 4823–4838, jul 2008. [Online]. Available: https://doi.org/10.1080%2F01431160801950162

[16] C. Huang, C. Zhang, Y. He, Q. Liu, H. Li, F. Su, G. Liu, and A. Bridhikitti, "Land cover mapping in cloud-prone tropical areas using sentinel-2 data: Integrating spectral features with ndvi temporal dynamics," *Remote Sensing*, vol. 12, no. 7, p. 1163, apr 2020. [Online]. Available: https://doi.org/10.3390%2Frs12071163

[17] M. Claverie, J. Ju, J. G. Masek, J. L. Dungan, E. F. Vermote, J.-C. Roger, S. V. Skakun, and C. Justice, "The harmonized landsat and sentinel-2 surface reflectance data set," *Remote Sensing of Environment*, vol. 219, pp. 145–161, dec 2018. [Online]. Available: https://doi.org/10.1016%2Fj.rse.2018.09.002

[18] M. E. D. Chaves, M. C. A. Picoli, and I. D. Sanches, "Recent applications of landsat 8/OLI and sentinel-2/MSI for land use and land cover mapping: A systematic review," *Remote Sensing*, vol. 12, no. 18, p. 3062, sep 2020. [Online]. Available: https://doi.org/10.3390%2Frs12183062

[19] N. Kussul, M. Lavreniuk, S. Skakun, and A. Shelestov, "Deep learning classification of land cover and crop types using remote sensing data," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 5, pp. 778–782, may 2017. [Online]. Available: https://doi.org/10.1109%2Flgrs.2017.2681128

[20] Q. Wang, G. A. Blackburn, A. O. Onojeghuo, J. Dash, L. Zhou, Y. Zhang, and P. M. Atkinson, "Fusion of landsat 8 OLI and sentinel-2 MSI data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 7, pp. 3885–3899, jul 2017. [Online]. Available: https://doi.org/10.1109%2Ftgrs.2017.2683444

[21] A. Bégué, D. Arvor, B. Bellon, J. Betbeder, D. de Abelleyra, R. P. D. Ferraz, V. Lebourgeois, C. Lelong, M. Simões, and S. R. Verón, "Remote sensing and cropping practices: A review," *Remote Sensing*, vol. 10, no. 2, p. 99, jan 2018. [Online]. Available: https://doi.org/10.3390%2Frs10010099

[22] S. Talukdar, P. Singha, S. Mahato, Shahfahad, S. Pal, Y.-A. Liou, and A. Rahman, "Land-use land-cover classification by machine learning classifiers for satellite observations—a review," *Remote Sensing*, vol. 12, no. 7, p. 1135, apr 2020. [Online]. Available: https://doi.org/10.3390%2Frs12071135

[23] G. M. Afrasinei, M. T. Melis, C. Buttau, C. Arras, A. Zerrim, M. Guied, M. Ouessar, B. Essifi, M. B. Zaied, A. Jlali, H. Jarray, and G. Ghiglieri, "Classification methods for detecting and

evaluating changes in desertification-related features in arid and semi-arid environments," in *Water and Land Security in Drylands*. Springer International Publishing, 2017, pp. 269–289. [Online]. Available: https://doi.org/10.1007%2F978-3-319-54021-4_23

[24] J. Pan and T. Li, "Extracting desertification from landsat TM imagery based on spectral mixture analysis and albedo-vegetation feature space," *Natural Hazards*, vol. 68, no. 2, pp. 915–927, apr 2013. [Online]. Available: https://doi.org/10.1007%2Fs11069-013-0665-3

[25] M. Lamchin, J.-Y. Lee, W.-K. Lee, E. J. Lee, M. Kim, C.-H. Lim, H.-A. Choi, and S.-R. Kim, "Assessment of land cover change and desertification using remote sensing technology in a local region of mongolia," *Advances in Space Research*, vol. 57, no. 1, pp. 64–77, jan 2016. [Online]. Available: https://doi.org/10.1016%2Fj.asr.2015.10.006

[26] A. M. F. and, "LAND DEGRADATION DETECTION USING GEO-INFORMATIONTECHNOLOGY FOR SOME SITES IN IRAQ," *Journal of Al-Nahrain University Science*, vol. 12, no. 3, pp. 94–108, sep 2009. [Online]. Available: https://doi.org/10.22401%2Fjnus.12.3.13

[27] A. A. Sahar, A. A. Alhadithi, M. A. Hassan, and A. A. Jasim, "Integrated remote sensing and GIS for developed new spectral index for estimating sandy land and its potential hazards. case study: north-east al-muthanna province area, south of iraq," *Arabian Journal of Geosciences*, vol. 14, no. 3, jan 2021. [Online]. Available: https://doi.org/10.1007%2Fs12517-021-06490-z

[28] A. A. Sahar, M. J. Rasheed, D. A. A.-H. Uaid, and A. A. Jasim, "Mapping sandy areas and their changes using remote sensing. a case study at north-east al-muthanna province, south of iraq," *Revista de Teledetección*, no. 58, p. 39, jul 2021. [Online]. Available: https://doi.org/10.4995%2Fraet.2021.13622

[29] A. Karnieli, "Development and implementation of spectral crust index over dune sands," *International Journal of Remote Sensing*, vol. 18, no. 6, pp. 1207–1220, apr 1997. [Online]. Available: https://doi.org/10.1080%2F014311697218368

[30] F. Chen, T. V. de Voorde, D. Roberts, H. Zhao, and J. Chen, "Detection of ground materials using normalized difference indices with a threshold: Risk and ways to improve," *Remote Sensing*, vol. 13, no. 3, p. 450, jan 2021. [Online]. Available: https://doi.org/10.3390%2Frs13030450

[31] F. A. Zwain, T. T. Al-Samarrai, and Y. I. Al-Saady, "A study of desertification using remote sensing techniques in basra governorate, south iraq," *Iraqi Journal of Science*, pp. 912–926, mar 2021. [Online]. Available: https://doi.org/10.24996%2Fijs.2021.62.3.22

[32] A. J. Al-Taie, Y. J. Al-Shakarchi, and A. A. Mohammed, "INVESTIGATION OF GEOTECHNICAL SPECIFICATIONS OF SAND DUNE SOIL: A CASE STUDY AROUND BAIJI IN IRAQ," *IIUM Engineering Journal*, vol. 14, no. 2, nov 2013. [Online]. Available: https://doi.org/10.31436%2Fiiumej.v14i2.408

[33] Z. Wang, J. Liu, J. Li, and D. Zhang, "Multi-spectral water index (MuWI): A native 10-m multi-spectral water index for accurate water mapping on sentinel-2," *Remote Sensing*, vol. 10, no. 10, p. 1643, oct 2018. [Online]. Available: https://doi.org/10.3390%2Frs10101643

[34] L. Li, C. Wang, J. Chen, and J. Ma, "Refinement of hyperspectral image classification with segment-tree filtering," *Remote Sensing*, vol. 9, no. 1, p. 69, jan 2017. [Online]. Available: https://doi.org/10.3390%2Frs9010069

[35] J. Yang, Z. Jiang, S. Hao, and H. Zhang, "Higher order support vector random fields for hyperspectral image classification," *ISPRS International Journal of Geo-Information*, vol. 7, no. 1, p. 19, jan 2018. [Online]. Available: https://doi.org/10.3390%2Fijgi7010019

[36] I. A. Cruz-Guerrero, R. Leon, D. U. Campos-Delgado, S. Ortega, H. Fabelo, and G. M. Callico, "Classification of hyperspectral in vivo brain tissue based on linear unmixing," *Applied Sciences*, vol. 10, no. 16, p. 5686, aug 2020. [Online]. Available: https://doi.org/10.3390%2Fapp10165686

[37] M. D. Gonzalez-Lima and C. C. Ludeña, "Using locality-sensitive hashing for SVM classification of large data sets," *Mathematics*, vol. 10, no. 11, p. 1812, may 2022. [Online]. Available: https://doi.org/10.3390%2Fmath10111812

[38] D. Horvath, J. Brown, G. Marcou, and A. Varnek, "An evolutionary optimizer of libsvm models," *Challenges*, vol. 5, no. 2, pp. 450–472, nov 2014. [Online]. Available: https://doi.org/10.3390%2Fchalle5020450

[39] H. Ni, L. Lu, M. Sun, X. Bai, and Y. Yin, "Research on fault diagnosis of PST electro-hydraulic control system of heavy tractor based on support vector machine," *Processes*, vol. 10, no. 4, p. 791, apr 2022. [Online]. Available: https://doi.org/10.3390%2Fpr10040791

[40] H. Zhang, W. Shao, S. Qiu, J. Wang, and Z. Wei, "Collaborative analysis on the marked ages of rice wines by electronic tongue and nose based on different feature data sets," *Sensors*, vol. 20, no. 4, p. 1065, feb 2020. [Online]. Available: https://doi.org/10.3390%2Fs20041065

[41] J.-S. Chou, T. T. P. Pham, and C.-C. Ho, "Metaheuristic optimized multi-level classification learning system for engineering management," *Applied Sciences*, vol. 11, no. 12, p. 5533, jun 2021. [Online]. Available: https://doi.org/10.3390%2Fapp11125533

[42] F. Al Farrajii, R. M. Al-Aazawy, A. S. Muhimeed, S. A. Salim, A. Alanbary, and A. T. Shihab, "Iraq - land degradation neutrality target setting," Report, 2017. [Online]. Available: https://www.unccd.int/documents/ldn-tsp-country-report-69

[43] J. R. Al-Obaidi, M. Y. Allawi, B. S. Al-Taie, K. H. Alobaidi, J. M. Al-Khayri, S. Abdullah, and E. I. Ahmad-Kamil, "The environmental, economic, and social development impact of desertification in iraq: a review on desertification control measures and mitigation strategies," *Environmental Monitoring and Assessment*, vol. 194, no. 6, may 2022. [Online]. Available: https://doi.org/10.1007%2Fs10661-022-10102-y

[44] Q. M. Ajaj, B. Pradhan, A. M. Noori, and M. N. Jebur, "Spatial monitoring of desertification extent in western iraq using landsat images and GIS," *Land Degradation &amp Development*, vol. 28, no. 8, pp. 2418–2431, aug 2017. [Online]. Available: https://doi.org/10.1002%2Fldr.2775

[45] I. D. S. A. Ameri, R. M. Briant, and S. Engels, "Drought severity and increased dust storm frequency in the middle east: a case study from the tigris–euphrates alluvial plain, central iraq," *Weather*, vol. 74, no. 12, pp. 416–426, mar 2019. [Online]. Available: https://doi.org/10.1002%2Fwea.3445

[46] A. Whyte, K. P. Ferentinos, and G. P. Petropoulos, "A new

synergistic approach for monitoring wetlands using sentinels -1 and 2 data with object-based machine learning algorithms," *Environmental Modelling &amp Software*, vol. 104, pp. 40–54, jun 2018. [Online]. Available: https://doi.org/10.1016%2Fj.envsoft.2018.01.023

[47] M. Pecha and D. Horák, "Analyzing l1-loss and l2-loss support vector machines implemented in PERMON toolbox," in *Lecture Notes in Electrical Engineering*. Springer International Publishing, apr 2019, pp. 13–23. [Online]. Available: https://doi.org/10.1007%2F978-3-030-14907-9_2

[48] A. Abdiansah and R. Wardoyo, "Time complexity analysis of support vector machines (SVM) in LibSVM," *International Journal of Computer Applications*, vol. 128, no. 3, pp. 28–34, oct 2015. [Online]. Available: https://doi.org/10.5120%2Fijca2015906480

[49] C.-C. Chang and C.-J. Lin, "LIBSVM," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 1–27, apr 2011. [Online]. Available: https://doi.org/10.1145%2F1961189.1961199

[50] Y. Chen, B.-L. Lu, and H. Zhao, "Parallel learning of large-scale multi-label classification problems with min-max modular LIBLINEAR," in *The 2012 International Joint Conference on Neural Networks (IJCNN)*. IEEE, jun 2012.

[51] N. Kranjčić, D. Medak, R. Župan, and M. Rezo, "Support vector machine accuracy assessment for extracting green urban areas in towns," *Remote Sensing*, vol. 11, no. 6, p. 655, mar 2019. [Online]. Available: https://doi.org/10.3390%2Frs11060655

[52] C. Jiang, M. Lin, and H. Wei, "A study of the technology used to distinguish sea ice and seawater on the haiyang-2a/b (HY-2a/b) altimeter data," *Remote Sensing*, vol. 11, no. 12, p. 1490, jun 2019. [Online]. Available: https://doi.org/10.3390%2Frs11121490

[53] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, sep 1995. [Online]. Available: https://doi.org/10.1007%2Fbf00994018

[54] C. J. Burges, *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121–167, 1998. [Online]. Available: https://doi.org/10.1023%2Fa%3A1009715923555

[55] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, no. 3, pp. 199–222, aug 2004. [Online]. Available: https://doi.org/10.1023%2Fb%3Astco.0000035301.49549.88

[56] S. S. Keerthi, S. Sundararajan, K.-W. Chang, C.-J. Hsieh, and C.-J. Lin, "A sequential dual method for large scale multi-class linear svms," in *Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD 08*. ACM Press, 2008.

[57] C.-J. Lin, R. C. Weng, and S. S. Keerthi, "Trust region newton methods for large-scale logistic regression," in *Proceedings of the 24th international conference on Machine learning - ICML '07*. ACM Press, 2007.

[58] R. Congalton and R. Mead, "A review of three discrete multivariate analysis techniques used in assessing the accuracy of remotely sensed data from error matrices," *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-24, no. 1, pp. 169–174, jan 1986. [Online]. Available: https://doi.org/10.1109%2Ftgrs.1986.289546

[59] G. M. Foody, "Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification," *Remote Sensing of Environment*, vol. 239, p. 111630, mar 2020. [Online]. Available: https://doi.org/10.1016%2Fj.rse.2019.111630

[60] A. Ajay, K. D. M. Dixon, V. Sowmya, and K. P. Soman, "Aerial image classification using GURLS and LIBSVM," in *2016 International Conference on Communication and Signal Processing (ICCSP)*. IEEE, apr 2016.

Ehsan Ali Al-Zubaidi is a Ph.D. student at the Department of Computer Science and Mathematics, University of Kufa. He graduated with a master's degree in computer science from Baghdad University in 2014. He is currently working as a faculty member at the Faculty of Physical Planning, University of Kufa. Researches interest in Data mining and machine learning.

Furkan Rabee is a faculty member in the Department of Computer Science, College of Computer Science and Mathematics, University of Kufa. He obtained a B.S. in computer engineering and an M.S. in computer engineering with a specialization in computer networks from AL-Nahrian University in 2000 and 2008. He obtained his Ph.D. in 2015 from UESTC's School of Computer Science and Engineering in Chengdu, China. He is particularly interested in real-time scheduling algorithms, real-time locking protocols, and real-time operating systems, as well as computer networks, smart cities, cloud computing, and the Internet of Things.

Ahmed H. Al-Sulttani is a Professor at the Department of Environmental Planning, Faculty of Physical Planning, University of Kufa, Iraq. He has a Ph.D. in Geomorphology and Environment hazards from Mustansiriyah University, Baghdad, Iraq, in 2006 and a Postdoctoral in Remote Sensing and Image processing from the Department of Geoinformatics, Cartography, and Remote Sensing, Faculty of Geography and Regional Planning, the University of Warsaw in 2014. His research field is in remote sensing, satellite image processing, modeling, spatial analysis, GIS, karst geomorphology, environmental hazards, and drought.