Article

Integrating SAR, Optical, and Machine Learning for Enhanced Coastal Mangrove Monitoring in Guyana

Kim Chan-Bagot 1,2,*, Kelsey E. Herndon 3,4, Andréa Puzzi Nicolau 2,5, Vanesa Martín-Arias 3,4, Christine Evans 3,4, Helen Parache 6, Kene Mosely 1, Zola Narine 1 and Brian Zutta 2,5

1 National Agricultural and Research Extension Institute (NAREI), Georgetown, Guyana; kmoseley@narei.gov.gy (K.M.); znarine@narei.gov.gy (Z.N.)
2 SERVIR Amazonia, Cali 76001, Colombia; apnicolau@sig-gis.com (A.P.N.); bzutta@sig-gis.com (B.Z.)
3 Earth System Science Center, University of Alabama in Huntsville, Huntsville, AL 35899, USA; keh0023@uah.edu (K.E.H.); vm0043@uah.edu (V.M.-A.); cae0004@uah.edu (C.E.)
4 NASA SERVIR Science Coordination Office, NASA Marshall Space Flight Center, Huntsville, AL 35812, USA
5 Spatial Informatics Group (SIG), San Francisco, CA 94566, USA
6 NASA IMPACT, NASA Marshall Space Flight Center, Huntsville, AL 35812, USA; helen.blue.parache@nasa.gov

* Correspondence: kbagot19@gmail.com

Abstract: Mangrove forests are a biodiverse ecosystem known for a wide variety of crucial ecological services, including carbon sequestration, coastal erosion control, and prevention of saltwater intrusion. Given the ecological importance of mangrove forests, a comprehensive and up-to-date mangrove extent mapping at broad geographic scales is needed to define mangrove forest changes, assess their implications, and support restoration activities and decision making. The main objective of this study is to evaluate mangrove classifications derived from a combination of Landsat-8 OLI, Sentinel-2, and Sentinel-1 observations using a random forest (RF) machine learning (ML) algorithm to identify the best approach for monitoring Guyana’s mangrove forests on an annual basis. Algorithm accuracy was tested using high-resolution planet imagery in Collect Earth Online. Results varied widely across the different combinations of input data (overall accuracy, 88–95%; producer’s accuracy for mangroves, 50–87%; user’s accuracy for mangroves, 13–69%). The combined optical–radar classification demonstrated the best performance with an overall accuracy of 95%. Area estimates of mangrove extent ranged from 908.4 to 3645.0 hectares. A ground-based validation exercise confirmed the extent of several large, previously undocumented areas of mangrove forest loss. The results establish that a data fusion approach combining optical and radar data performs marginally better than optical-only approaches to mangrove classification. This ML approach, which leverages free and open data and a cloud-based analytics platform, can be applied to mapping other areas of mangrove forests in Guyana. This approach can also support the operational monitoring of mangrove restoration areas managed by Guyana’s National Agricultural and Research Extension Institute (NAREI).

Keywords: remote sensing; blue carbon; wetlands; sustainable development goals; Google Earth Engine; random forest

1. Introduction

Mangroves are a woody halophyte inhabiting the upper intertidal zones of saltwater areas, primarily in tropical and subtropical regions [1–3]. These forests grow along muddy beaches with low hydrodynamic intensity, particularly in areas with extensive river estuaries and deltas where sediments are continuously supplied by upland catchments [4]. Globally, mangroves are estimated to occupy approximately 152,000 km² across 124 countries [5]. Mangrove species can be found in almost 75% of the world’s coastlines, between latitudes of 25° north and south [1,5,6]. Mangroves have a number of attributes that make them physically and ecologically distinct. These characteristics include aerial
roots, high rates of canopy production, extremely efficient nutrient storage mechanisms, capacity to adapt to saline water, and rapid carbon sequestration, among other morphological characteristics [7].

Mangrove forests are among the most productive and important ecosystems in the world because they provide a wide range of ecological and socioeconomic benefits to human society [7,8]. They provide valuable ecosystem services, including the prevention of coastal erosion, saltwater intrusion, and flooding of coastal communities. Mangrove forests also help to stabilize the coastline and subsequently mitigate the devastating impacts of natural disasters, such as tsunamis, erosion caused by hurricanes, and flooding [9]. In addition, mangroves regulate nutrients and sequester carbon. In fact, mangroves have a potential carbon sequestration rate that is 50 times greater than that of many other types of tropical forests due to their ability to store high levels of carbon below ground and in mangrove sediment soils [10]. Mangroves also support biodiversity by providing vital breeding and nursing grounds for many fishes, invertebrate species, and coastal animals. By providing habitats for a diverse array of marine and terrestrial animals, mangrove forests have the potential to boost ecotourism and support the local economy [11,12].

The three primary species of mangroves in Guyana are black mangroves (Avicennia germinans), red mangroves (Rhizophora mangle), and white mangroves (Laguncularia racemosa). Significant portions of the coast are productive monoculture stands of black mangroves [13,14], with secondary establishments of red mangrove usually within riverine environments. White mangroves vary in density along the coast [14]. Conocarpus erectus, popularly known as buttonwood, is considered a mangrove associate and usually grows at higher elevations within the mangrove ecosystem [15].

In Guyana, mangrove ecosystems play a vital role in forming a natural soft defense for coastal protection [16]. Approximately 90% of Guyana’s population and economic activities reside on the coastal plain, which lies between 0.5 to 1 m below high and spring tide levels from the Atlantic Ocean, making the coast very vulnerable to flooding, sea level rise, and erosion from the ocean [17]. The mapping and monitoring of mangroves are important steps in gathering information on geomorphologic and environmental conditions [18,19]. This paper seeks to address the challenges of mapping mangrove ecosystems by using free and open remote sensing datasets and leveraging a cloud computing platform. Mangroves have distinct spectral and spatial characteristics that are distinguishable in satellite images, making remote sensing an effective technique for mapping and monitoring this important ecosystem [9]. Remote sensing techniques can be used to estimate the mangrove forest cover that extends over large areas, especially in areas that are inaccessible or costly to access in field surveys [19].

Background

The National Agricultural Research and Extension Institute (NAREI) in Guyana has mapped mangroves manually from 2010 to 2013 as part of the Guyana Mangrove Restoration Project initiative. Subsequently, the mandate to restore and manage mangroves was permanently integrated into NAREI [20], with 2018 being the first year remote sensing was introduced into NAREI’s monitoring operations. While fieldwork is an essential component of management efforts, the personnel, time, equipment, and logistical organization needed to conduct fieldwork make it a costly initiative. Remote sensing provides a cost-effective method by which to map mangroves, expanding the scope and scale of traditional fieldwork. The ability to use freely available data on cloud-based platforms can revolutionize the workflow of regularly monitoring environments like mangrove forests. Prior remote sensing work for mangroves in Guyana was carried out by Conservation International (CI), who developed a 2018 mangrove extent map based on a random forest (RF) classification of Sentinel-2 optical imagery [21]. The 2018 CI classification was run with multiple vegetation indices, and the resulting polygon of mangrove extent was manually edited using expert, in-region knowledge. This initial effort of mapping Guyana’s mangrove extent using remote sensing noted the exclusive use of optical data as a limitation in the study and
recommended incorporating lidar or radar data in mapping work, given the considerable cloud cover over Guyana’s coastline.

In partnership with SERVIR, a joint program between the United States Agency for International Development (USAID) and NASA aimed at connecting decision makers across the world with earth observations, NAREI, in the context of SERVIR-Amazonia, is working to use remote sensing technology to develop a national mangrove monitoring system. This study supports these efforts by evaluating an optical, synthetic aperture radar (SAR), and fusion approach to generate annual mangrove extent maps as part of NAREI’s monitoring and restoration efforts.

There have been extensive efforts to map global distributions of mangrove ecosystems [9,21–26]. Giri et al. (2005) [27] produced a global mangrove extent product for the year 2000 using Landsat Thematic Mapper (TM) imagery by applying an unsupervised ISODATA clustering algorithm combined with an iterative labeling process. Since then, numerous projects have attempted to improve upon these results and have expanded the types of remotely sensed data used to create mangrove classifications [21–23,26]. More recently, global products evaluating drivers of mangrove loss [24,28] have been developed to support decision making within conservation and restoration initiatives. The publication led by Bunting et al. (2018) [21] described the use of an extremely randomized tree classification to both SAR and optical data to produce the 2010 baseline for Global Mangrove Watch (GMW). Some studies employed a combination of optical and SAR data applied to a supervised machine learning (ML) approach [26,29]. Gumbricht et al. (2017) [25] exclusively used optical data in a supervised ML approach to map wetlands and peatlands throughout South America. Thomas et al. (2017) [24] conducted a manual visual interpretation of exclusively L-band SAR composites validated with optical imagery. Other studies extracted mangrove extent from several pre-existing databases, such as Terrestrial Ecosystems of the World, to create synthesized databases [22,23].

Techniques for mapping mangrove extent are clearly varied. In many cases, approaches exclusively using optical remote sensing [9,30] to map mangroves are insufficient due to persistent cloud cover in the tropics, the primary location of mangrove forests. Cloud-free optical imagery for mangrove mapping on an annual or subannual basis can be difficult to obtain [31]. Figure 1 illustrates the availability of Landsat-8 observations after cloud and cloud shadow are masked [32]. While Landsat-8 and Sentinel-2 data availability is relatively scarce here, the number of valid observations also varies greatly by year, demonstrating an opportunity to incorporate data unencumbered by persistent cloud cover to help fill these gaps. Specifically, as Figure 1 shows for the year 2020, there are many areas with fewer than approximately 10 cloud-free observations.

SAR data present multiple options and advantages for mapping mangrove forest ecosystems. Two critical characteristics of SAR data are different wavelengths, particularly X-band, C-band, and L-band listed from shortest to longest wavelength, and a variety of backscatter mechanisms [33]. The different wavelengths have different strengths and optimal applications, with L-band having the greatest capacity to penetrate vegetation canopies to reach the ground. Scattering mechanisms, namely, rough surface scattering, volume scattering, and double-bounce scattering, vary based on the wavelength and the surface characteristics of the terrain interacting with the radar signal [33]. Double-bounce scattering occurs when the signal bounces from a tall vertical structure to the ground, then back to the sensor. The double-bounce mechanism can be leveraged to detect understory floods or vegetated areas with high moisture content, such as mangrove forests [34].

Several other studies have demonstrated the suitability of SAR data across a variety of wavelengths for a wide range of applications, including mangrove mapping [29,35–39]. The study presented in this paper builds on these previous efforts by comparing a data fusion approach with optical-only and radar-only approaches. It also expands the evaluation of the algorithms to include an independent accuracy assessment and field validation exercise.
Figure 1. Cloud-free optical imagery (Landsat-7, Landsat-8, Sentinel-2) over Guyana for the year 2020.

The extent of Guyana’s coastal mangroves varies greatly across global mangrove extent products, and guidelines for identifying which dataset is the most appropriate are limited (see the Sustainable Development Goals [40] for one perspective). In addition, these global datasets may not support local, national, or regional mangrove management and monitoring efforts as localized accuracy of the classification may not be optimized. The definitions and uncertainty of what is actually being mapped may also not meet local requirements (e.g., definition of mappable mangrove, appropriate time step, etc.). This emphasizes the need for more localized mangrove mapping products in some cases.

Since work on the restoration effort began in 2010, there has been a noticeable gap in Guyana’s mangrove ecosystem systematic monitoring. For sustainable mangrove management, it is essential to comprehend the extent of the mangrove cover, the change (gains and losses), and the factors for change. Prior to this study, NAREI’s mangrove monitoring program relied on field observations and Google Earth data to locate possible mangrove restoration locations and areas of natural deterioration and regeneration. Because losses were not identified until they were apparent from the coastline, reliance on ground assessments led to reactive rather than proactive management. This resulted in coastal flooding and the loss of arable land owing to saltwater intrusion in numerous cases. The study’s methodology will enable yearly systematic monitoring of Guyana’s coastal mangroves, providing NAREI information to support decision making regarding critical sites that need additional protection, eroding sites that need intervention, baseline data to support carbon storage evaluation, and areas that can support community livelihood projects.

This paper also provides the first comparison of mangrove extent mapping approaches in Guyana using remotely sensed data. This analysis is needed to identify the most appropriate technique to develop a mangrove monitoring service for Guyana in the context of NAREI and the SERVIR-Amazonia program. We evaluate the performance of an ML algorithm to classify three different data combinations to determine the best approach for mapping annual mangrove extent in a tropical, cloud-persistent, coastal ecosystem. The results from this analysis will inform an operational national mangrove monitoring system for Guyana.

2. Materials and Methods

2.1. Study Area

Guyana consists of 6 coastal regions: Barima-Waini (Region 1), Pomeroon-Supenaam (Region 2), Essequibo Islands-West Demerara (Region 3), Demerara-Mahaica (Region 4), Mahaica-Berbice (Region 5), and East Berbice-Corentyne (Region 6). Our study area in
Pomeroon-Supenaam (Region-2) is approximately 6195 km² and is bounded by the Atlantic Ocean to the north, the Pomeroon River Estuary to the west, the Essequibo River Estuary to the east, and extends all the way to the Supenaam River to the south (Figure 2) [41]. This region is of interest given that NAREI began mangrove restoration efforts in this region in 2013.

The coast of Pomeroon-Supenaam experiences a humid tropical climate, with an annual average rainfall that varies between 1778 and 2800 mm and an average daily temperature of 25 degrees Celsius. The warmest months are usually September and October, while the coldest are January and February [42]. The area has a relatively flat, low-lying elevation, with most of the population concentrated along the coast. The major land use and land cover (LULC) types include mangrove forests, wetlands, cropland, water, and human settlements. Rice cultivation is the primary economic activity in the region [41].

2.2. Data Acquisition and Preprocessing

The mapping workflow (Figure 3) encompasses the use of optical and radar datasets and an ML algorithm to map the extent of mangrove forests in the study area. The workflow includes dataset acquisition and preprocessing, ML model training and classification, model validation, and accuracy assessment.
Data from the Landsat-8 Operational Land Imager (OLA), Sentinel-2 MultiSpectral Instrument (MSI), and Sentinel-1 C-band SAR ground-range detected (GRD) product were acquired through Google Earth Engine (GEE) [43]. GEE is a cloud-based analysis and visualization platform with access to a large catalog of geospatial datasets. Given that satellite imagery, particularly SAR imagery, makes for very large files that can be slow to download and process on a local machine, a cloud-based platform was essential for this analysis and is required for NAREI’s future mangrove monitoring system.

In GEE, the Landsat-8 Collection 2 Tier 1 dataset and the Sentinel-2 datasets were filtered to the year 2020. We applied a cloud mask to each scene using the quality assessment bands. We created annual composites using the median value of each pixel across the full year of imagery (Table 1).
Table 1. Number of scenes acquired from Landsat-8, Sentinel-2, and Sentinel-1 satellites, along with the sensors’ repeat cycle and acquisition dates.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Repeat Cycle</th>
<th>Number of Scenes</th>
<th>Acquisition Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landsat-8 (optical)</td>
<td>16 days</td>
<td>7</td>
<td>01-01-2020 to 31-12-2020</td>
</tr>
<tr>
<td>Sentinel-2 (optical)</td>
<td>10 days</td>
<td>216</td>
<td>01-01-2020 to 31-12-2020</td>
</tr>
<tr>
<td>Sentinel-1 (radar)</td>
<td>12 days</td>
<td>59</td>
<td>01-01-2020 to 31-12-2020</td>
</tr>
</tbody>
</table>

In GEE, the Sentinel-1 dataset contains GRD scenes that are preprocessed with the Sentinel-1 Toolbox [44] using the following steps: thermal noise removal, radiometric calibration, and terrain correction using Shuttle Radar Topography Mission (SRTM) 30. The final terrain-corrected values are converted to decibels via logarithmic scaling. We used the Sentinel-1 analysis-ready data (ARD) script made available by [45] to filter and process Sentinel-1 images for our area of study. The script was set to filter images acquired in 2020 with VV and VH polarizations and from both ascending and descending orbits, resulting in 59 total images. Furthermore, we applied the refined Lee speckle filter [46] using a 7 by 7 kernel to each image individually with the aim of despeckling the SAR data [47]. We did not apply terrain flattening as our area of study is generally flat and recent work suggests minimal impact on SAR data values [48]. The resulting images were exported as a GEE asset on the linear scale. We calculated the modified Radar Forest Degradation Index [49] for each image (reference Section 2.2.1) and created a final composite with six bands as follows: (1) VV ascending using a minimum reducer, (2) VH ascending using a minimum reducer, (3) VV descending using a minimum reducer, (4) VH descending using a minimum reducer, (5) mRFDI ascending using a median reducer, and (6) mRFDI descending using a median reducer. Each band was then normalized to values from zero to one per best practices [50]. The final composite was reprojected to have a 30 m spatial resolution to maintain consistency across datasets.

2.2.1. Spectral and Backscatter Indices

Multiple red–green–blue composites and standard spectral indices were created to employ in the ML classification as input variables (Table 2). The spectral indices include the Normalized Difference Vegetation Index (NDVI), which measures vegetation greenness as a proxy for health [51]; Normalized Difference Moisture Index (NDMI), which quantifies vegetation with high water content [52]; Modified Normalized Difference Water Index (MNDWI), which distinguishes open water features [53]; Green Chlorophyll Vegetation Index (GCVI), to quantify leaf chlorophyll content [54]; and modified Radar Forest Degradation Index (mRFDI), which can distinguish between different vegetation varieties using the difference between the VV and VH polarizations in SAR data [49] to potentially distinguish mangroves from other vegetation types. We also calculated additional indices, including the simple ratio (SR), to quantify vegetation density [55]; Band Ratio 54, which is helpful in distinguishing water bodies from vegetation [56]; and Band Ratio 35, to enhance urban areas and distinguish water from vegetation [57].

These indices have been selected based on previous studies that mapped mangroves [29,58] and the fact that mangrove ecosystems are a combination of coastal vegetation and tidal conditions; therefore, the vegetation and water indices are expected to help distinguish between mangroves and nonmangroves.
Table 2. Band combinations and formulas for spectral indices NDVI, NDMI, MNDWI, SR, Band Ratio 54, Band Ratio 35, GCVI for Optical and mRFDI Radar Images.

<table>
<thead>
<tr>
<th>Index</th>
<th>Sensor</th>
<th>Bands</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDVI</td>
<td>Landsat-8</td>
<td>Band 4 and Band 5</td>
<td>$\frac{\text{NIR} - \text{RED}}{\text{NIR} + \text{RED}}$</td>
</tr>
<tr>
<td></td>
<td>Sentinel-2</td>
<td>Band 4 and Band 8</td>
<td>$\frac{\text{NIR} - \text{RED}}{\text{NIR} + \text{RED}}$</td>
</tr>
<tr>
<td>NDMI</td>
<td>Landsat-8</td>
<td>Band 7 and Band 3</td>
<td>$\frac{\text{SWIR}2 - \text{GREEN}}{\text{SWIR}2 + \text{GREEN}}$</td>
</tr>
<tr>
<td></td>
<td>Sentinel-2</td>
<td>Band 12 and Band 3</td>
<td>$\frac{\text{SWIR}2 - \text{GREEN}}{\text{SWIR}2 + \text{GREEN}}$</td>
</tr>
<tr>
<td>MNDWI</td>
<td>Landsat-8</td>
<td>Band 3 and Band 3</td>
<td>$\frac{\text{SWIR}1 - \text{GREEN}}{\text{SWIR}1 + \text{GREEN}}$</td>
</tr>
<tr>
<td></td>
<td>Sentinel-2</td>
<td>Band 11 and Band 3</td>
<td>$\frac{\text{SWIR}1 - \text{GREEN}}{\text{SWIR}1 + \text{GREEN}}$</td>
</tr>
<tr>
<td>SR</td>
<td>Landsat-8</td>
<td>Band 5 and Band 4</td>
<td>$\frac{\text{NIR}}{\text{RED}}$</td>
</tr>
<tr>
<td></td>
<td>Sentinel-2</td>
<td>Band 8 and Band 4</td>
<td>$\frac{\text{NIR}}{\text{RED}}$</td>
</tr>
<tr>
<td>Band Ratio 54</td>
<td>Landsat-8</td>
<td>Band 6 and Band 5</td>
<td>$\frac{\text{SWIR}1 - \text{NIR}}{\text{SWIR}1 + \text{NIR}}$</td>
</tr>
<tr>
<td></td>
<td>Sentinel-2</td>
<td>Band 11 and Band 8</td>
<td>$\frac{\text{SWIR}1 - \text{NIR}}{\text{SWIR}1 + \text{NIR}}$</td>
</tr>
<tr>
<td>Band Ratio 35</td>
<td>Landsat-8</td>
<td>Band 4 and Band 6</td>
<td>$\frac{\text{RED}}{\text{SWIR}}$</td>
</tr>
<tr>
<td></td>
<td>Sentinel-2</td>
<td>Band 4 and Band 11</td>
<td>$\frac{\text{RED}}{\text{SWIR}}$</td>
</tr>
<tr>
<td>GCVI</td>
<td>Landsat-8</td>
<td>Band 5 and Band 3</td>
<td>$\frac{\text{GREEN}}{\text{RED}} - 1$</td>
</tr>
<tr>
<td></td>
<td>Sentinel-2</td>
<td>Band 8 and Band 3</td>
<td>$\frac{\text{GREEN}}{\text{RED}} - 1$</td>
</tr>
<tr>
<td>mRFDI</td>
<td>Sentinel-1</td>
<td>VV and VH</td>
<td>$\frac{\sigma_{VV} - \sigma_{VH}}{\sigma_{VV} + \sigma_{VH}}$</td>
</tr>
</tbody>
</table>

2.3. Existing Guyana Mangrove Extent Map Comparison

We compared our results with a 2018 mangrove extent map produced by Earth and Marine Environmental Consultants (EAME) acquired from a report by NAREI and CI. EAME used Sentinel-2 imagery to classify mangroves along Guyana’s coastline. Their results were used to verify the spatial distribution of mangrove forests within our study area.

Machine Learning Algorithm

Random forest is an ML algorithm [59,60] that builds large numbers of decision trees when training the sample data and returns a particular class depending on the maximum number of occurrences. This reduces the chance of overtraining the sample data [59,61,62]. The standard principles for RF training are selecting a random sample with a replacement of the training dataset and fitting trees to the samples by creating a split in the sample data. This provides the algorithm with matching responses for the training data [62]. Predictions are then established after training the sample by designating a particular classification of the trees [62].

Mangrove experts at NAREI generated training polygons for both mangrove and nonmangrove land cover types by visually inspecting Landsat-8 and Sentinel-2 imagery in GEE. Nonmangrove samples comprised coastal seagrasses, agriculture, water, grassland, barren ground, and settlement. We then sampled the spectral and backscatter information from the Landsat-8 composited image only, Sentinel-2 composited image only, Sentinel-1 composited image only, and combined composited images (data fusion). These samples were used for training the algorithm and assessing model performance in GEE. The stratified random sampling method was used for this research to reduce sampling bias [62,63]. When choosing a reference sample location, we considered the sample sizes for the two land cover types, where the larger land cover type (nonmangrove) required more training samples than the smaller land cover type (mangrove). The collected reference samples were divided into 60% for training the ML model algorithm and 40% for validation for each image at 30 m resolution (Table 3).
To determine the optimal number of trees with which to parameterize the random forest classification model for each dataset, hyperparameter tuning was conducted. Hyperparameter tuning can result in statistically significant improvements in supervised classification model performance [64]. In this case, every RF classifier was trained with a range of 5 to 200 decision trees in increments of 5 to determine the number of trees that leads to the greatest accuracy for the least amount of computational power. The accuracy of model performance for each iteration was graphed (Figure 4). Notably, Sentinel-1 (C) had a lower overall accuracy than other model iterations, and an adjusted y-axis range was used accordingly.

Table 3. Pixel samples from each data source.

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Mangrove</th>
<th>Nonmangrove</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landsat-8</td>
<td>546</td>
<td>5859</td>
<td>6405</td>
</tr>
<tr>
<td>Sentinel-2</td>
<td>556</td>
<td>6122</td>
<td>6678</td>
</tr>
<tr>
<td>Sentinel-1</td>
<td>557</td>
<td>6922</td>
<td>7479</td>
</tr>
<tr>
<td>Data Fusion</td>
<td>546</td>
<td>5648</td>
<td>6194</td>
</tr>
</tbody>
</table>

Figure 4. RF accuracy per number of decision trees for each dataset method.

For the validation, we used the F1-score (Equation (1)) as the metric for model performance. The F1-score is calculated using the precision and recall, also known as producer’s and user’s accuracy [65,66]. The precision is the ratio of correctly predicted positive observations to the total predicted positive observations. The recall represents the ratio of correctly predicted positive observations to all the observations in the class. The F1-score is the harmonic mean of precision and recall, considering both the false positives and false negatives.

\[ F1 - score = \frac{2 \times (\text{recall} \times \text{precision})}{\text{recall} + \text{precision}} \]  

(1)

2.4. Independent Accuracy Assessment and Unbiased Area Estimates

A systematic quantitative accuracy assessment is necessary to measure a map’s accuracy and estimate each class’s area based on reference sample data. We used methods described by Olofsson et al. (2014) [67] to extract the sample-based producer’s accuracy, user’s accuracy, overall accuracy (Equations (2)–(4)), and unbiased area estimates (±95% confidence interval).
Producer’s accuracy = \frac{\sum \text{correctly classified reference data for a class}}{\text{Total reference points for that class}} \quad (2)

User’s accuracy = \frac{\sum \text{correctly classified points for a class}}{\text{Total classified points for that class}} \quad (3)

Overall accuracy = \frac{\sum \text{correctly classified points for all classes}}{\text{Total points}} \quad (4)

A total of 401 reference sample points within the study area were generated in GEE using a simple random sampling design \[67\]. The points were classified into mangrove and nonmangrove classes using the freely available Norway’s International Climate and Forests Initiative (NICFI) Planet high-resolution (4.77 m spatial resolution) base maps. All monthly base maps for the year of classification (2020) were used in image interpretation. This process was completed using a visual interpretation sampling approach \[68,69\] in Collect Earth Online (CEO), a free and open-source interpretation system connected with high-resolution satellite image display functionality. CEO was created as a collaborative platform by the Food and Agriculture Organization (FAO) and SERVIR \[70\]. Image interpretation was conducted by remote sensing analysts at NAREI and the SERVIR Science Coordination Office, who have combined expertise in Guyana’s mangroves and in interpreting remotely sensed imagery. Reference points in which the visual interpreter had low confidence were flagged and reviewed by multiple analysts for final classification.

3. Results

3.1. Mangrove Mapping Extent Results

The RF machine learning algorithm produced 2020 mangrove extent maps at 30 m for the study area using Landsat-8, Sentinel-2, Sentinel-1, and data fusion. Smaller mangrove stands within riverine environments and extensively fragmented patches can be detected and mapped in considerable detail using these high-resolution data and RF ML algorithm. Based on the unbiased area estimates, Landsat-8 only produced the largest mangrove extent, with a total of 3645.0 ± 836.7 hectares, accounting for 8.5% of the study area (Table 4). On the other hand, Sentinel-1 data produced the lowest mangrove extent with a total of 908.4 ± 579.7 hectares, accounting for 2.1% of the study area (Table 4).

Table 4. Comparison of total mangrove coverage (ha) and percentage coverage for optical and radar images at 30 m resolution, Landsat-8, Sentinel-2, Sentinel-1, and data fusion.

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Nonmangrove Coverage (ha)</th>
<th>Mangrove Coverage (ha)</th>
<th>95% Confidence Interval of Area (ha)</th>
<th>Mangrove Percentage Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landsat-8</td>
<td>39,360.8</td>
<td>3645.0</td>
<td>±836.7</td>
<td>8.5%</td>
</tr>
<tr>
<td>Sentinel-2</td>
<td>39,734.0</td>
<td>3271.8</td>
<td>±712.0</td>
<td>7.6%</td>
</tr>
<tr>
<td>Sentinel-1</td>
<td>42,097.3</td>
<td>908.4</td>
<td>±579.7</td>
<td>2.1%</td>
</tr>
<tr>
<td>Data fusion</td>
<td>39,817.0</td>
<td>3188.8</td>
<td>±707.0</td>
<td>7.4%</td>
</tr>
</tbody>
</table>

3.2. Mangrove Map Extent Comparison

The resulting mangrove extent map outputs were compared with the EAME 2018 mangrove baseline that was generated using only Sentinel-2 data. The visual comparisons demonstrated how closely our mangrove extent maps matched the 2018 baseline (Figure 5). The Landsat-8, Sentinel-2, and data fusion outputs were consistent with the 2018 baseline.
Figure 5. Mangrove extent map comparison for the study area based on Landsat-8, Sentinel-2, Sentinel-1 data, and data fusion. (Larger) Mangrove extent agreement between all datasets (Smaller); separated mangrove extents.

The EAME 2018 mangrove baseline displayed the spatial extent of mangroves at the estuary of the Pomeroon River, along both banks of the river, and along both sides of the Atlantic coast. The baseline also revealed that mangroves were expanding further inland of the Pomeroon estuary. While the distribution of mangroves between Landsat-8 and Sentinel-2 mangrove extents was similar to that of the 2018 baseline, Landsat-8 detected more mangroves in the Pomeroon riverine area and inland patches than Sentinel-2 (Figure 5). Sentinel-1’s mangrove extent revealed sparse mangrove distribution both within mangrove areas highlighted by Landsat-8, Sentinel-2, data fusion, and baseline data, and within areas where mangroves were not highlighted by the other datasets and where mangroves may not be present with certainty (Figure 5). The mangrove extent produced using data fusion revealed a general distribution that was consistent with Landsat-8, Sentinel-2, and baseline data. Data fusion, Landsat-8, and Sentinel-2 extents intersected well, highlighting the areas where mangroves were likely to be found and coincided with the baseline data.

3.3. Accuracy

3.3.1. Model Performance

Using the 40% reference points dedicated to assessing the model performance (see Section 2.4), F1-scores were calculated based on confusion matrices of point counts for the four different products (Landsat-8 only, Sentinel-2 only, Sentinel-1 only, and data fusion). The results are summarized in Table 5.

<table>
<thead>
<tr>
<th>F1-Score (%)</th>
<th>Nonmangrove</th>
<th>Mangrove</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landsat-8</td>
<td>99.5</td>
<td>93.5</td>
</tr>
<tr>
<td>Sentinel-2</td>
<td>99.5</td>
<td>94.2</td>
</tr>
<tr>
<td>Sentinel-1</td>
<td>96.0</td>
<td>28.6</td>
</tr>
<tr>
<td>Data Fusion</td>
<td>99.6</td>
<td>95.8</td>
</tr>
</tbody>
</table>

Model performance varied, with F1-scores ranging between 96.0% and 99.6% for the nonmangrove class and 28.6% and 95.8% for the mangrove class. The Sentinel-1 only product had the lowest performance for both mangrove and nonmangrove classes, while the optical and data fusion outputs performed better for the mangrove class. The data fusion approach showed the highest F1-score for both mangrove (95.8%) and nonmangrove
(99.6%), with the latter being a 0.1% difference compared with the Sentinel-2 only and Landsat-8 only outputs (99.5%). However, model performance results based exclusively on the reference data do not guarantee accurate maps. Additional results from an independent accuracy assessment are needed and summarized in the next section.

3.3.2. Independent Accuracy Assessment

The independent accuracy assessment we conducted compares the mangrove extent results with a random sample of points collected over monthly Planet base maps from 2020. Overall accuracy varied from 88% (Sentinel-1 only) to 95% (data fusion) (Tables 6–9). User’s accuracy for the mangrove classification varied greatly (13% to 69%), as did producer’s accuracy (50% to 87%). The poorest performance across these metrics was from the Sentinel-1 only classification. Performances across the Landsat-8 only, Sentinel-2 only, and data fusion approaches were more consistent, with the data fusion slightly outperforming the other model iterations in overall, user’s, and producer’s accuracy (Tables 6, 7 and 9). Tables 6–9 also report biased area totals ($A_m$) and class proportion ($W_h$).

### Table 6. Landsat-8 only.

| Pixel Count | Predicted | Total | $A_m$ (ha) | $W_h$
|-------------|-----------|-------|------------|------|
| Observed Nonmangrove | 344 | 9 | 353 | 39,030.9 | 91%
| Mangrove | 16 | 32 | 48 | 3974.8 | 9%
| Total | 360 | 41 | 401 | 43,005.8 | 100%
| OA: 94% M Users A: 67% | M Producers A: 78% | NM Users A: 97% | NM Producers A: 96% |  |

### Table 7. Sentinel-2 only.

| Pixel Count | Predicted | Total | $A_m$ (ha) | $W_h$
|-------------|-----------|-------|------------|------|
| Observed Nonmangrove | 348 | 5 | 353 | 39,051.4 | 91%
| Mangrove | 15 | 33 | 48 | 3954.4 | 9%
| Total | 363 | 38 | 401 | 43,005.8 | 100%
| OA: 95% M Users A: 69% | M Producers A: 87% | NM Users A: 99% | NM Producers A: 97% |  |

### Table 8. Sentinel-1 only.

| Pixel Count | Predicted | Total | $A_m$ (ha) | $W_h$
|-------------|-----------|-------|------------|------|
| Observed Nonmangrove | 347 | 6 | 353 | 41,362.7 | 96%
| Mangrove | 42 | 6 | 48 | 1643.1 | 4%
| Total | 389 | 12 | 401 | 43,005.8 | 100%
| OA: 88% M Users A: 13% | M Producers A: 50% | NM Users A: 98% | NM Producers A: 89% |  |
Table 9. Data fusion.

<table>
<thead>
<tr>
<th>Pixel Count</th>
<th>Predicted</th>
<th>Total</th>
<th>( A_m ) (ha)</th>
<th>( W_h )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed</td>
<td>Nonmangrove (NM)</td>
<td>348</td>
<td>353</td>
<td>39,174.6</td>
</tr>
<tr>
<td></td>
<td>Mangrove (M)</td>
<td>15</td>
<td>33</td>
<td>3831.1</td>
</tr>
<tr>
<td>Total</td>
<td>Nonmangrove (NM)</td>
<td>363</td>
<td>38</td>
<td>43,005.8</td>
</tr>
<tr>
<td></td>
<td>Mangrove (M)</td>
<td>38</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>OA: 95%</td>
<td>M Users A: 69%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M Producers A: 87%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NM Users A:</td>
<td>NM Users A: 99%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NM Producers A: 96%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.3.3. Field Verification Exercise

The ground truth fieldwork activity featured three targeted mangrove sites for cross validation, which NAREI conducted from 30 September to 1 October 2021. These sites were identified based on observation of visible changes between the 2018 and 2020 extent maps. In order to visualize the land cover and land use within the areas, the chosen target sites were studied using high-resolution still photography taken from a DJI Phantom 4 Advanced drone. The drone was first stationed in the Pomeroon River close to the estuary (target site 1), where agriculture was the predominant use of the surrounding land, along with dispersed settlements (Figure 6A). At the second target site of located La Union, mangroves were not visible in the photographed foreshore, but other tiny shrubs and seagrasses were observed in that location (Figure 6B). Target site 3 was located at Supenaam Stelling, where communities were observed more inland from the river bank to the north and south of the drone station (Figure 6C).

Figure 6 illustrates 2018 and 2020 map extents, which highlighted notable changes within target areas. For target site 1, the 2020 mangrove extent showed significant mangrove losses as a direct result of agricultural expansion in the area compared with the 2018 mangrove extent. Figure 7A,B show a delineation of a small fringes of mangroves along the river banks and agricultural lands immediately adjacent to mangroves. Furthermore, evidence of mangrove displacement can be observed at the boundary between the mangrove fringe and agricultural lands Figure 7E. These changes were detected using our mangrove
mapping method. At target site 2, the 2018 mangrove extent showed mangrove fringe located along the La Union Village foreshore; however, the 2020 mangrove extent and Figure 7C,D showed absence of mangrove. Nonetheless, instead of mangroves, seagrasses can be observed at the site. At target site 3, the 2018 mangrove extent showed mangrove inland; however, our method showed mangrove fringes on the river banks, with clustered settlements located in the proximity of mangroves (Figure 7E,F). Furthermore, the 2018 extent showed other vegetation species detected inland, while our model detected mangrove fringe, which was verified by the photographs.

Figure 7. (A–F) Land cover type within target areas depicted in Figure 6 captured using drone photography.

4. Discussion
4.1. Key Findings

Overall, the classification based on a combination of optical and radar data resulted in higher user’s, producer’s, and overall accuracies than models trained with a single dataset. Nevertheless, the performance of the model using only optical data was strong, likely due in part to the sufficient data coverage of Region 2 in 2020 from Landsat-8 and Sentinel-2. The Sentinel-1 product did not perform well alone, but results were enhanced when it was combined with the optical datasets.

When the map outputs from Sentinel-2, Landsat-8, and the data fusion in our study were compared with existing ones, results indicated that the classification of mangroves in the EAME 2018 output identified mangroves where they were not likely to be found. Ground truth activities corroborated this finding, since large swaths of mangroves were not observed along the banks of the Supenaam River; rather, very sparse fringes of mangroves

Figure 7. (A–F) Land cover type within target areas depicted in Figure 6 captured using drone photography.

4. Discussion
4.1. Key Findings

Overall, the classification based on a combination of optical and radar data resulted in higher user’s, producer’s, and overall accuracies than models trained with a single dataset. Nevertheless, the performance of the model using only optical data was strong, likely due in part to the sufficient data coverage of Region 2 in 2020 from Landsat-8 and Sentinel-2. The Sentinel-1 product did not perform well alone, but results were enhanced when it was combined with the optical datasets.

When the map outputs from Sentinel-2, Landsat-8, and the data fusion in our study were compared with existing ones, results indicated that the classification of mangroves in the EAME 2018 output identified mangroves where they were not likely to be found. Ground truth activities corroborated this finding, since large swaths of mangroves were not observed along the banks of the Supenaam River; rather, very sparse fringes of mangroves
were observed along the eastern bank of the Supenaam River and along the west bank of Essequibo. Nonetheless, the EAME 2018 extent managed to clearly detect patches of mangroves within the Pomeroon Riverine environment.

Moreover, a comparison between the 2018 and 2020 map products detected significant changes, particularly mangrove losses at the Pomeroon estuary and upriver (Figure 6). Results from ground truth activities showed significant mangrove forest loss due to the large expansion of coconut farms northeast of the estuary and the large clearing of mangroves for construction purposes (Figure 6).

We also compared the accuracies of our models with those of similar mangrove extent studies. We assessed GMW [21] in GEE and found that their overall, producer’s, and user’s accuracies were 91.8%, 70.8%, and 64.2%, respectively, for Region 2 using Advanced Land Observing Satellite (ALOS) Phased Array Type L-Band Synthetic Aperture Radar (PALSAR) L-Band and Landsat data, which are lower than the accuracies of our data fusion approach. Additionally, the GMW 2020 data misclassify a clearly deforested area of mangroves (see Figure 6A). We also assessed the European Space Agency’s WorldCover Mangrove Layer [26] for 2020 in Region 2, and their accuracies were 89.0%, 79.2%, and 52.8%, respectively. This product is also produced with optical and SAR data, specifically from Sentinel-1 and 2. Our data fusion outperformed this product as well. Lastly, we assessed the mangrove extent in the MapBiomas Amazon 2020 product from Collection 5 [71] for Region 2 and found that the accuracies were 95.0%, 58.8%, and 71.4%, respectively. The overall accuracy is the same, and the producer’s and user’s accuracies are lower than our data fusion results. It should be noted that the MapBiomas product does not cover the entire Region 2 coastline that we assessed in our study.

4.2. Limitations and Future Work

While radar data have been an adequate data source for mangrove classification in other studies [29], our results demonstrate poor performance using only Sentinel-1 for classification relative to optical data. The poor performance may be due to the nature of C-band data, which may not be able to penetrate the mangrove forest canopy. A study by Simard et al. (2002) [72] successfully used L-band in the HH polarization to capitalize on the double-bounce backscatter mechanism and clearly discriminate flooded forests from other forest classes. In order to have a clearer double-bounce signal, L-band is required in order to penetrate vegetation and reach the ground [73]. If most of the C-band signal scattering is caused by canopy volume scattering rather than double-bounce scattering from interactions between tree trunks and the ground, the signal may not be as discernible [73]. In this case, the exclusive use of C-band data was likely a limiting factor in the effectiveness of SAR to map the mangrove extent [74]. Future work would involve employing free, open, and regularly available L-band data in this analysis, potentially from the new NISAR mission by NASA and the Indian Space Research Organisation set to launch in early 2024 [75].

There is potential for the future estimation of carbon stock in mangrove forests based on the maps produced by these methods. Coastal forests are one of the richest carbon sinks globally, sequestering more carbon than most inland rainforests. Mangroves are considered a blue carbon ecosystem due to their accumulation of rich soils and ability to store organic carbon both internally and externally [76]. Their complex root system traps sediment buildup, sequestering the carbon over a long period of time, while their dense, leafy canopy sequesters over a shorter period of time but throughout the living biomass of the forest [77].

Carbon reporting is a crucial component of the United Nations Framework Convention on Climate Change framework that guides country participants seeking to reduce deforestation and forest degradation emissions. These reports are frequently based on land cover datasets produced by the country or on global datasets produced with remotely sensed data. SERVIR’s Science Coordination Office has begun implementing the SERVIR Carbon Pilot Project (S-CAP), which derives carbon emission estimates from a wide array of land cover extent datasets and compares them to better support countries in selecting
the most efficient method for their applications. The S-CAP team plans to incorporate their methodologies into this mangrove monitoring service to better quantify the emission and sequestration of Guyana’s mangroves. The S-CAP approach would employ the mangrove change extents generated through this study and apply various biomass datasets to estimate the emissions of cleared mangroves. The integration of these methods will assist NAREI in monitoring and understanding the climate impacts of these changes.

5. Conclusions

This study used Landsat-8, Sentinel-2, Sentinel-1, and a combination of those datasets to produce high-resolution mangrove forest data products for Pomeroon-Supenaam (Region 2) based on the RF ML algorithm in GEE. The resulting mangrove extent maps revealed that Sentinel-2, Landsat-8, and the data fusion product produced higher accuracies that were consistent with the general spatial distribution of mangrove forests in the EAME 2018 mangrove forest baseline, although smaller mangrove stands were additionally detected in the riverine environment. On the other hand, Sentinel-1 showed sparse mangrove forest distribution that was inconsistent with Landsat-8, Sentinel-2, the fusion dataset, and the EAME 2018 baseline.

Based on these results, a random forest model applied in a cloud computing platform like GEE is a suitable approach for mapping the mangrove extent along Guyana’s coast. Time series analytics can require considerable computing power; therefore, a cloud-based approach is necessary, especially in regions with intermittent internet connection. Moreover, GEE provides quick access to free and openly available datasets across multiple sensors and data providers, facilitating these analyses. In terms of the computing power needed to run these models, we estimate the use of 48–53 EECU hours (Earth Engine Compute Unit) or USD 24–31 to replicate this same analysis to another year of study, proving that operationalizing this approach in a cloud-based environment is feasible and cost-effective.

This analysis was needed to identify the most appropriate technique to develop a mangrove monitoring service for Guyana that provides updated annual data on Guyana’s mangrove extent, gains, losses, and the impact of NAREI’s restoration efforts. It was also necessary for the data produced by the analysis to be freely accessible to a number of agencies, including the Sea and River Defence Division, Guyana Forestry Commission, Conservation International, World Wildlife Fund, and other stakeholders.

This paper provides the first comparison of mangrove extent mapping approaches in Guyana using remotely sensed data. We evaluate the performance of an ML algorithm to classify three different data combinations to determine the best approach for mapping the annual mangrove extent in a tropical, cloud-persistent, coastal ecosystem. Using only free and open data, this method empowers NAREI to produce mangrove extent datasets to their own specifications with the confidence that it will perform on par with other mangrove monitoring efforts and meet their mandate to monitor and protect this critical and vibrant ecosystem.


Funding: Funding for this work was provided through the cooperative agreement 80MSFC22N0004 between NASA and UAH. SERVIR is a joint NASA- and USAID-led program.

Data Availability Statement: The Landsat-8, Sentinel-2, and Sentinel-1 imageries used in this study are open and freely available in Google Earth Engine. The NICFI monthly base maps are available for noncommercial use through Planet and on Google Earth Engine. The training and testing datasets employed are available in the linked Google Earth Engine scripts.
Acknowledgments: Special thanks to Marc Simard, Abigail Barenblitt, Lola Fatoyinbo, and NAREI.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations
The following abbreviations are used in this manuscript:

- ALOS: Advanced Land Observing Satellite
- EECU: Earth Engine Compute Unit
- FAO: Food and Agriculture Organization
- GCVI: Green Chlorophyll Vegetation Index
- GEE: Google Earth Engine
- GMW: Global Mangrove Watch
- GRD: Ground Range Detected
- LULC: Land Use and Land Cover
- ML: Machine Learning
- MNDWI: Modified Normalized Difference Water Index
- mRFDI: Modified Radar Forest Degradation Index
- MSI: MultiSpectral Instrument
- NAREI: National Agricultural and Research Extension Institute
- NASA: National Aeronautics and Space Administration
- NDMI: Normalized Difference Moisture Index
- NDVI: Normalized Difference Vegetation Index
- NICFI: Norway’s International Climate and Forests Initiative
- OLI: Operational Land Imager
- PALSAR: Phased Array Type L-Band Synthetic Aperture Radar
- RF: Random Forest
- RTC: Radiometric Terrain Corrected
- SAR: Synthetic Aperture Radar
- S-CAP: SERVIR Carbon Pilot
- SDG: Sustainable Development Goals
- SEDAC: NASA Socioeconomic Data and Applications
- SR: Simple Ratio
- SRTM: Shuttle Radar Topography Mission
- TM: Thematic Mapper

References


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.