Mangrove Extent Mapping Using Machine Learning and a Fusion of Optical and Radar Images

Daniel Aja[1*], Michael Miyittah[1, 2] and Donatus Bapentire Angnuureng[1]

1. Africa Center of Excellence in Coastal Resilience, Center for Coastal Management, University of Cape Coast, Ghana

2. Department of Environmental Science, University of Cape Coast, North Campus, University Avenue, Science Building, Cape Coast, Ghana

* Correspondence: author: daniel.aja@stu.ucc.edu.gh

Abstract

Mangrove complexity and cloud cover effects, among others, make it difficult to classify mangrove forests in tropical coastal zones using simply passive remote sensing techniques. The method described in this paper combines optical and radar data to overcome the difficulties of identifying mangrove stands in cloudy areas. Google Earth Engine geospatial processing platform was used to extract multiple scenes of Landsat surface reflectance Tier1 and synthetic aperture radar. The images were enhanced by creating a feature that removes clouds from the optical data and using speckle filters to remove noise from the radar data. The random forest algorithm was used for mangrove classification. Classification was evaluated using three scenarios: classification of optical data only, classification of radar data only, and combination of optical and radar data. The scenario that uses both optical and radar data fared better, according to our findings. For the classification of optical data only, radar data only, and a combination of optical and radar data, the overall accuracy for 2019 was 98.9 %, 84.6 %, and 99.1%, respectively. This research has shown that it is possible to map mangrove correctly, enabling on-site conservation practices in the climate change environment.

Keywords: Algorithm; Decision Trees; Cloud Cover; NDVI; Radar; Tropical Coastal Zone

1. Introduction

In a habitat that serves as a buffer between terrestrial and marine ecosystems, mangroves flourish and thrive. Mangroves are globally significant ecosystems that are only found in coastal regions with mean monthly air temperatures above 20 °C, with very few exceptions (Maurya et al., 2021). Although mangrove forests only cover 1% of the total land area in the tropics, they are thought to be the region's most carbon-rich ecosystems (Maurya et al., 2021). Mangrove ecosystems have lately been incorporated into the IPCC's strategy for combating climate change (Lucas et al., 2014). Mangrove forests occupy around 2,746,500 ha in Africa in 2010 and support vulnerable coastal populations (Bunting et al., 2018). These forests also provide significant ecological services, such as improving water quality, natural coastal protection, and alternative livelihoods (Mondal et al., 2018). These ecosystems offer numerous vital functions, but they have undergone severe degradation and are in danger of disappearing (Breithaupt et al., 2012).

Mapping of mangrove ecosystem is crucial for an improved understanding of many coastal and sea processes. Remote sensing, a space-based technology, offers a tremendous ability to map and track changes in mangrove forests since data can be collected from an environment that is ordinarily difficult to access (Son et al., 2015). For more precise assessment of mangrove extent, several authors suggest combining Synthetic Aperture Radar with optical satellite data (Hu et al., 2020). The Google Earth Engine (GEE) offers a simple but effective mangrove assessment and monitoring framework with a high level of automation (Ghorbanian et al., 2021). GEE has a variety of machine learning classifiers such as Random Forest (RF), Naïve Bayes, and Support Vector Machine (SVM) which are suggested for the classification of multi-source data in complex environments (Maurya et al., 2021). Random forest (RF) algorithm has the

benefit of using a number of decision trees to vote the best classification for all pixel within the imagery to reduce the risk of overfitting, training time, sensitivity to outliers in training data and runs efficiently for huge dataset (Maurya et al., 2021).

The coast of Ghana, like other coastal zones, faces the threat of flooding and erosion with rising sea levels. Mangroves in this area may be threatened by sea level rise. There is a need to understand where mangroves are currently found and how they have changed over time. Therefore, the main objective of this study is to assess the performance of different satellite imageries for mangrove mapping and monitoring.

2. Materials and methods

2.1. Study location

This study area is located at Anlo Beach Wetland complex which is situated along the coastline of Ghana in Shama District, Western Region as shown in Figure (1). The area covers about 50.42km^2, lying approximately within latitudes 5º1'30"N and 5º3'5"N, and longitudes 1°34'30"W and 1°37'30"W and it is covered by mangroves which have been comparatively disturbed (Friends of the Nation. 2014). Hydrologically, the area lies within the plains of Pra River, which opens directly into the ocean.

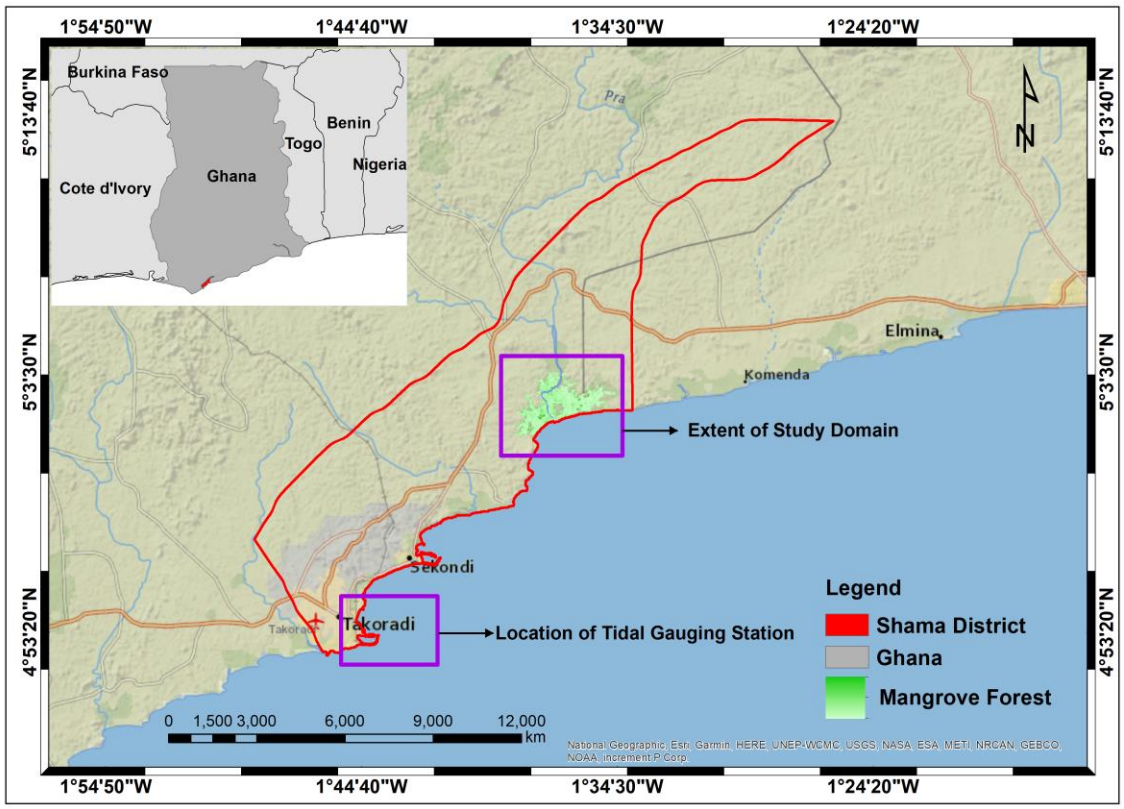

Figure 1. Map of study location.

2.2. Datasets and sources

Table 1. Description of datasets used for Mangrove extent mapping and quantification.

S/N	Data Type & Date	Description	Source
1	Sentinel-1 (2019)	A synthetic aperture radar (C-Band) with interferometric wide swath mode (IW), having a descending pass, a resolution of 25 m, dual polarization of VV and VH. Image Collection ID: ee.ImageCollection("COPERNICUS/S1_GRD"), more details can	Google Earth Engine Platform Database

S/N	Data Type & Date	Description	Source
		be found at https://developers.google.com/earth-engine/guides/sentinel1.	
2	Landsat 8 Surface Reflectance Tier 1 (2019)	Has been atmospherically corrected and contain five visible and near-infrared bands, two short wave infrared bands and two thermal infrared bands. Image Collection ID: ee.ImageCollection('LANDSAT/LC08/C01/T1_SR'). More details at https://www.usgs.gov/landsat-missions/landsat-surface-reflectance	Google Earth Engine platform database
3	Global Mangrove distribution vector (GMW) (2010)	A baseline global distribution map of mangroves for year 2010. GMW was produced by Aberystwyth University in collaboration with solo earth observation (soloEO) It provides geospatial information about mangrove extent and changes.	https://data.unep-wcmc.org/datasets/45

2.3. Mangrove extent mapping

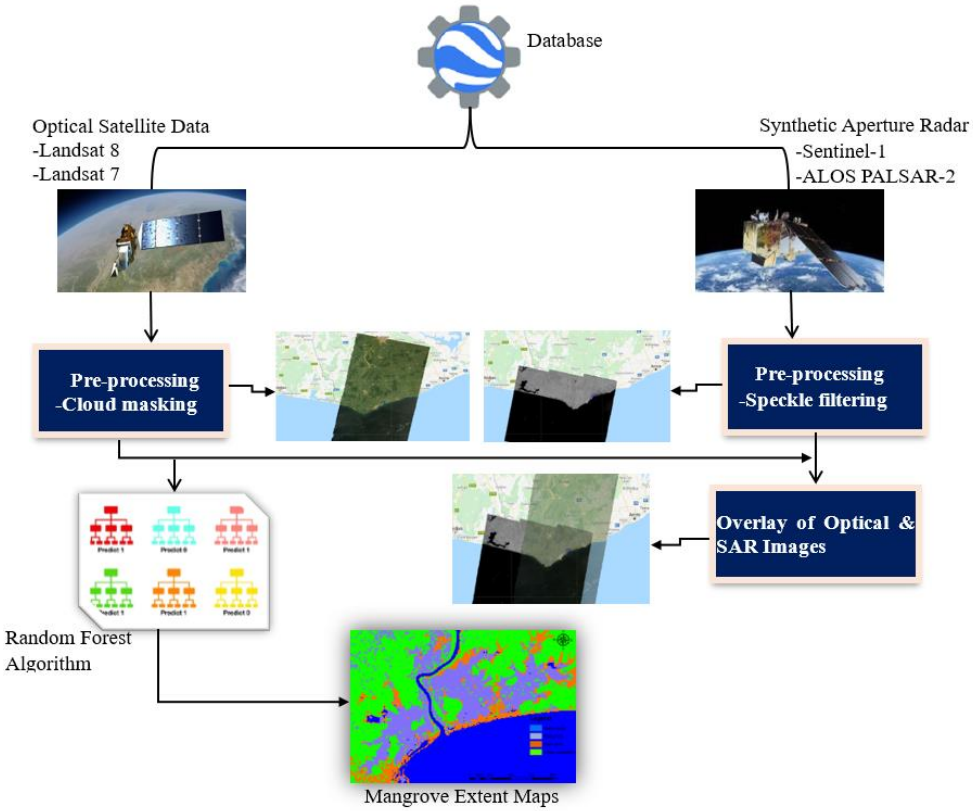

Figure 2. Flowchart of data extraction and iteration process for random forest model.

For this study, Google Earth Engine (GEE) cloud-based platform and random forest classification algorithm were used. Mangrove extent maps were generated by classifying both optical and radar images separately and in combination. The maps were created for the year 2019, to examine the performance of different

satellite data. The Google Earth Engine database was filtered for Sentinel-1 imagery that is in Wide Swath (IW) interferometric mode, descending pass, 25 m resolution, VH and VV polarization and falls within the region of interest.

The Sentinel-1 dataset was filtered by date to retrieve image from 2019. The corresponding Landsat 8 images were also extracted from the Google Earth Engine database. The images were loaded into the Google Earth Engine code editor and visualized in the layer bar. A speckle filter was applied to the Sentinel-1 image to minimize speckle noise (Ayman et al., 2017) and the speckle filtered image was also added to the layer bar (Figure 2). The optical image was improved by creating a function that masks cloud shadows (Giri et al., 2011). The normalized difference vegetation index (NDVI) was calculated from the optical image to obtain a composite image, which was used as an information layer to inform the classifier (Nathan et al., 2018).

2.3.1. Construction of random forest model

Supervised classification was conducted using Random Forest algorithm as shown in Figure 2. The approach used follows the method described by Erika et al., (2020), which involves collecting representative samples of backscatter values for each landcover class of interest. The random forest algorithm (RF) was run using 100 trees and 5 randomly selected predictors per split (Ghorbanian et al., 2021). Training and validation data is based on field campaigns conducted between December 2020 and April 2021. The global mangrove distribution vector (Table 1) was needed as a reference data to guide the creation of training samples.

2.3.2. Synthetic aperture radar (SAR) classification

The speckle-filtered Sentinel-1 (VH) image was displayed in the map section of GEE and a polygon symbol was selected from the geometry imports box next to the geometry drawing tool to add training data. Each new layer created, for example ‘open water’ represents one class within the training data and it was saved as featureCollection called landcover. Training samples were selected for four different land cover classes: open water, mangroves, bare land/built-up, and other vegetation/wetlands. The defined classes were then merged into a single collection called ‘new FeatureCollection’. The ‘new FeatureCollection’ created was used to extract backscatter values for each landcover identified for the Sentinel-1 image. The Sentinel-1 (SAR) image was defined with the code in bracket (var final = ee.Image.cat(SARVV_filtered,SARVH_filtered)) and the training data was created by overlaying the training points (new FeatureCollection) on the image. This created a ‘training point’ statistics based on the classes (new FeatureCollection) and was used to ‘train’ the random forest classifier. The classification was ‘run’ and result displayed on the ‘layers’ bar.

2.3.3. Landsat image classification

The ‘new FeatureCollection’ created was used to extract reflectance values for each landcover class from the defined Landsat 8 image (var trainingl8 = composite.select(bandsl8).sampleRegions({}). The ‘training’ data was created by overlaying the training points (new FeatureCollection) on the image and used 'B1', 'B2', 'B3', 'B4', 'B5', 'B6', 'B7', 'NDVI' to generate the statistics. This was used to ‘train’ the random forest classifier. The classification was ‘run’ and result displayed on the ‘layers’ bar.

2.3.4. Both Landsat and SAR classification

Now, the ‘new FeatureCollection’ created was used to extract reflectance and backscatter values for each landcover classes from the Landsat 8 and Sentinel-1 images to be used in the classification. Both the Landsat 8 and Sentinel-1 images were defined (var opt_sar = ee.Image.cat(composite, SARVV_filtered,SARVH_filtered)) and the ‘training’ data was created by overlaying the ‘training’ points on the defined image (Figure 2). This was used to ‘train’ the random forest classifier. The classification was ‘run’ and result displayed on the ‘layers’ bar.

2.3.5. Classification accuracy assessment

A total of 2131 training sample points were created. The sample points were randomly split for 'training' and validation; 80% (1705 points) of the sampling points were used to 'train' the model while 20% (426 points) was used for validation.

This was done to remove any systematic error as a result of using the same pixels to train and validate classifiers (Pimple et al., 2018). The accuracy of the classification in GEE was assessed using confusion matrix based on the classifier. The overall accuracy and the kappa coefficient were calculated using the following formular:

$$Overall\ Accuracy = \frac{Number\ of\ correctly\ classified\ pixels\ (sum\ of\ diagonal)}{Number\ of\ total\ sampled\ pixels} * 100 \quad (1)$$

$$Kappa = \frac{P_0 - P_C}{1 - P_C} \quad (2)$$

Where.

P_0= Observed accuracy and P_C= Chance agreement.

3. Results

Several iterations were run using Sentinel-1, Landsat 8 separately and a combination of both (i.e., Sentinel-1+ Landsat 8) which represents the three scenarios for 2019. The results of mangrove extent and other land cover changes are presented in Figure 3 and 4.

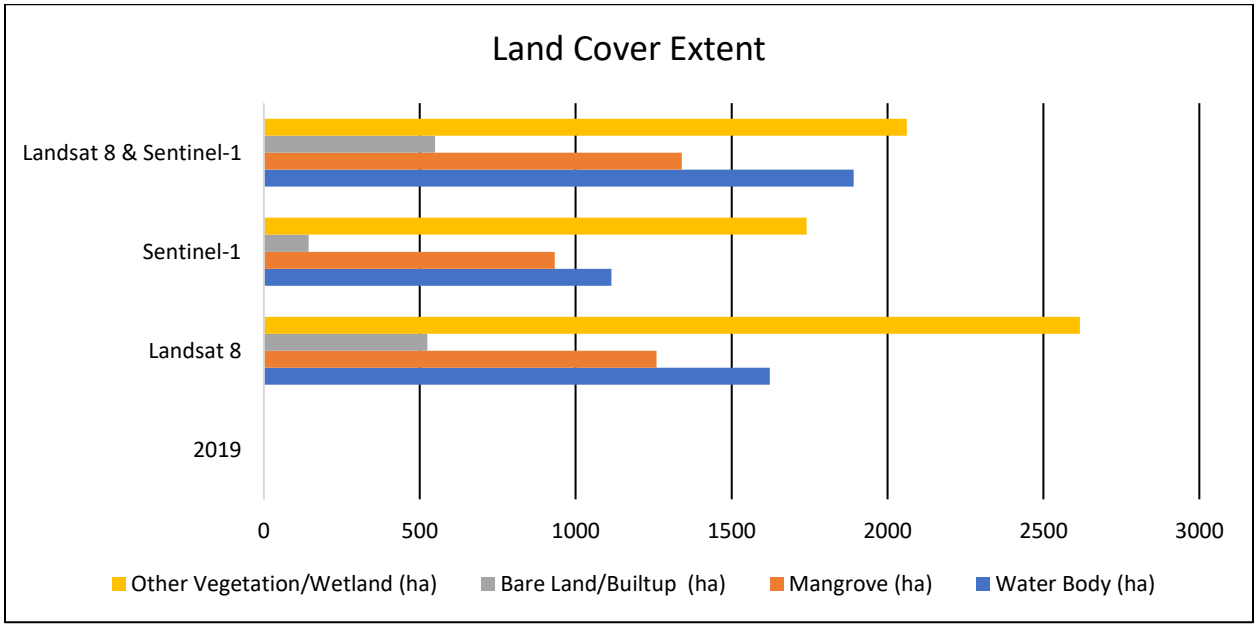

Figure 3. LULC extent for different classification scenario.

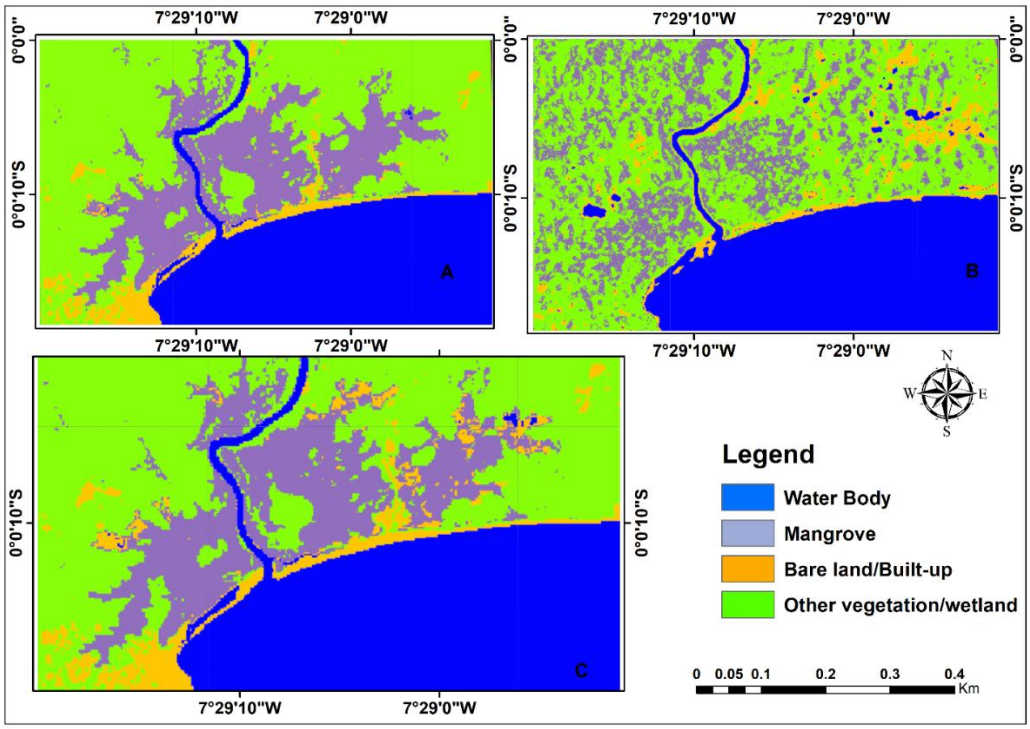

Figure 4. Different scenarios of mangrove extent maps for the year 2019. (a) classification scenario using optical image only (Landsat 8); (B) Classification scenario using sentinel-1 image only; (C) Classification scenario using both optical (Landsat 8) and Sentinel-1.

Three main classification scenarios were established to map the extent of mangroves and other land cover classes for the period under study: classification of optical data only, classification of SAR data only, and the third scenario combined both optical and SAR data (Figure 4 and 5). The result of Landsat 8 data only (2019) showed Mangrove extent of 1259 ha, water body extent of '1622 ha', Bare land extent of 524 ha and other vegetation extent of 2617 ha (Figure 3). The overall classification accuracy for Landsat 8 was 99.1% with Kappa Coefficient of 0.797. The second classification scenario showed Mangrove extent of 933 ha, water body extent of '1115 ha', Bare land extent of 144 ha and other vegetation extent of 1741 ha for Sentinel-1 (Figure 3). The overall classification accuracy for Sentinel-1 classification was 84.6% with Kappa Coefficient of 0.687. The third classification scenario showed Mangrove extent of 1340 ha, water body extent of '1891 ha', Bare land extent of 549 ha and other vegetation extent of 2062 ha for Sentinel-1and Landsat 8 combined. The overall classification accuracy for both Sentinel-1 and Landsat 8 when combined together was 98.9% with Kappa Coefficient of 0.828 (Table 2 - 4).

Table 2. Confusion/error matrix of land cover classification using landsat 8 image.

Classes	Open Water	Mangroves	Bare Land	Vegetation/ Wetland	Row Total	User's Accuracy (%)
Open Water	**81**	0	0	0	81	100
Mangroves	0	**642**	1	4	647	99.2
Bare Land	0	0	**32**	0	32	100
Vegetation/Wetland	0	9	0	**936**	945	99
Column Total	81	651	33	940	**1705**	
Producer's Accuracy (%)	100	98.6	96.9	99.5		

Overall Accuracy = 99.18%; Kappa Coefficient = 0.797.

Table 3. Confusion/error matrix of land cover classification using sentinel-1 data.

Classes	Open Water	Mangroves	Bare Land	Vegetation/ Wetland	Row Total	User's Accuracy (%)
Open Water	**79**	0	2	0	81	97.5
Mangroves	0	**521**	0	126	647	80.5
Bare Land	2	0	**23**	7	32	71.9
Vegetation/ Wetland	0	119	6	**820**	945	86.8
Column Total	81	640	31	953	**1705**	
Producer's Accuracy (%)	97.5	81.4	74.2	86		

Overall Accuracy = 84.6%; Kappa Coefficient = 0.687.

Table 4. Confusion/error matrix of land cover classification using a combination of landsat 8 and sentinel-1 images.

Classes	Open Water	Mangroves	Bare Land	Vegetation/ Wetland	Row Total	User's Accuracy (%)
Open Water	**81**	0	0	0	81	100
Mangroves	0	**635**	0	12	647	99.2
Bare Land	4	0	**28**	0	32	100
Vegetation/Wetland	0	3	0	**942**	945	99
Column Total	85	638	28	954	**1705**	
Producer's Accuracy (%)	95.3	99.5	100	98.7		

Overall Accuracy = 98.9%; Kappa Coefficient = 0.828.

4. Discussion

The results show that there are differences in all three classification scenarios (Figure 4). For example, classification using Synthetic Aperture Radar data showed that most structural aspects were captured but underestimated the vegetation cover and this is consistent with observations in other studies (Carreiras et al., 2013). In contrast, optical satellite image classification captured the tree canopy more (Erika et al., 2020) but seems to overestimated (Mondal et al., 2019) the extent.

We used confusion matrices to provide detailed statistical information for each classification scenario. The confusion matrix for Sentinel-1 image classification alone showed that out of 647 pixels which were identified as Mangrove, 521 pixels were correctly classified while the confusion matrix for the corresponding optical image alone showed that out of 647 pixels which were identified as Mangroves, 642 pixels were correctly classified (Table 2 - 4). It was also revealed that the Sentinel-1 image alone tend to underestimate the mangrove vegetation canopy (Figure 4). On the other hand, Landsat 8 alone tend to overestimate the vegetation cover. The overall classification accuracy for the Sentinel-1 image was 84.6%, while the overall accuracy for the Landsat 8 alone was 99.1%. The overall accuracy when both images were combined was 98.9% (Table 2 - 4). The third classification scenario that combines optical and radar data yielded the best classification results for 2019 as the classes were relatively well distributed, capturing both clustered mangroves and mangrove patches near the water body (Figure 4). The visual interpretation showed that the third classification scenario achieved a better result, indicating the high potential of this

mangrove mapping method and agrees with the work of Ghorbanian et al. (2021). Despite the fact that the same 'training sample' was used to train the classifier, the accuracies differ depending on the scenario. The scenario that combined both optical and radar data showed better agreement and less confusion compared to when either optical or radar data was used.

This study confirms that combining synthetic aperture radar data with optical satellite data improves the accuracy of mangrove mapping, as recommended by several authors (Attarchi and Gloaguen, 2014; Ayman et al., 2017). The resulting classification is consistent with other studies that used random forest algorithms for land cover classification (Ghorbanian et al., 2021; Beselly et al., 2021).

5. Conclusion

One of the most pressing environmental calamities of our time is the worldwide devastation of tropical and subtropical mangrove ecosystems. Without tackling deforestation and promoting increased restoration of mangroves and other forests, it's possible that the world won't be able to meet the Sustainable Development Goals. In order to produce more precise mangrove extent maps, we developed a method in this research for combining the pertinent database in a spatial framework using the Google Earth Engine platform and a random forest algorithm. In this study, it was shown that cloud computing methods and machine learning algorithms, such Google Earth Engine, have the ability to accurately quantify mangrove stands as well as a variety of other land uses, particularly in cloud-prone regions. This would enable more precise estimate of mangrove changes at local and regional levels.

Acknowledgements

The authors would like to thank the Management of Africa Center of Excellence in Coastal Resilience (ACECoR), UCC for funding this project.

References

Ayman AH, Olena D, Islam AE, Gunter M (2017) Integration of SAR and Optical Remote Sensing data for mapping of mangroves extents. In: From Science to Society: The Bridge provided by Environmental Informatics. (Ed. Benoît Otjacques, Patrik Hitzelberger, Stefan Naumann, Volker Wohlgemuth) Germany: Shaker Verlag GmbH, p. 1 - 8

Beselly, S.M.; van der Wegen, M.; Grueters, U.; Reyns, J.; Dijkstra, J.; Roelvink, D. (2021**)** Eleven Years of Mangrove–Mudflat Dynamics on the Mud Volcano-Induced Prograding Delta in East Java, Indonesia: Integrating UAV and Satellite Imagery. Remote Sensing, 13, 1084. https://doi.org/10.3390/rs13061084

Breithaupt J L, Smoak J M, Smith T J, Sanders C J, Hoare A (2012) Organic carbon burial rates in mangrove sediments: Strengthening the global budget. Global Biogeochemical Cycles 26(3).

Carreiras JMB, Melo JM, Vasconcelos MJ (2013) Estimating the above-ground biomass in Miombo Savanna woodlands (Mozambique, East Africa) using L-band synthetic aperture radar data. Remote Sensing, 5(4):1524–1548

Friends of the Nation (2014), Assessment of flora and fauna of ecological and socioeconomic significance within the Anlo Beach Wetland Complex for improved management and livelihood outcomes, Parks and Gardens, Adiembra, 60 pp.

Ghorbanian, A.; Zaghian, S.; Asiyabi, R.M.; Amani, M.; Mohammadzadeh, A.; Jamali, S. (2021) Mangrove Ecosystem Mapping Using Sentinel-1 and Sentinel-2 Satellite Images and Random Forest Algorithm in Google Earth Engine. Remote Sensing, 13, 2565. https://doi.org/10.3390/rs13132565

Giri C, Ochieng E, Tieszen LL, Zhu Z, Singh A, Loveland T, Masek J, Duke N (2011) Status and distribution of mangrove forests of the world using earth observation satellite data. Global Ecology and Biogeography, 20: 154–159.

Hu T, YingYZ, Yanjun S, Yi Z, Guanghui L, Qinghua G (2020) Mapping the Global Mangrove Forest Aboveground Biomass Using Multisource Remote Sensing Data. Remote Sensing, 12: 1690. doi:10.3390/rs12101690

Lucas R, Rebelo L, Fatoyinbo L, Rosenqvist A, Itoh T, Shimada M, Simard M, Pedro WS, Nathan T, Carl T, Arnon A, Joao C, Lammert H (2014) Contribution of L-band SAR to systematic global mangrove monitoring. Marine and Freshwater Research, 65:589–603 http://dx.doi.org/10.1071/MF13177

Maurya K, Mahajan S, Chaube N (2021) Remote sensing techniques: mapping and monitoring of mangrove ecosystem - a review. Complex & Intelligent Systems, 7:2797–2818. https://doi.org/10.1007/s40747-021-00457-z

Mondal P, Trzaska S, de Sherbinin A (2018) Landsat-Derived Estimates of Mangrove Extents in the Sierra Leone Coastal Landscape Complex during 1990–2016. Sensors, 18:12.

Nathan T, Peter B, Richard L, Andy H, Ake R, Temilola F (2018) Mapping Mangrove Extent and Change: A Globally Applicable Approach. Remote Sensing, 10:1466. doi:10.3390/rs10091466

Pimple U, Simonetti D, Sitthi A, Pungkul S, Leadprathom K, Skupek H, Som-ard J, Gond V, Towprayoon S (2018) Google Earth Engine Based Three Decadal Landsat Imagery Analysis for Mapping of Mangrove Forests and Its Surroundings in the Trat Province of Thailand. Journal of Computer and Communications, 6: 247-264. https://doi.org/10.4236/jcc.2018.61025