Spatiotemporal Mapping and Monitoring of Mangrove Forests
Changes From 1990 to 2019 in the Northern Emirates, UAE Using
Random Forest, Kernel Logistic Regression and Naive Bayes Tree
Models

Samy I. Elmahdy
American University of Sharjah

Tarig A. Ali
American University of Sharjah

Mohamed M. Mohamed
United Arab Emirates University

Fares M. Howari
Zayed University

Mohamed Abouleish
American University of Sharjah

Follow this and additional works at: https://zuscholars.zu.ac.ae/works

Recommended Citation
Elmahdy, Samy I.; Ali, Tarig A.; Mohamed, Mohamed M.; Howari, Fares M.; Abouleish, Mohamed; and
Simonet, Daniel, "Spatiotemporal Mapping and Monitoring of Mangrove Forests Changes From 1990 to
2019 in the Northern Emirates, UAE Using Random Forest, Kernel Logistic Regression and Naive Bayes
Tree Models" (2020). All Works. 3170.
https://zuscholars.zu.ac.ae/works/3170

This Article is brought to you for free and open access by ZU Scholars. It has been accepted for inclusion in All
Works by an authorized administrator of ZU Scholars. For more information, please contact
Yrjo.Lappalainen@zu.ac.ae, nikesh.narayanan@zu.ac.ae.
Spatiotemporal Mapping and Monitoring of Mangrove Forests Changes From 1990 to 2019 in the Northern Emirates, UAE Using Random Forest, Kernel Logistic Regression and Naive Bayes Tree Models

Samy I. Elmahdy*, Tarig A. Ali1,2, Mohamed M. Mohamed3,4, Fares M. Howari5, Mohamed Abouleish6 and Daniel Simonet7

1 GIS and Mapping Lab, College of Engineering, American University of Sharjah, Sharjah, United Arab Emirates, 2 College of Engineering, American University of Sharjah, Sharjah, United Arab Emirates, 3 Department of Civil Engineering, College of Engineering, United Arab Emirates University, Al Ain, United Arab Emirates, 4 National Water Center, United Arab Emirates University, Al Ain, United Arab Emirates, 5 College of Natural and Health Sciences, Zayed University, Abu Dhabi, United Arab Emirates, 6 Department of Biology, Chemistry and Environmental Sciences, College of Arts and Sciences, American University of Sharjah, Sharjah, United Arab Emirates, 7 Department of Management, School of Business Administration, American University of Sharjah, Sharjah, United Arab Emirates

Mangrove forests are acting as a green lung for the coastal cities of the United Arab Emirates, providing a habitat for wildlife, storing blue carbon in sediment and protecting shoreline. Thus, the first step toward conservation and a better understanding of the ecological setting of mangroves is mapping and monitoring mangrove extent over multiple spatial scales. This study aims to develop a novel low-cost remote sensing approach for spatiotemporal mapping and monitoring mangrove forest extent in the northern part of the United Arab Emirates. The approach was developed based on random forest (RF), Kernel logistic regression (KLR), and Naive Bayes Tree machine learning algorithms which use multitemporal Landsat images. Our results of accuracy metrics include accuracy, precision, and recall, F1 score revealed that RF outperformed the KLR and NB with an F1 score of more than 0.90. Each pair of produced mangrove maps (1990–2000, 2000–2010, 2010–2019, and 1990–2019) was used to image difference algorithm to monitor mangrove extent by applying a threshold ranges from +1 to −1. Our results are of great importance to the ecological and research community. The new maps presented in this study will be a good reference and a useful source for the coastal management organization.

Keywords: NUAE, mangrove, FMNF, remote sensing, change detection, Landsat
INTRODUCTION

Mangroves are woody plants that are extensively distributed in intertidal and estuarine zones and their forests cover thousands of hectares along the shorelines (Sherrod and McMillan, 1985; Field et al., 1998). Mangroves are able to store 1200 Mg of carbon per hectare (Donato et al., 2011; Pham et al., 2019a,b). They form an essential component of the coastal ecosystem consisting of salt-tolerant plants with aerial breathing roots that supply a microenvironment to several marine species (Snedaker, 1982; Upadhyay and Mishra, 2008).

Mangroves provide a wide range of benefits to the economy and the environment as they play a vital role in ecology. Some ways they do this, for example, are by supplying a safe breeding ground and suitable environment for fish species and birds nesting, as well as and stabilization and protection for shorelines from hurricanes (Fry and Cormier, 2011; Giri et al., 2011; Inoue et al., 2011; Barua et al., 2014).

Although mangroves have declined globally by 2% per annum, the mangroves of the United Arab Emirates appear to have rapidly increased. This is due to localized plantation activities and increased public awareness and conservation efforts during the last decade (Food and Agriculture Organization of the United Nations, 2007; Loughland et al., 2007; Howari et al., 2009). The most common mangrove species in the United Arab Emirates is Avicennia marina or gray mangrove and its protection is consequently an urgent conservation priority (Figure 1).

Our knowledge on the spatial variation, causes of degradation of the mangroves are still poor due to; (i) the northern part of the United Arab Emirates (NUAE) mangrove database suffering from significant errors and it is sometimes unavailable and has usage restricted, (ii) inaccessibility of the tidal fluctuation mangrove forests in isolated islands, and (iii) the majority of previous studies have been conducted based on field surveys and measurements over a local scale with manual screen digitizing (Moore et al., 2013; Alsumaiti, 2014). Understanding the spatiotemporal changes of the NUAE mangrove over a regional scale could provide valuable information for the ecologists and natural resources specialists in the United Arab Emirates and this requires a potential method.

Remote sensing data has proven to be a good tool for mapping and monitoring mangrove changes over a regional scale as it has low-cost and is not time-consuming (Boardman and Kruse, 1994; Guild et al., 2004; Ghanavati et al., 2008; Mondal et al., 2018, 2019; Elmahdy et al., 2019, 2020a,b; Spruce et al., 2020). The multi-temporal Landsat images have been widely used for mapping and monitoring mangrove changes due to their suitable of spatial and temporal resolutions, as well as their easy availability and accessibility (Birth and McVey, 1968; Edwards et al., 2007; Klemas, 2009; Vo et al., 2013; Nguyen et al., 2013; Kanniah et al., 2015; Chen et al., 2017; Elmahdy and Mohamed, 2018; Buitre et al., 2019).

The most common method for mapping vegetation are the vegetation indices such as the Simple Ratio Index (SRI) of Birth and McVey (1968), Normalized Difference Vegetation Index (NDVI) of Edwards and Richardson (2004), and the Normalized Difference Moisture Index (NDMI) is the oldest and most well known and most frequently used by several researchers (Fang and Liang, 2003; Huete et al., 2010). These indices were designed to enhance the sensitivity of the spectral reflectance contribution of vegetation while minimizing the soil background reflectance or atmospheric effects (Fang and Liang, 2008; Huete et al., 2010) and widely used in the literature (Díaz and Blackburn, 2003; Ishil and Tateda, 2004; Jean-Baptiste and Jensen, 2006; Kovacs et al., 2009; Rodriguez-Romero et al., 2011; Laongmanee et al., 2013; Nascimento et al., 2013; Pereira et al., 2018; Otero et al., 2019; Liu et al., 2020). These indices sometimes cannot discriminate between mangrove and non-mangroves areas such as grass and algae (Howari et al., 2009; Elmahdy and Mostafa Mohamed, 2013a,b). Generally, the classification methods can be divided into five types: (i) unsupervised learning; (ii) supervised; (iii) advanced learning; (iv) object-based image analysis (OBIA); and (v) sub-pixel. Recently, several machine learning algorithms have been designed for mapping and classifying land use land cover (LULC). Ensemble machine learning algorithms such as Random forest (RF) is widely used in LULC classification and mangrove mapping from Landsat images (Erfemeijer and Hamerlynck, 2005; Pal, 2005; Sesnie et al., 2008; Mountrakis et al., 2011), mangrove and sea grass mapping (Heumann, 2011; Hossain et al., 2015; Buitre et al., 2019; Diniz et al., 2019; Small and Sousa, 2019; Toosi et al., 2019), prediction in water resources (Zhao et al., 2012; McGinnis and Kerans, 2013; Naghibi et al., 2016; Naghibi and Dashtpaderi, 2017), and prediction of land subsidence (Elmahdy et al., 2020a). Further studies combined image transformation and supervised classification to map and classify mangrove forests (Yokoya and Iwasaki, 2010; Ouerghemmi et al., 2018). Locally, only limited numbers of studies have been carried out using remote sensing data (Embabi, 1993; Moore et al., 2013; Almahasheer, 2018). However, these studies were based only on traditional techniques of traditional classifiers and visual interpretation of Landsat and spot images, sometimes coupled with manual digitization techniques for mapping mangrove forests. These techniques introduced a level of human bias and can be subjective and time-consuming owing to the rapid changes in the United Arab Emirates mangroves (Crouvi et al., 2006; Adam and Hutchings, 2010). Therefore, to construct and maintain a flexible regional database for the NUAE mangroves, it is urgent to modify an integrated powerful machine learning classifier.

Integration of RF, Kernel logistic regression (KLR), Naive Bayes Tree (NBT), and Image difference (ID) have shown the ability to achieve classification over a regional scale and precise monitoring extent of the mangroves (Colkesen and Kavzoglu, 2017). These techniques can reduce the variance and overfitting of the classification maps and assess many variables separately compared to traditional classifiers, such as maximum likelihood (Ha et al., 2020). Thus, integration of RF and KLR and ID was adopted to extract spatiotemporal information about the NUAE mangrove forests. The use of machine learning algorithms decreases the overfitting and variance in the classified maps (Belgiu and Dăguță, 2016; Feng et al., 2018; Mondal et al., 2019; Elmahdy et al., 2020b; Ha et al., 2020). Thus, the main goals of this study were to present a novel ensemble machine learning approach which integrates RF with KLR and NBID
Elmahdy et al. Enhancing RF Performance

FIGURE 1 | Photos of the NUAE mangrove forests showing their spatial patterns and environment.

algorithms and uses Landsat images for spatiotemporal mapping of the NAEU mangroves, comparing the performance of these algorithms, and implementing a novel image to image change detection technique for monitoring mangrove changes over multiple scales.

STUDY AREA

The study area is in the eastern and western coastal areas of the United Arab Emirates. In particular, it extends between the longitudes E 53°56' 23'' and E 60°15' 22'' and latitudes 24° 21' 2'' and 26° 12' 20'' N (Figures 2, 3). The area consists of two mangrove ecosystems; the Khor Fakkan mangrove is located in the Gulf of Oman (5.5 km length) and the Dubai’s Greek, Ajman’s Creek, Hammraih’s Creek, Umm Al Quwain and Ras Al Khaimah (RAK) Estuaries (Figures 2, 3). They are facing the Gulf of Oman in the east and the Arabian Gulf in the west and dominated by gray mangroves (A. marina) and locally referred to as Al Qurm, which is one of the most habitat-tolerant mangroves (Boer and Aspinall, 2005; Howari et al., 2009; Alsaideh et al., 2013). The United Arab Emirates mangroves are able to grow in saline soil that exceeds the seawater salinity of the Arabian Gulf and Gulf of Oman (Alsumaiti, 2014). An Assessment of A. marina Forest Structure and Aboveground Biomass in Eastern Mangrove Lagoon National Park, Abu Dhabi
Elmahdy et al. Enhancing RF Performance

FIGURE 2 | RGB 541 band combination of Landsat image of the UAE showing the spatial distribution of the UAE mangroves and in the study area (red polygons).

The height of the United Arab Emirates mangrove trees ranges from a few centimeters to 3–8 m (Moore et al., 2013; Alsumaiti, 2014). The mangrove soil texture is loamy clay and silt and characterized by high porosity and low permeability, which allows keeping the seawater in the soil porosity for a long time, especially during tidal times (Cintron et al., 1978; Bashitialshaaer et al., 2011).

DATASETS AND METHODS

Datasets and Preprocessing

Two remotely sensed data were used in this study. The first dataset was the Landsat Thematic Mapper (TM) acquired on August 23, 1990, the Landsat Enhanced Thematic Mapper (ETM+) acquired on August 23, 2000 and August 19, 2010 and the Operational Landsat Imager (OLI) Landsat 8 acquired on August 15, 2019 (Path 160, rows 42 and 43). The data were obtained via the USGS Earth Resources Observations and Science (EROS) Center through the Global Visualization Viewer1. The mangrove forests in Landsat images (coastal areas) were then clipped into five subsets to reduce the image processing time by avoiding unnecessary calculations (Amarsaikhan et al., 2009; Diniz et al., 2019; Ma et al., 2019).

We used the Landsat images due to their suitability spectral and spatiotemporal resolutions, free of charge, easy accessibility, and time-series availability (Chander, 2009; Irons et al., 2012; Elmahdy and Mohamed, 2018; Milani, 2018; Toosi et al., 2019). Additionally, Landsat images are provided in orthorectification format at a lower level of cloud cover (Darvishsefat, 1995).

The second dataset includes the QuickBird images with a spatial resolution of 0.6 m acquired on August 22, 2019, the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), and Earth Observing-1 (EO-1) images with a spatial resolution of 30 m acquired on August 13, 2019 and July 29, 2019, respectively.

The spectral coverage of ASTER sensor consists of four visible and near-infrared bands (0.52–0.86 µm) with a spatial resolution of 15 m, six shortwave infrared (SWIR) bands (1.6–2.43 µm) with a spatial resolution of 30 m and five thermal bands (8.125–1165 µm) with a spatial resolution of 90 m. The EO-1 Advanced Land Imager (ALI) provides Landsat type panchromatic and multispectral bands. These bands have been designed to mimic six Landsat bands with three additional bands covering 0.433–0.453, 0.845–0.890, and 1.20–1.30 µm. The ALI also contains wide-angle optics designed to provide a continuous 15° × 1.625° field of view for a fully populated focal plane with a 30-m resolution.

1www.glovis.usgs.gov
for the multispectral pixels and 10-m resolution for the panchromatic pixels.

These datasets were downloaded via the USGS EROS Center through the Global Visualization Viewer (see footnote 1). We used these various datasets to collect training datasets as input to the classifiers and visual verification of the obtained mangrove maps and compare the textural features (mangroves and non-mangrove areas), mapped from Landsat images against those mapped from the ASTER and EO-1 images using RF, KLR, and NB algorithms.

After collecting the remote sensing data, the datasets were all re-sampled to UTM WGS 84 projection, at a spatial resolution
of 30 m with mean root square error values of less than 0.40 pixels. All the United Arab Emirates images were assigned to a UTM zone 40 N and datum WGS 48 (Jensen, 1996; Foody, 2002) followed by an atmospheric correction. The atmospheric correction was performed by Fast Line-of-sight Atmospheric Analysis of Hypercubes (FLAASH) implemented in Envi v. 4.6 software. Fast Line-of-sight Atmospheric Analysis of Hypercubes process consists of radiometric calibration and dark subtraction. In radiometric calibration, beta nought calibration, all digital number (DN) values were converted into Top Of Atmosphere (TOA), reflectance. Top Of Atmosphere was performed using four parameters, namely, calibration type (reflectance), output interleave (BSQ), output data type (float), and scale factor value of 1. In dark objects subtraction, TOA was converted into surface reflectance (SR) using band minimum (Green et al., 1988; Foody, 2002; Todd and Chris, 2010).

Training Data Collections
The training data collection is a very important step in producing a higher quality of classification, especially when the classifiers trained with training datasets with higher spatial resolutions (Elmahdy and Mohamed, 2018). The training datasets were collected from QuickBird images with a spatial resolution of 0.6 m using a straight random sampling or proportional method, which reduces error and bias. This method divides the population into homogenous groups and produces training sample sizes that are directly related to the size of the classes. We used this method due to its ability to reduce bias and errors (Van Niel et al., 2005; Elmahdy and Mohamed, 2018).

Random sampling collection was performed using Envi v. 4.5 software. In total, 536 training datasets were collected and divided into 375 (70%) for training and mangrove mapping and 161 (30%) for validation of this study. The collected training datasets were checked using visual interpretation and knowledge and background of the authors and where they live (Figure 4).

Machine Learning for Image Classification
Random Forest
To classify and map mangrove forests in an accurate and low-cost way, it is important to employ machine learning algorithms, then learn these algorithms with training datasets with a higher spatial resolution as well as algorithm optimal parameterization (Huang et al., 2009; Elmahdy and Mohamed, 2018).

The RF algorithm can provide a higher quality of classification than linear classifiers and has been employed previously to map and classify mangroves in Iran based on Landsat images (Kamal and Phinn, 2011; Kamal et al., 2015). The algorithm performs well for mangroves mapping over a regional scale and its ability to handle data with unclassified pixels (Prasad et al., 2006; Taalab et al., 2018; Thomas et al., 2018). Random forest algorithm is a non-parametric ensemble machine learning and one of the most popular algorithms that provide high-quality of mangroves classification and environmental modeling (Strobl et al., 2008; Vyas et al., 2011; Bachmair and Weiler, 2012; Torres and Qiu, 2014; Fu et al., 2017). It is a combination of classification and regression tree (CART).

A regression tree is a hierarchy and comprises of a root node, node separator (decision rules) and end of a leaf node (desired classes). The algorithm involves an ensemble of tree- structure classifiers and each tree starts by choosing a set of random model units selected separately. Then, it creates a forest using the vote on all decision “tree” in the “forest.” Then, two-thirds of the set was employed to make the decision tree and the remaining data are then used to estimate the model performance based on a bootstrapping procedure (Breiman, 2001; Schapire, 2003). Finally, the best model was chosen based on the majority vote approach (Yu et al., 2011). The model is flexible and can be used to solve problems related to classification and exploit the information provided by several condition variables (Catani et al., 2013; Ha et al., 2020). The model uses a random selection of predictor variables to divide each node of the trees. After that, each tree was developed in such a way as to reduce errors in the classification process. However, this type of random selection influences the results, hence creating a very unsuitable single-tree classification (Figure 5).

The algorithm is widely employed in environmental modeling (Strobl et al., 2008; Bachmair and Weiler, 2012), although only rarely in land subsidence and sinkholes susceptibility mapping (Vorpahl et al., 2012).

For b = 1 to B:

(b) Grow a random-forest tree Tb from the bootstrapped data, by recursively repeating the following steps for each terminal node of the tree, until the minimum node size nmin is reached.

1. Select m variables at random from the p variables.
FIGURE 6 | A comparison of the precision (A), recall (B), and F1 scores (C) for the Landsat image (August 15, 2019) using RF, KLR, and NB classifiers and total area in ha mapped using RF, KLR, and NB (D) and feature extracted from the Landsat images (E), ASTER images (F), and EO-1 (G) for the year 2010.
2. Pick the best variable/split-point among the \( m \) variables.
3. Split the node into two daughter nodes.

**Kernel Logistic Regression**

Kernel logistic regression is a discriminative machine learning classifier that differentiates the mangrove and non-mangrove classes perfectly where probabilistic output are evaluated based on minimizing the negative log-likelihood function using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) optimization (Yokoya and Iwasaki, 2010). It is described as the kernel version of logistic regression capable of converting into a high-dimensionality feature space of the original input feature space kernel functions (Tanaka et al., 2011).

The kernel function can be defined as the basic function in which \( \phi \) is supposed to be unidentified:

\[
K (x, x') = \phi(x)^T \phi(x')
\]  

(1)

where \( T \) is the inner product in the \( Z \) space.

Suppose a set of training dataset \( \{x_i, y_i\}_{i=1}^N \) with \( x_i \in \mathbb{R}^n \) as input parameters with \( n \) variables and \( N \) data samples. Here, the input parameters are dense mangrove, dispersed mangrove and non-mangrove.

\( y_i \in \{1,0\} \) is the related label that indicates mangrove and non-mangrove classes. The KLR was applied to built a non-linear decision boundary that may divide the two classes in the feature space based on the following equation:

\[
p(x) = \frac{e^{\beta(x)}}{1 + e^{\beta(x)}} = \sum_{i=1}^N \alpha_i K(x_i, x_j) + b
\]  

(2)

where \( y(x) \) is the logistic function with values in \([0,1]\); \( \alpha_i \) is a vector of dual model parameters, whereas \( b \) is the intercept; \( K(x_i, x_j) \) is the kernel function.

For our study, radial basis function (RBF), which is used widely in the literature, was chosen due its flexibility and reliability (Bui, 2015; Hong et al., 2015).

**Naive Bayes Tree**

Naive Bayes (NB) is a machine learning classifier that creates a probability-based model. It works based on Bayes Theorem, which is known as Naive Bayes. The NB uses a decision tree (DT) for its structure and organizes the NB model on every leaf node of the constructed DT. The NBT exhibits a significant classification performance and accuracy (Liang et al., 2006; Wang et al., 2015).

During the NB process, the impact of an attribute value on a specific class is an independent value of another attribute and known as class conditional independence. This conditional independence of NB makes the datasets to train quicker and it considers all the vectors as independent and applies the Bayes rule (Farid et al., 2014).

The theorem of the Bayes can be explained as follows:

\[
P(A|B) = P(B|A) P(A) / P(B)
\]

where:

- \( P(A|B) \) = conditional probability of A given B;
- \( P(B|A) \) = conditional probability of A given B;
- \( P(A) \) = probability of event A;
- \( P(B) \) = probability of event B.
This classifier was chosen due to its: (i) quick training and classification, (ii) powerful to irrelevant features, (iii) easy to use and understand, and (v) ability to work with few numbers of training datasets (Ho, 1998).

The model starts by estimating the probability of each class in the model, calculating covariance and variance matrix, and building the discrimination function for each class (Pham and Yoshino, 2015; Wang et al., 2015).

For the KLR and NB, mapping and monitoring of mangroves were performed using SATISTICA v. 7 (Fleiss et al., 2003; Hill and Lewicki, 2006) and Salford system (Friedman, 2001; Friedman, 2002). These tools have a stochastic gradient boosting tree which is used widely for regression problems related to predict and map a continues dependent variables (Hill and Lewicki, 2006). After that, the setting and optimizing of all parameters were performed. These parameters namely; learning rate, the number of additive

FIGURE 8 | Spatiotemporal variation of mangrove forests on the coastal area of RAK for the years 1990 (A), 2000 (B), 2010 (C), and 2019 (D) and zooms of Mina Al Arab (southwest), Dafan creek (middle), and Sham (northeast). Red polygons and dark green highlight dense mangroves.
FIGURE 9 | Spatiotemporal variations of mangrove forests for the years in Ajman-Hammriah (A–D), and Ras Al Khor (E–H) natural reserves for the years 1990, 2000, 2010, and 2019. Red polygons and dark green highlight dense mangroves.

trees, the proportion of subsampling, and so forth. Here, the optimal value for the learning rate was set as 0.1, additive trees were 185 and the maximum size of the tree was 5. These values may lead to precise result accuracy (Friedman, 2001).

To get the optimum RF parameter values, we tested several RF parameter values. These parameters include the maximum total of tree depth, the minimum numbers of samples per node (sample), and maximum tree number (tree number). The tested values of tree depth and minimum sample per node were 0, 1, 5, 10, 20, 50, 100, 500, and 1000, while the maximum tree number values were 50, 100, 200, 400, 800, and 1000. After that, the best band combination of RF parameters and the highest overall accuracy was chosen and applied. RF was applied to each Landsat image separately using 536 training data sets collected from QuickBird images with a spatial resolution of 0.6 m. The classification process was performed using the R package (Catani et al., 2013).

Ensemble of RF, KLR and NB
Among several ensemble methods, the stacking algorithm was employed to build the ensemble model with a threshold of 0.9 used to detect mangrove pixels. The algorithm builds a model on of the RF, KLR, and NB machine learning algorithms based on the selected training datasets R package via the “glmnet”.

Evaluation of the Performance of the Classifiers
To evaluate the performance of the classifiers, maps of mangroves were standardized and compared based on a pixel by pixel producing numerical values for mangrove commission, mangrove omission, total incorrect pixels, percentage of incorrect pixels, precession, recall, and F1 score (Congalton et al., 1983; Raschka, 2018; Raschka and Mirjalili, 2019). Once the classification process was achieved, it is important to evaluate classifiers’ performance. Confusion metrics include accuracy, precision, recall, F1 score was found to the best technique and used widely in this literature (Ha et al., 2020). The calculation of accuracy, precision, recall, and F1 score is based on four parameters namely; true positive (TP), true- negative (TN), false-positive (FP), and false-negative (FN). Accuracy, precision, recall and F1 score can be calculated via the following equations:

\[
\text{Accuracy} = \frac{TP}{TP + FP + FN + TN}
\] (3)

or

\[
\text{Accuracy} = \frac{1}{n_{\text{samples}}} \sum_{i=0}^{n_{\text{samples}-1}} 1 (y_{\text{predi}} = y_i)
\]

where \( y_{\text{predi}} \) is the predicted value and \( y \) is the corresponding true value

\[
\text{Kappa} = p_o - p_e / 1 - p_e
\] (4)

where \( p_o \) is the observed agreement ratio and \( p_e \) is the expected agreement

\[
\text{Precision} = \frac{TP}{TP + FP}
\] (5)

\[
\text{Recall} = \frac{TP}{TP + FN}
\] (6)

\[
F1 = 2 \times \text{precision} \times \text{recall}/\text{precision} + \text{recall}
\] (7)

where TP is the true positive; FP is the false positive; and FN is the false negative.
The performance of machine learning algorithms were evaluated using the open source R 4.0.0 software.

Further evaluation of the performance of the RF, KLR, and NS algorithms were performed by applying them on the ASTER and EO-1 with a spatial resolution of 30 m and calculate the total area of mangrove (Mondal et al., 2019) and the textural features (mangroves and non-mangrove areas) evident from Landsat images were compared against those from the ASTER and EO-1 images with a spatial resolution of 30m using RF, KLR, and NB algorithms. This method was applied to the Kalba mangrove natural reserve on the eastern side of the study area. These two methods can be considered as an alternative way of field observation, especially when the mangrove forests are located in remote and inaccessible areas (Elmahdy and Mostafa Mohamed, 2013a,b; Elmahdy and Mohamed, 2018; Estoque et al., 2018).

**Change Detection**

Among several change detection techniques, ID algorithm was chosen to monitor mangrove changes. The ID algorithm was chosen due to its ability to locate the changes in mangroves within each class (dense and sparse) (Mishra et al., 2017; Elmahdy and Mohamed, 2018). The ID algorithm determines the difference in the DN values in each pair of mangrove images (1990–2000, 2000–2010, 2010–2019, and 1990–2019). This can serve as an indicator of mangrove change which has occurred over the NUAE. Monitoring changes start with proper input parameters such as the number of classes, threshold value and change detection type. Threshold value locates the changes in mangroves within each class, depending on the type of remote sensing data (Singh, 1986).
The threshold values are evenly spaced between \((-1)\) and \((+1)\) for simple difference (the initial state image is subtracted from the final state image). The positive changes (positive value) represent the first \((n/2)\) classes, while the negative changes represent the last \((n/2)\) classes. The no-change class \([\lfloor n/2 \rfloor + 1]\) represents the middle class (Figure 4), and normalizing the images (raster maps) by subtracting the image minimum. On the final change detection maps, a positive change was identified when pixels took on a red color code (final state image), while a negative change was identified when pixels took on a blue color code (initial state image). Monitoring the NUAE mangrove changes were performed using a change detection tool implemented in the Envi v.4.5 software.

**RESULTS**

**Optimal Parameterization and Evaluation of Machine Learning Performance**

Figures 6A–C show different precision, recall, and F1 values for RF, KLR, and NB. Random forest and KLR yielded the highest values (0.95) for F1 score, while NB presented the lowest value (0.75) for F1. Slight differences in F1 score between RF and KLR algorithms were observed. This difference appears to be due to the difference in precision and recall scores and this difference to be due to the difference in overall incorrect pixels between RF, KLR, and NB. Both RF and KLR showed a strong ability in discriminating between dense mangrove compared with NB. However, all models showed less ability.
to distinguish dispersed mangroves as indicated from F1 score (<0.8), and thus, lower ability to detect mangrove forests. This lower accuracy due to lower ability in discriminating between mangrove and non-mangrove areas and the lower ability in discriminating owing to the lower reflectance. Results of cross-validation showed that the RF and KLR classifiers predicted and classified mangrove from the EO1 ALI better than NB classifier. The extent predicted mangrove areas by the RF and KLR have a wider range compared to those predicted by NB (Figure 6D).

Both RF and KLR predicted an area of 72.86 and 71.7 ha, respectively. However, the mangrove extent predicted by the NB has a lower range of mangroves from the Landsat, ASTER and EO-1 images. Although these images have the same spatial resolution, the ASTER and EO-1 miss the regular time span and geographical coverage of the NUAE mangrove. Further validation of the proposed approach was performed by comparing textural features from Landsat images against those extracted from the ASTER and EO-1 images using RF, KLR, and NB algorithms (Figures 6E–G).

### Mangrove Mapping and Classification

Maps of mangrove forests produced from the Landsat images using RF classifier with a population minimum of 100 and 8 neighbors are shown in Figures 7–10. The figures comprise of two color codes to facilitate visual interpretation. The first color code is dark green corresponding to dense mangroves (healthy mangroves). These dense mangroves are located on the gentle slopes of tidal flats, easily discriminated and concentrated in the middle parts of the mangrove patches. The second color code is light green, corresponding to disperse mangroves and are difficult to discriminate and map and are clustered around the dense mangroves and submerged in the water. The figures show that the mangrove forests are mainly spatially distributed in creeks (Khors in local language) of Kalba (Figures 7, 11B), Ras Al Khaimah (Figures 8, 11C), Ajman-Hammria and Dubai (Figure 9), and Umm Al Quwain (Figure 10), which represent about 35% of the total area of the United Arab Emirates mangroves (Elmahdy and Mostafa Mohamed, 2013a,b).

In Khor Kalba (the Emirate of Sharjah), mangrove areas are spatially distributed as linear and curvilinear strips with a length of 5 km and a width of 300 m. Their common trends were found to be in the NNW-SSE and NNW-SSE directions (Figures 7, 11A). From 1990 to 2019, sparse mangroves have slightly increased from 28.89 ha (1.53%) in 1990 to 34.56 ha (2.037%) in 2019, while dense mangroves increased from 49.95 ha (2.7%) in 1990 to 61.74 ha (4.11%) in 2010 and 63.9 ha (4.80%) in 2019. Most of the mangrove areas are mainly distributed in creeks and estuaries (that now are natural reserves) along the coastal area of the Arabian Gulf.

The highest spatial distribution of mangroves was observed to be in the estuarine area of Umm Al Quwain, covering an area ranging from 700 to 1200 ha, while the lowest spatial distribution of mangrove was observed to be in Ras Al Khor of Dubai covering an area ranging from 4 to 9 ha, where the saline tidal flats distributed in an area is currently dominated by intensive
urbanization. Moving to the west, in Ras Al Khaima, mangroves are mainly distributed in Mina Al Arab (south), Dafan Creek (middle) and Shaam estuarine areas (north) (Figures 8, 11C).

In Ajman and Hammria Creeks, dispersed mangroves occupied an area of 2.1 ha (0.0021%) in 1990, 28.42 ha (0.39%) in 2000 and 36 ha (0.41%) in 2019. Between 1990 and 2000, the mangroves were dispersed and occupied an area of 28.42 ha (0.39%) (Figures 9A–D, 11D). In 2010, both types of mangroves were clearly observed. Dense and dispersed mangroves occupied an area of 29.88 (0.41%) and 60.39 ha (0.83%), respectively. Like the Ajman-Hammria area, small patches of dispersed and dense mangroves were observed in the 1990 map. Moving to Dubai, mangrove forests were observed to be the smallest mangrove area (Figures 9E–H).

The total area of dispersed mangroves in these three locations was 468 ha (1.38%) in 1990, 149.22 ha (0.305%) in 2000 and 163.53 ha (0.33%) in 2019 (Figure 11D). Similarly, dense mangrove occupied an area of 217.71 ha (0.44%) in 1990 to 187.65 ha (0.832%). These small patches of mangroves were observed in the western edges of Ras Al Khor and occupied an area of about 18.99 ha in 2010 and 40.95 ha in 2019, respectively. Dense mangroves were observed to be clustered in the southern edge of the Ras Al Khor and easily definable in remote sensing data than the dispersed mangrove distributed in the western edge of the Ras Al Khor. Since 1990, dense and dispersed mangroves have gradually increased from 4.3 ha in 1990 to 94.23 ha in 2019 (Figure 11E) and spatially distributed in the Ras Al Khor area. The dispersed mangroves were observed to be distributed in the western part of the Ras Al Khor, while dense mangroves were observed to be distributed in the southwestern edge of Ras Al Khor (Figures 9, 11E). They occupied an area of about 3.22 ha (0.51%) and 0.97 ha (0.071%), respectively.

In Umm Al Quwain estuarine, dispersed mangrove occupied an area of 590.22 ha (2.29%) in 1990, 453.24 ha (1.764%) in 2000 and 715.32 (2.784%) in 2019 (Figures 10, 11F). Similarly, the dense mangrove occupied an area of 194.13 ha (0.75%) increased to 366.57 ha (0.42%) in 2000 and 518.58 ha (2%), then slightly
declined to 449.55 (1.74%) in 2019. In 2019, dispersed and dense mangroves increased, occupying an area of 36 ha (0.46%) and 76.14 ha (0.91%), respectively.

Change Detection of the NUAE Mangrove Forests

The results of monitoring mangrove changes in four periods: 1990–2000, 2000–2010, 2010–2019, and 1990–2019 with a time-span of 10 years are shown in Figures 12–16 and their statistics are shown in Figure 17. In Khor Kalba (Figure 12), the total area of mangroves increased by about 10.35 ha from 1990 to 2000 and 7.47 ha (0.26%) between 2000 and 2010, and then decreased slightly by about 0.36 ha between 2010 and 2019. It can be noticed that although the Khor Kalba mangroves expansion is small. The rate of growth of the Khor Kalba mangrove remains relatively stable during the last 29 years (Figures 12, 17).

In RAK areas, there is a negative change (mangrove loss) of -414.27 and -334.53 ha during the periods from 1990 to 2000 and from 1990 to 2019, respectively (Figures 13A–D). From 2000 to 2010, the mangrove increased slightly by 20.79 ha and more than double (58.95 ha) during the period from 2010 to 2019.

In Umm Al Quwain estuarine, similar changes (mangrove loss) were observed during the period from 2000 to 2010 (Figures 14E–H). The estimated lost area was about -122.76 ha. Conversely, the mangrove forests increased sharply by approximately 467.82 ha from 2010 to 2019. Between 1990 and 2019, the mangrove forests increased by about 380.52 ha (Figure 17).

Further south, the mangrove area in Ajman and Hammria creeks increased by about 26.32 ha from 1990 to 2000, then increased sharply by about 61.85 ha between 2000 and 2010 (Figures 15A–D). From 2010 to 2019, the mangroves increased slightly by about 28.87 ha. Since 1990, the total area of the mangroves increased by about 110.04 ha (1.88%).
In Ras Al Khor of Dubai, the lowest positive changes were observed to be during the periods from 1990 to 2000 (25.18 ha) and from 2000 to 2010 (4.36 ha), respectively. The largest changes were observed from 2010 to 2019 of about 60.39 ha (Figures 16, 17). Since 1990, an increase (positive change) in the mangrove area was reported. The estimated area of mangrove was 89.93 ha which represents the third after Umm Al Quwain and Ajman and Hammria Creeks. Generally, the mangrove areas showed an increase in all locations except the RAK and Umm Al Quwain areas. Comparing between the two mangrove ecosystems there appears to be stability in the mangrove area faced to the Gulf of Oman than those facing the Arabian Gulf.
DISCUSSION

Evaluation of the Performance of the Classifiers

The use and comparison of the performance of RF, KLR, NB, and ID algorithms permitted precisely mapping, classifying, and monitoring mangrove changes over 29 years in different ecological systems. Among the machine learning, RF yielded high precision, high recall and F1 score. This means that the RF has a powerful ability to map the mangrove forests in different ecosystems. These results also indicated that the RF and KLR are able to detect patches of dense mangrove that represent the majority of mangrove forests. On the other hand, NB yielded the lowest value for the F1 score (<0.8), and thus, lower ability to detect mangrove forests (Figure 6).

Random forest iterations predicted a slightly wider range of mangrove forest extent compared to those by KLR with F1 scores of 0.93 and 0.90, respectively. These findings were in good agreement with several studies that have indicated that RF algorithm has higher accuracy compared to CART (Mondal et al., 2019), support vector machine (SVM; Chen et al., 2020), and maximum likelihood (Ha et al., 2020) who compared machine learning methods against traditional classifiers and assist potentiality of groundwater.

These findings are also in accordance with Toosi et al. (2019) who compared different classifiers for monitoring mangrove changes and concluded that the RF which uses freely available

Landsat images, had performed best. Feng et al. (2018) implemented a novel change detection approach based on visual saliency and RF from multi-temporal high-resolution remote-sensing images. However, the opposite occurred for mapping mangrove using traditional classifiers.

The use of freely available Landsat images offered the ability to detect mangroves with overall accuracy comparable to the machine learning classifiers that involved the commercially remote sensing data such as LiDAR and WorldView-3 (Elmahdy and Mohamed, 2018; Toosi et al., 2019). Although the LiDAR, WorldView, and RapidEye images have high spatial resolution, the overall accuracy of RF only reached 82% (Ha et al., 2020). Other studies confirmed that the rotation forest (RoF) algorithm had a higher accuracy than the canonical correlation forest (CCF) in LULC classification, which has less sensitivity compared to the RF algorithm (Colkesen and Kavzoglu, 2017; Ha et al., 2020).

The proposed approach represents a significant modification in mangroves and LULC, as well as an enhancement of the performance of linear and non-linear classifiers. The machine learning algorithms, especially, when optimized, have their robust and a higher efficiency and quality of classification (Coppin and Bauer, 1996; Banfield et al., 2006; Wang et al., 2015; Gong et al., 2019).

The wide availability and ease of implementation of machine learning and free availability of Landsat images permitted an exceptional simplicity in processing vast amounts of remote sensing data promptly and at low costs (Elmahdy et al., 2019; Mondal et al., 2019).

**Mangrove Classification**

The use of machine learning algorithms and Landsat images provides up-to-date mangrove maps over multiple scales and able to discriminate between mangrove and non-mangrove areas and dense mangrove and dispersed mangrove patches than currently available. Between 1990 and 2019, the largest increase in mangrove forest extent occurred in the Umm Al Quwain Estuarine (1200 ha), while the largest decrease occurred in Dubai Natural Reserve (9 ha). Locally, these results are consistent with Su et al. (2009), Martins et al. (2016), Elmahdy and Mohamed (2018), and Gong et al. (2019) who optimized SVM, SAM, and RF classifiers and much more precisely than those mapped by Moore et al. (2013) using manual screen digitizing and field observations. This result is not consistent with our results, which indicated the mangrove of Khor Kalba was over 800 ha. The maps produced by Moore et al. (2013) had a lot of errors and bias and cannot be used as a reference. However, our results permitted better understanding of the NUAE mangrove's extent and can be updated in any future study, as well as building a new framework of the change detection. The proposed approach provides a better accuracy by reducing the level of noise associated with remote sensing data and helping in discriminating between mangrove and non-mangrove areas.

**Mangrove Change Detection**

During change detection, we observed that the ID algorithm using pair of bi-temporal mangrove classification of images with a thresed value between -1 and +1, is better than other algorithms such as an image to map and map to map (Elmahdy and Mohamed, 2018; Ma et al., 2019). The ID minimizes the errors created in one or two of the classification maps detecting greater presence pixels (+) and absence pixels (-) than those detected by image to map and map to map approaches. These approaches suffer from error propagation and differences in image calibration between Landsat sensors.

The rapid changes in mangrove forests during the period from 2010 to 2019 are much more than those during the period from 1990 to 2000. These changes are due to localized plantation activities and increased public and the local authorities
and non-mangrove areas and had an overall accuracy of more than 90% with an increase of 6.7% compared with supervised classifications. The results also demonstrated that the NUAE mangroves increased in all locations over the last 29 years. The majority of the NUAE mangroves are spatially distributed in the western coastal area than those in the eastern coastal area and the dense mangrove is much more than dispersed mangrove and much easier to discriminate and map. The mangrove forests of Khor Kalba facing the Gulf of Oman are more stable than those facing the Arabian Gulf. The highest distribution of mangroves was observed to be in the Umm Al Quwain estuarine, occupying an area of 1200 ha, while the lowest distribution of mangrove was reported to be in Ras Al Khor of Dubai covering an area of about 9 ha. Although there is a rapid change in the built-up area along with the eastern and western coastal areas of the NUAE, the mangroves had increased in sites of the NUAE and predicted to increase further over the next decade. The proposed approach can be applied to any area in arid and semi-arid regions and the results can be used as reference maps and updated by future studies as well as to provide useful ecological and environmental information.

**DATA AVAILABILITY STATEMENT**

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

**AUTHOR CONTRIBUTIONS**

SE processed the data and developed the RF. TA supervised the project and provided data. MM provided data and reviewed the methods. FH coded the RF algorithm. MA contributed to the scientific review. DS reviewed the manuscript. All authors contributed to the article and approved the submitted version.

**FUNDING**

The research has received funding under financial grant SCR1 18 Grant EN0-284 and the grant of UAE Space Agency z01-2016-001.

**ACKNOWLEDGMENTS**

We would like to thank the American University of Sharjah, United Arab Emirates and UAE Space Agency for supporting this research.

**REFERENCES**


The world's food and agriculture production is threatened by deforestation and other land-use practices that threaten coastal seagrass communities. This threat is compounded by the need to understand the impact of climate change on ecosystems and the services they provide. The Food and Agriculture Organization of the United Nations (2007) has noted that deforestation and land conversion have led to the loss of pasture land and arable land, and have increased the risk of soil erosion and desertification. These changes have also had a negative impact on the health and well-being of local communities.

In addition, the rapid development of cities in fast growing economies, such as coast regions in southern Peninsular Malaysia, has led to the conversion of wetland vegetation to fast growing economic regions. This has led to the loss of wetland vegetation and the degradation of coastal ecosystems, including seagrass beds.

Recent studies have provided insights into the processes that drive these changes. For example, a study by Giri et al. (2011) found that the rapid development of cities in coastal regions of southern Peninsular Malaysia has led to the conversion of wetland vegetation to fast growing economic regions. This has led to the loss of wetland vegetation and the degradation of coastal ecosystems, including seagrass beds.

Furthermore, the use of advanced remote sensing technologies, such as hyperspectral data, has provided a new tool for monitoring changes in coastal ecosystems. These technologies have been used to map changes in seagrass beds and other coastal ecosystems, and to understand the impact of climate change on these ecosystems.

In conclusion, the need to understand the impact of climate change on ecosystems and the services they provide is crucial. This requires the development of new tools and technologies, such as hyperspectral data, to monitor changes in coastal ecosystems. These tools will help us to better understand the impact of climate change on these ecosystems and to develop strategies to mitigate these changes.


**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Elmahdy, Ali, Mohamed, Howari, Abouleish and Simonet. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.