Ann. For. Res. 66(1): 3114-3142, 2023
ISSN: 18448135, 20652445

ANNALS OF FOREST RESEARCH
www.e-afr.org

CLASSIFICATION AND ABOVE GROUND BIOMASS MAPPING OF INDIAN LANDLOCKED MANGROVE FOREST THROUGH SAR DATA

Archana Singh[1,2*] and Seema Mahajan[2,1]

[1*]Computer Engineering, Indus University, Ahmedabad, Gujarat, India.
[2]Computer Engineering, Indus University, Ahmedabad, Gujarat, India.

*Corresponding author(s). E-mail(s): archanasingh.rs@indusuni.ac.in;
Contributing authors: ce.hod@indusuni.ac.in;

Abstract

One of the most productive ecosystems, mangroves have numerous ecological advantages. They are the largest carbon absorbers, which means they are vital to preserving the balance of the environment. In order to safeguard, conserve, and plan for the replanting of these priceless natural resources, reliable thematic maps of mangrove ecosystems and models for mangrove above-ground biomass (AGB) assessment are essential. In this study, the Google Earth Engine was used to create a thorough mangrove ecosystem map and AGB model for the Guneri mangrove forest in the Kutch area of Gujarat, India, using Sentinel-1 A satellite images in conjunction. This research emphasizes the capacities of Sentinel-1A, SAR images for mangrove characteristics mapping, and the model generated is verified using ground truth data collected from the field survey of 127 sample points. For a more precise model, synthetic data is generated using Generative Adversarial Networks (GANs), and the min-max normalization technique is employed to normalize the data. Mangrove vegetation is mapped using a pixel-based random forest (RF) classifier, with an average overall classification accuracy of 91% and RMSE of 0.506. For AGB model generation the machine learning techniques applied to the dataset are Extra Trees Regressor, XGB Regressor, Random Forest Regressor, Bagging Regressor, and Decision

Tree Regressor. Comparatively, it is found that Extra Trees Regressor demonstrated a good validation accuracy of 66% with 0.10 RMSE. This work validates the applicability of Random Forest (RF) and Extra Trees Regressor algorithms for mapping and estimating AGB for a unique landlocked mangrove site of Guneri, and it is observed that the results and robustness of the model are highly affected by the usage of a larger dataset and the geographical parameters of the study site.

Keywords: AGB, SAR, remote sensing, Machine learning, synthetic data , Normalization,

Introduction

One of the most biodiverse habitats along tropical seacoasts and bays is mangrove forests, which are made up of salt-tolerant plants with aerial breathing roots that act as sediment traps and support a variety of marine creatures [1, 2]. Mangroves serve as barriers to stop hurricanes, tsunamis, and ocean waves from destroying the coastal area. Additionally, mangroves act as a natural filter to enhance the quality of the water and serve as a habitat for a variety of aquatic life forms. Mangroves are substantial carbon sinks in coastal environments, and they perform key functions in this regard. Mangrove serves to control coastal floods and erosion and safeguards inland farms, ranches, and other settlements that are close to the coast from storms like cyclones and hurricanes [3, 4]. Additionally, mangroves directly support the economics and way of life of coastal populations by supplying honey, fuel, traditional medicines, and suitable sites for aquaculture and fisheries. The fact that the ocean and coastal regions only contain 0.05% of the total plant biomass on land, but may nonetheless absorb a similar amount of carbon annually, is an interesting one. Studies show that mangroves have better primary yield than other types of forests. Mangroves have some of the greatest levels of biomass carbon in the tropics. When compared to other tropical forests throughout the world, mangroves can store up to four times more carbon [5].
The mapping of the world's carbon stores is becoming increasingly accurate, and fluxes have accelerated significantly in recent years. However, mangroves have mostly been disregarded in these evaluations due to their modest area and difficult conditions [6]. It has long been acknowledged that most of the regions with a diverse range of mangroves are either inaccessible or logistically challenging to research in the field and that this takes a lot of time [7–10]. As dwarf mangroves' aboveground biomass (AGB) may be as little as 8 t/ha to as high as >500 t/ha in the Indo-Pacific region's riverine and fringe mangroves. One of the main markers of a forest's carbon content is its AGB, which is simple to measure in the field but necessitates the removal of trees, hence a different approach is needed [7]. Consequently, there was a need for a better, more efficient, and quicker way to investigate mangrove ecosystems. In India, there have been numerous attempts to map and classify mangroves. The use of remote sensing data is advantageous because it enables change detection studies to be conducted much more easily than field-based estimates. This is because remote sensing data can provide quick synoptic coverage with a high temporal resolution [11–14]. Monitoring and mapping the changes that have taken place in mangroves are absolutely necessary in order to gain a better understanding of how mangroves react to changes in climate as well as changes that have been caused by humans. Estimating mangrove AGB is necessary for planning the preservation and sustainable use of mangrove resources. Remote sensing data is useful for mapping and monitoring vegetation, land cover, and land-use change, it remains a key source of information because it is challenging to access dense mangrove areas. Radar sensors allegedly have the ability to identify the volumetric properties of dense foliage [15–17]. Due to polarisation, sensitivity to moisture (dielectric constant), surface roughness, varying incident angles, and strong penetration capabilities, SAR has the potential to be used for forest investigations [18].

Synthetic Aperture Radar (SAR) has been used in several biomass estimation studies for forest vegetation. The purpose of this study was to develop a robust model for distinguishing mangroves from non-mangroves and a model for mangrove above-ground biomass estimation using spectral signatures (produced by SAR remote sensing) and morphological characteristics of mangroves. This model's foundation is built on the mangrove vegetation, which was measured for height, width, latitude, and longitude. The model prepared here takes the sigma values generated by the SAR sensors and the ground truth values as input to distinguish mangroves and non-mangroves regions and then estimates the AGB for mangroves, with the help of machine learning algorithms and also validates allometric models for mangrove above ground biomass estimation The research was conducted in the village of Guneri's mangrove forest. Due to their landlocked location, these mangroves are a unique variety. And because these mangroves are being investigated for the first time for this kind of research, this study is unique.

Remote Sensing: Mangrove classification and Mangrove AGB estimation

For the management and monitoring of mangroves, remote sensing has become essential. The majority of the existing mangrove mapping research relies on optical remote sensing data, the most common of which are medium-resolution multispectral pictures like those from Landsat [19–22]. The accuracy of mangrove extraction will be impacted by the existence of mixed pixels, which is determined to be inevitable by the medium spatial resolution. Mangrove extraction uses high-resolution remote sensing data increasingly, including UAV data [23, 24] and satellites like IKONOS [25], ZY-3 [26], and GF-1 [27]. Data from hyperspectral sensors like the Hyperion are anticipated to show the potential for accurate and thorough mangrove forest classification. Accurate tropical mangrove species discrimination may be one step closer with the use of hyperspectral technologies. Many uses in the structure, composition, and physiology of plants have already shown the promise of hyperspectral imaging and image processing. The benefit is mostly attributable to its capacity to measure reflectance and absorption in a significant number of continuous and narrow spectral bands. The 400 to 2500 nm region of the electromagnetic spectrum is covered by airborne and satellite hyperspectral data. As a result, it is now much easier to characterize and describe the full spectrum of the many floristic classes that make up a mangrove forest. Henceforth, hyperspectral data may enhance our capacity to distinguish mangroves from other terrestrial forests and then identify these at the species level [28–30]. Coastal wetlands, particularly in tropical and subtropical areas, have a lot of overcast and foggy days as well as complicated land cover types. Optically available images are highly limited. The advantages of SAR satellite imaging include excellent resolution and extensive coverage. With the exception of extreme weather, it is not constrained by climatic or meteorological circumstances and can reliably collect periodic data. The unique imaging technique of SAR results in issues such as inherent speckle noise, shadow, and overlap in SAR images, as well as a low signal-to-noise ratio. Because of this, mangrove recognition and extraction using SAR data alone are considered less accurate by researchers across the world, but SAR photos are still a helpful addition to optical imaging. Numerous studies have demonstrated the utility of longer wavelength L-band SAR data for estimating AGB in tropical forests. Nevertheless, since L-band data from the Advanced Land Observing Satellite (ALOS) is not publicly available, it is challenging to estimate and frequently monitor the AGB of a vast area using solely L-band SAR data. C-band SAR interferometric coherence has shown great promise in estimating biophysical parameters like growing stock volume while using semi-empirical models, such as the Water Cloud Model (WCM) and the Interferometric WCM, even though C-band SAR backscatter is not suitable for AGB estimation of tropical forests. This research is conducted with two objectives: mangrove discrimination and mangrove AGB estimation using SAR. Researchers are working in this direction for years, yet there is no general way to classify mangroves and estimate AGB, (Table 4 and Table 2) are summaries of the most recent work done in this direction.

Mangrove dynamic monitoring and management depend heavily on accurate mangrove mapping. Optical remote sensing images are the primary data source for the generation of mangrove indices and mangrove mapping, and numerous studies have shown positive findings in this area. [31] The study created the OSCMI (Optical and SAR Images Combined Mangrove Index) based on the concept of multi-feature fusion. For efficient and accurate mangrove mapping, an OSCMI-based classification scheme was put forth. Using Sentinel1 SAR and Sentinel-2 optical images, extraction experiments were carried out in four different types of typical Chinese mangrove environments. In [32], using the Google Earth Engine (GEE) cloud computing platform, Sentinel-1 and Sentinel-2 satellite pictures were combined to create a precise mangrove

Table 1 Mangrove discrimination from other vegetation using Sentinel imaging and machine learning

Data	Region/ Forest Type	Method	Model Performance	Reference
Sentinel 1 and Sentinel 2	4 mangrove forests in China (2022)	OSCMI	92%-96%	42
Sentinel 1 and Sentinel 2	Qeshm, Iran (2021)	Pixel based RF	93%	43
LANDSAT 8 OLI	Sundarban and Andaman, India (2018)	CMRI	73.43%	44
EO-1 Hyperion	Sundarban, India (2017)	SVM	98%	45

ecosystem map of the Hara protected area, Qeshm, Iran, with a 10 m spatial resolution. In this context, seasonal optical and synthetic aperture radar (SAR) characteristics were created using 86 Sentinel-1 and 41 Sentinel-2 data that were gathered in 2019. Following that, seasonal features were added to a pixel-based random forest (RF) classifier. This produced an accurate mangrove ecosystem map with average overall accuracy (OA) and Kappa coefficient (KC) of 93.23% and 0.92, respectively. Mangrove species are difficult to distinguish from non-mangrove flora, particularly in locations where they coexist with other plant species. In [33], an attempt is made to create a better index using data from the Bhitarkanika Mangrove Forest in Odisha, India's Normalized Difference Vegetation Index (NDVI) and Normalized Difference Water Index (NDWI). The correlation between these indices is adverse (r = -0.988; p 0.01). Additionally, at the pixel level, the NDWI values were subtracted from the NDVI data. Due to the negative correlation between the outputs, subtracting them broadens the upper and lower bounds of the overall output as well as the distinct values of two classes with closely related spectral signatures. On the Sundarbans and Andaman mangroves (r = -0.987 and -0.989, respectively), the CMRI algorithm was used. Terrestrial vegetation has a significant impact on the mangroves' ability to be accurately identified in remote sensing. The paper [34] discusses the classification of mangroves into floristic composition classes and the application of specific vegetation indices for extracting mangrove forests using Earth Observing-1 Hyperion images over a stretch of the Indian Sundarbans. Both of these topics are covered in greater detail. The decision tree algorithm relied on the following five vegetation indices in order to generate the mangrove mask: the Mangrove Probability Vegetation Index; the Normalized Difference Wetland Vegetation Index; the Shortwave Infrared Absorption Index; the Normalized Difference Infrared Index; and the Atmospherically Corrected Vegetation Index. Then, using the information included in the mask, three full-pixel classifiers, Minimum Distance, Spectral Angle Mapper, and Support Vector Machine (SVM) were assessed. SVM outperformed the other two, with a 99.08% total precision. The research shows that Optical images have a better capability to classify mangroves.

Table2 Mangroveabovegroundbiomass(AGB)estimationusingSentinelimagingandmachinelearning

DataRegion/ForestTypeMethodModelPerformanceReference				
Sentinel1& Sentinel2	BhitarkanikaWildlifeSanctuary (BWS inEasternIndia(2021)	XGB,GBM,RFR20.73-0.7546		
Sentinel-1A	Kutch,Gujarat, India(2021)	PRVIR20.5647		
ALOS-2& PALSAR-2	EastKalimantan, Indonesia(2020)	LLOCVR20.8948		
RISAT-1& ALOS-PALSAR	GujaratStateVegetation, India(2017)	semi-empiricalmodelbased onmulti-linearregression	C-bandSAR(RISAT-1) Trop.moistdeciduousR20.52 Trop.drydeciduousR20.71 NortherntropicalthornR20.79 L-bandSAR(ALOS-PALSAR) Trop.moistdeciduousR20.78 Trop.drydeciduousR20.76 NortherntropicalthornR20.59	49

In recent years, the process of estimating forest AGB has undergone a revolution due to the free availability of high geographic and temporal resolution Sentinel series data and a variety of machine learning algorithms. A study [35] was carried out on the eastern Indian coast, in the Bhitarkanika Wildlife Sanctuary. The AGB maps that are generated by the Deep Learning models and the Interferometric Water Cloud Model (IWCM) are distinct from one another due to the fact that they are dependent on different sets of factors. When it came to coherence, the ground and vegetation components were particularly important to IWCM, whereas for the Deep Learning model, canopy height was the most important factor. However, the insignificance of the changes brought about by Deep Learning-based AGB maps may be attributable to a lack of comprehension regarding the significance of coherence and VH backscatter. Due to the low canopy penetration power of C-band SAR, strong temporal decorrelation caused by a large time gap between interferometric picture pairs, and the great spatial variability of mangrove forests, it was found that IWCM was an inadequate method for AGB estimation. It is interesting to note that in the Bhitarkanika Wildlife Sanctuary, a Deep Learning algorithm was able to interpret the precise link between predictor factors and mangrove AGB. A greater emphasis should be placed, rather than employing semi-empirical models in AGB estimate studies in mangrove forests using Sentinel data, on applying machine learning algorithms such as Deep Learning. The use of remote sensing technologies for the assessment of has increased in recent years due to the difficulties associated with accessibility within the forests (AGB). For the mangroves of Mundra taluka in Kachchh district, western India, [36] uses an allometric model to offer a novel method of estimating AGB from dual-pol Sentinel-1A (SAR) data. Polarimetric Radar Vegetation Index (PRVI) is used to simulate AGB . The work incorporates AGB time series change detection within the study region and argues that combining field measurements, allometric equations, and remote sensing technology may be the key to modelling the AGB of mangrove forests. One of the available methods, L-band Synthetic Aperture Radar (SAR), can reliably estimate AGB in terrestrial tropical forests. Mangrove ecosystems, however, typically have poor accuracy. [37] A study was done to model and map AGB using backscatter coefficients from the Advanced Land Observing Satellite-2 (ALOS-2) Phased Array L-band SAR-2 in a section of the restored mangrove forest at Mahakam Delta, Indonesia (PALSAR-2). The model's performance was assessed using three different methods: independent validation (R2 of 0.89 and RMSE of 23.16 tonnes ha1), random k-fold cross validation (R2 of 0.89 and RMSE of 24.59 tonnes ha1), and leave location out cross-validation (LLO CV) (R2 of 0.88 and RMSE of 24.05 tonnes ha1). [38] a semi-empirical model based on multi-linear regression coefficients of HH and HV polarisation backscatter with field-measured forest biomass was used to retrieve temporal forest AGB of Gujarat, India, from global SAR mosaic products in HH/HV polarisations produced from Japanese ALOS-PALSAR 1/2 data. This model was used to retrieve temporal forest AGB of Gujarat, India, from global SAR mosaic products in HH/HV polarisations. The primary types of forests that can be found in Gujarat are the tropical moist deciduous forest, the tropical dry deciduous forest, the northern tropical thorn forest, and the littoral and swamp forest. In order to produce Gujarat's biomass maps, various model coefficients had to first be established for the various types of Gujarat's forests. These were done based on in-depth ground measurements of the forest's parameters. High correlations between HV and HH/HV and field-measured biomass were found across various forest vegetation types, with biomass densities ranging from 20 to 120 t/ha. According to studies on several Indian mangrove species, AGB ranges from 20.9 t/ha in the Sundarbans region to 196.48 t/ha and 236 t/ha for mangroves on the Kerala coast, respectively. Since machine learning techniques have never been utilized for AGB estimation of mangroves, their results cannot be directly compared to those of semi-empirical models. A study of earlier studies reveals that semi-empirical and machine learning models are rarely used in AGB Estimation in the setting of Indian tropical forests [39– 41]. This research is conducted on the Landlocked mangrove forest of Gunneri (one of its own kind) for the first time, with the objective of discriminating the mangroves from other vegetation and estimating the AGB of classified mangroves. Mangrove is classified with a pixel-based Random Forest algorithm and AGB estimation model is formed using a machine learning algorithm. Several experiments were performed in order to map above-ground biomass (AGB) of mangroves, and it is observed that Extra Tree Regression is the most suited model.

Materials and Methods

Study Area

The Kuchchh district in Gujarat is home to one of the most extensive mangrove forests in India. Some of the mangrove species that can be found in the Kachchh district include Avicennia marina, Avicennis officinalis, Avicennia alba, Rhizophora mucronata, Ceriopstagal, and Aegiceras corniculatum. The coastal areas of the Kachchh district are home to a variety of mangrove association species, some of which include Salvadora oleoides, Suaeda fruticosa, and Suaeda nudiflora. They grow in the areas between land and sea known as intertidal zones and estuary mouths, which makes them an essential habitat for a wide variety of marine and terrestrial plant and animal life [42].

The study area Guneri, figure 1, is a natural inland mangrove site that spans 33 hectares, including the buffer zone, and is about a century old. Due to their unique natural condition, the Gujarat government is seeking to declare the inland mangrove site in Guneri village of Kutch on the Western IndoPakistan border as a Biodiversity Heritage Site (BHS). Mangroves are typically found in coastal locations. However, the Guneri location features a significant number of inland mangroves. It is also home to animals such as chinkara, ratel, and migratory birds. Mangroves are small plants or trees that thrive in

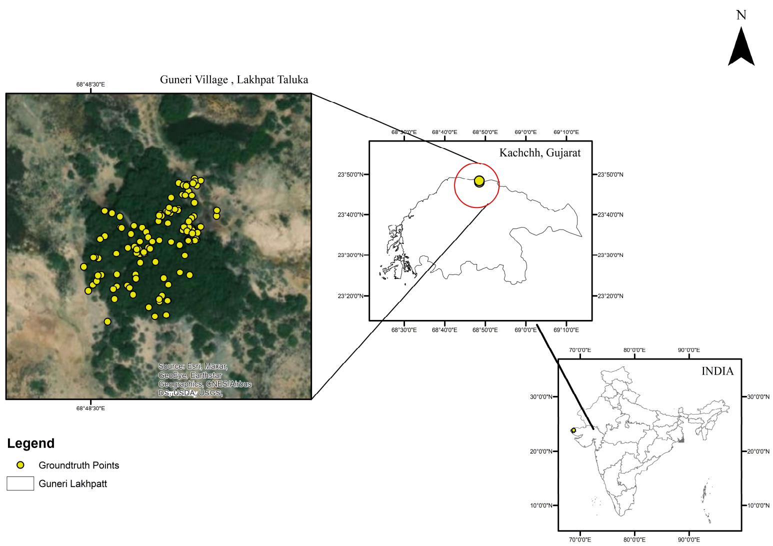

Fig. 1 The geographical extent of the Guneri Mangrove forest near Indo –Pak border, Kutch along with the spatial distribution ground truth sample points over the study area

saline or brackish water and are found in coastal areas, according to the definition. Guneri mangroves in Kutch, on the other hand, are fully landlocked and so unique. According to the Gujarat Institute of Desert Ecology's marine and coastal ecology section, Kutch's inland mangroves are one of only three or four of their sort in the world (GUIDE). Brazil, Peru, and South America are the others. The Guneri mangrove has a fascinating history. Approximately two millennia ago, the Rann of Kutch was a shallow sea. As the water level dropped over time, the dried soil was lifted owing to tectonic upheaval. The Arabian Sea moved away from the coast by 100-150 kilometers. As a result, the inland mangroves of Guneri have survived thanks to a subterranean source of brackish or saline water. These mangroves, according to researchers, are part of ancient bio-genetic pools. Seed distribution and propagation are aided by a plentiful subsurface supply of brackish water. These groves are fast deteriorating earlier they used to be 30-35 hectares in size, but now they have reduced to just around 13 hectares. It has ceased to regenerate; propagation has ceased. The hurricane of 1998 devastated a large portion of the groves, and the wood borer bug decimated the rest. It's also possible that the subsurface brackish water supply has been cut off. The precise cause is unknown; however, it could be due to human activity or climatic changes. As a result, the Guneri mangrove is an important study site[36, 43]. Avicennia Marina is the only species that dominates the site, as shown in figure 2. The ground truth was gathered between March 30ˆth and April 20ˆth, 2022. Random sampling was used to determine various data such as GPS positions, mangrove height, breast height breadth, and girth diameter.

Ann. For. Res. 66(1): 3114-3142, 2023
ISSN: 18448135, 20652445

Fig. 2 Actual images of Avicennia Marina, the only species that dominates the site

Data Collection

Ground Truth Data Collection

Remote sensing research can't be done without collecting data from the ground. It is usually done to calibrate a remote sensor, help correct remote sensing data, provide reliable data to identify each feature of interest in an image to help and guide the process of image interpretation/analysis, find representative areas of each image feature to generate their spectral signatures in order to model the spectral behaviour of specific Earth surface features, and check the accuracy of thematic maps created. Like most types of forests, mangroves can be divided into five carbon pools: 1) biomass of living plants above ground, 2) biomass of living plants below ground, 3) dead wood, 4) forest floor (litter), and 5) soil. Non-tree plants and litter are usually small parts of the ecosystem in mangroves, so they can often be left out of measurements without affecting how accurate the sample is. Trees are always part of the carbon pool study because they are easy to measure, there are good mounting calculations for them, and land use has a big effect on them [44]. The aboveground carbon pool is dominated by trees, which provide a clear sign of land-use change and ecological health. It is critical to measure trees precisely and completely. The basic idea is that allometric equations are used to estimate tree biomass and carbon stock by species using measurements of stem diameter (and sometimes height). The study site in the research is Guneri Mangroves of Kutch, which are different from coastal mangroves and are having tree-like structures instead of shrubs, hence they are a good carbon pool for above-ground biomass. The ground truth data is collected through simple random sampling. Here observation locations are picked at random as shown in Figure 3. Without any human bias, randomness assures that all portions of the study area have an equal probability of being sampled and a total of 123 trees were randomly picked and measured for their height and DBH. DBH was calculated by measuring the circumference of the stem 4 and a ½ feet above the ground as shown in the figure (Fig. b, c, d). The circumference was calculated to DBH using the equation 1:

$$DBH = circumfrence/\pi, (where\ \pi = 3.14). \qquad (1)$$

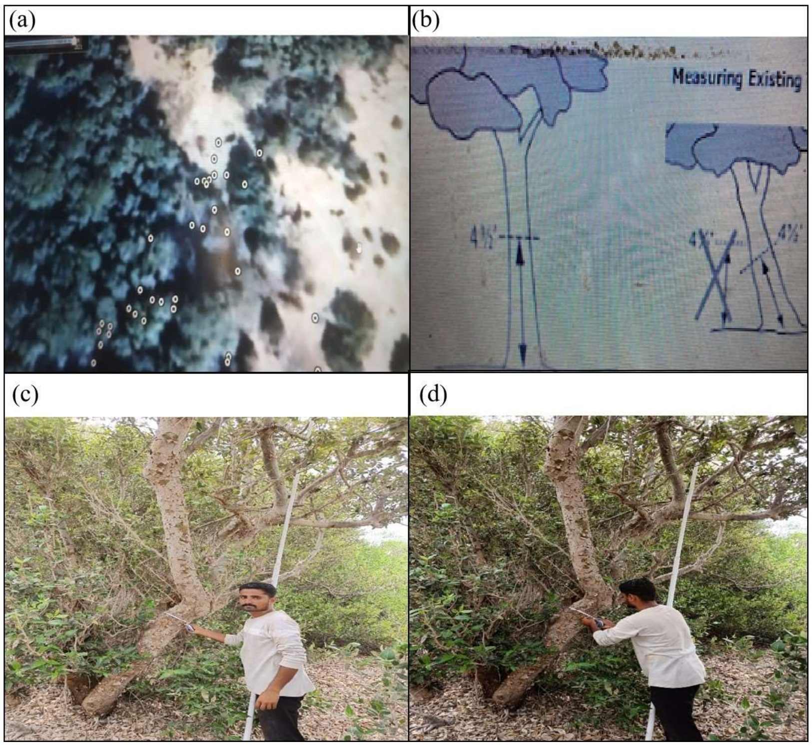

Fig. 3 (a) Random Sampling (b) technique to measure the circumference of stem for DBH calculation DBH (c) and (d) physical measurement circumference

The destructive traditional harvesting methods are the most effective and efficient method for estimating mangrove AGB. This method can be used to construct allometric equations based on the measured data from the harvested trees, such as the diameter at breast height (DBH), tree height, and timber volume. However, because mangroves are slow-growing trees, the conventional method of determining AGB is not recommended. As a result, remote sensing is playing an important role in mapping and estimating AGB of mangroves. In the current study to generate the ground truth AGB an existing allometric equation [45] is used, as shown in equation 2; the choice of an allometric model is crucial and should be based on the study's goal [46] and the dataset's characteristics. Allometric models should reflect the DBH range and ecology under investigation [47].

$$AGB = 0.162 * H^{1.81} * DBH^{1.24}. \qquad (2)$$

where, H is height and DBH is breast height diameter .

Reference Samples

Accurate visual interpretation in this work required gathering reference data from high-resolution satellite images. Additionally, earlier mangrove ecosystem maps and false-colour composite (FCC) satellite images were employed. To lessen the problem of mixed pixels by avoiding fragmented areas, homogeneous sites were taken into consideration for reference sample collecting. Three classes(mangrove, non-mangrove vegetation, and barren land) in total were created, each with sufficient reference samples and the right spatial distribution, as shown in figure no. .

The reference samples for the training and test were then arbitrarily split into two groups. Due to random splitting, the final classification scores exhibit little bias. The main issue with random sampling, however, is the information leak between training and test samples. In other words, due to pixel-level random sampling, the training and test datasets both include reference samples taken from the same polygons. The spatial auto-correlation of the training and test datasets increases as a result of this issue, which lowers the generality of the classifier and decreases the accuracy of evaluation results [48, 49]. As a result, the step was carried out at the polygon unit using a random splitting method, which also served to spatially separate the data used for training and testing. It is important to note that the random dividing phase was carried out ten times in order to facilitate the implementation of a cross-validation strategy for performance evaluation. This was done in order to demonstrate not only the applicability but also the resilience of the method that was suggested for accurate and comprehensive mangrove ecosystem mapping.

Satellite Data Collection

A mangrove ecosystem map was created using the time-series Sentinel-1 satellite photographs. The mangrove ecosystem's water level variations and tidal effects can also be taken into account using time-series data, which can improve the accuracy of the categorization outcomes [50]. The European SAR satellite Sentinel-1 has a temporal resolution of six days, and it collects data in

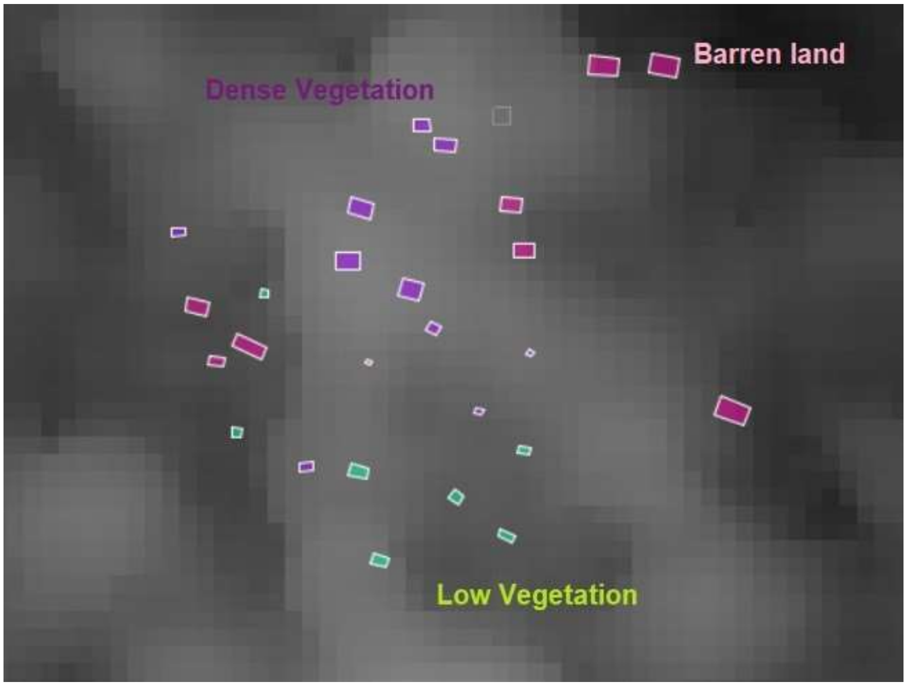

Fig. 4 Reference samples collected from study area using SNAP

the C-band with dual polarisation throughout the year. For this analysis, we utilised Level-1C ground range detected (GRD) images with a spatial resolution of 10 metres in both the ascending and descending directions. Sentinel-1A images were acquired for the year 2022 using the fine dual-polarization mode IW mode to produce the Ground Range Detected (GRD) product. The VV and VH polarisations are both present in the GRD. The images were selected with consideration given to the availability of data and the data collection taking place during periods of low tide. The Dual-Pol SAR that was utilised for the study is from the 17th of April in 2022. The detailed characterization of the image is per the table 3.

Table 3 Characteristics of Sentinel-1 SAR image

Product Type	**GRD**
Mission Sentinel	Sentinel-1A
Acquisition Mode	Interferometric Wide (IW)
Polarization	VV, VH
Mode	Fine Dual Pol
Duration	April 17th, 2022
Coverage (Swath Width)	250km

Geometric Resolution	5m by 20m
Pass Direction	Desencding
Geodetic CRS	WGS 84

SAR backscatter coefficient (VV & VH) extraction from ground truth data

As the backscatter coefficient indicates the microstructure of the things existing on the earth's surface, the correlation between the dielectric constant and the above ground biomass makes SAR data an excellent tool for determining how much vegetation is in the ground. For a given vegetation condition (forest, grassland, and desert etc.), there is a linear relationship between radar backscatter and volumetric tree structure [51]. Ground truth data is needed to estimate above-ground biomass using SAR image formats, including the height and width of the tree as well as the diameter at breast height, or DBH, which is a common way to represent the diameter of the base or branch of a growing tree. In this study, ground truth data (height, width and DBH) were collected and tagged with a GPS location representing the coordinates (latitude and longitude), these collected coordinates were projected on a pre-processed on Sentinel-1 SAR imagery, and VH VV backscatter coefficient values were extracted using QGIS 3.15 and python script. The extracted VH and VV values were correlated with the actual biomass data which was calculated from the collected observations from the sites. Overall, 123 separate tree samples were collected and AGB was calculated from the study region for the purpose of estimating the biomass of the trees. These observations were all geo-tagged and then superimposed on Sentinel-1 SAR images to obtain the values for the backscattered coefficient.

Synthetic Data Generation

Mangrove sampling and height and breadth measurements are extremely challenging since mangroves are prevalent in geographically challenging areas. Even so, we made an effort to gather as much real-world information as we could. Figure No. illustrates the site's problem. However, there weren't enough sample points collected to develop a model or train/test any machine learning models. So, in order to create a reliable model, we turned to the creation of synthetic data.

High volume, high velocity, and high diversity datasets are readily available, and they can be combined with sophisticated statistical methods for information extraction to enhance decision-making and speed up research and innovation. In addition, sharing many large-scale datasets that are highly sensitive (such as those in the fields of health or finance) may breach fundamental rights protected by current privacy laws (such as the GDPR or CCPA), and sometimes the data collection method is problematic due to difficult geographical dimensions. Numerous instances from the real world show that high-dimensional, frequently sparse datasets are naturally vulnerable to privacy attacks and that current anonymization techniques do not offer a sufficient defence. Due to this restriction on data sharing, machine learning and data science method development and application are slowed down [52].

A model produces fictitious data, frequently with the intention of substituting it for actual data. The end-user can theoretically change the amount of personal information disclosed by synthetic data and manage how closely it resembles actual data by regulating the data-generating process. In addition to addressing privacy issues, one can build convincing hypothetical situations and correct for biases in historical datasets. A significant amount of synthetic data can be produced from a small number of well-labeled data points, which would save the time and effort required to handle the enormous amount of realworld data. There are numerous methods for creating synthetic data, including SMOTE, ADASYN, Variational AutoEncoders, and Generative Adversarial Networks. We chose synthetic data creation since we had trouble finding additional data points during this research to help us create a model. In this study, generative adversarial networks were utilized to create artificial data (GAN). Many deep learning and machine learning architectures are vulnerable to adversarial manipulation, meaning that the models fall short when fed input that is different from the data used for training. Ian Goodfellow [53] introduced Generative Adversarial Networks (GANs) to address the adversarial challenge, and these networks are now widely used to produce synthetic data. A typical GAN has two parts, the discriminator and the generator, which are in competition with one another. The heart of the GAN is the generator, which uses attributes from real data to create false data that mimics genuine data. The discriminator compares the generated data to the real data and determines whether or not the generated data appears to be real. It then gives feedback to the generator so that it can enhance the quality of its data generation The actual ground points were inputted into GAN, to generate synthetic data. Figure 5 (a) is the principal component analysis where the comparison of the distribution of actual points with synthetically generated points is shown, 5(b) shows the analysis of width and latitude distribution and 5 (c) shows the analysis of height. Through these generated points we got a respectable amount of data points to build our model.

Methodology

Figure 5 presents the research framework, which gives a summary of the suggested strategy. This part is divided into three subsections, each of which provides a detailed explanation of the above-ground biomass estimating model, classification methodology, and preparation of satellite data.

SAR Data Pre-processing

Sentinel-1 With the fragment of (Image Collection ID: COPERNICUS/S1 GRD), GRD data are accessible within GEE. Because the GEE developers first apply a number of pre-processing processes to them, they are typically readyto-use data. These data have previously been ortho-rectified and transformed to the backscattering coefficient (dB). Each Sentinel-1 image underwent five

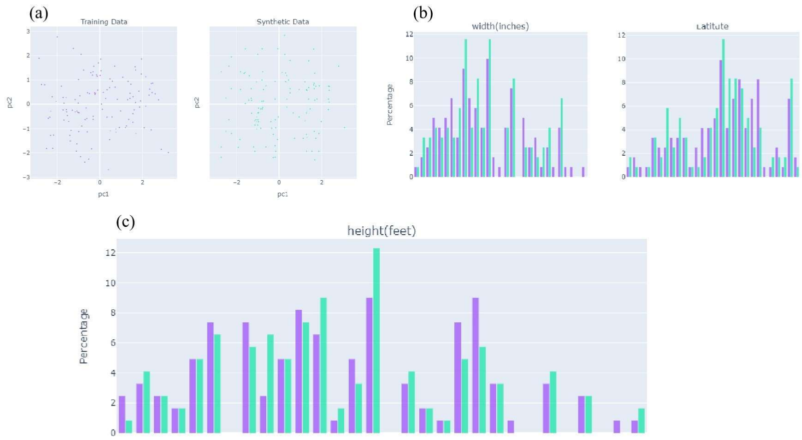

Fig. 5 Results of Synthetic data generation

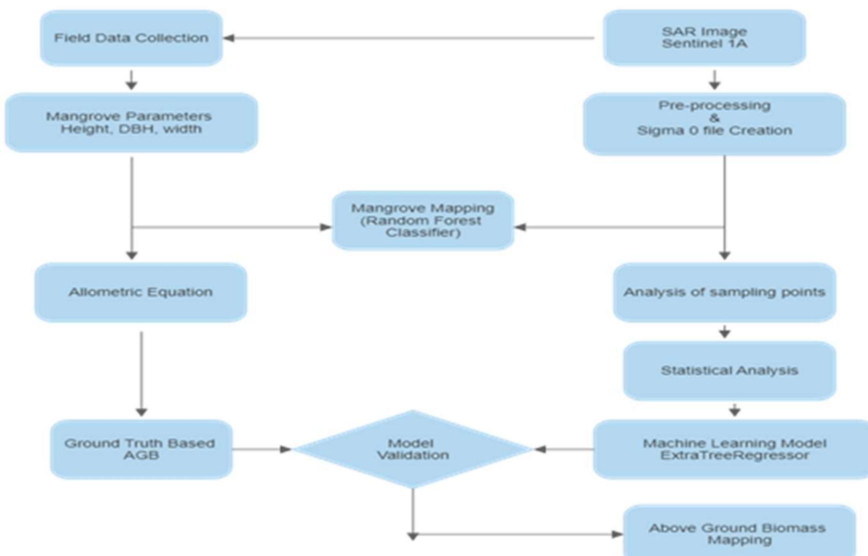

Fig. 6 Research Framework

pre-processing phases, including the following: orbit file correction, GRD border noise reduction, thermal noise removal, radiometric calibration, and terrain correction. Using the following equation, the digital numbers (DN) of SAR intensity data were transformed to Normalized Radar Cross section (NRCS or gamma-0) data (in dB):

$$y^{o}(dB) = 10 * log_{10}(DN)^{2} - CF \qquad (3)$$

CF is the calibration factor offered in the metadata file for each polarisation data point. The DN values of the SAR image were transformed into normalized backscatter values with the below coefficients [54]:

$$y^{o}(dB) = 10 * log_{10}(DN)^{2} - 83 \qquad (4)$$

Here, 83 is the calibration factor for dual polarized data.

Mangrove Classification

Using satellite imagery, various categorization techniques have been used to map mangroves. In this regard, selecting discriminative features and the best classifier is equally crucial and has a direct impact on the classification outcomes. Machine learning algorithms must be used to accurately and affordably categorize and map mangrove forests, and these algorithms must be learned using training datasets that have greater spatial resolution and algorithm optimal parameterization. The RF technique has been used in the past to map and categorise mangroves using remote sensing data and can offer a greater grade of classification than linear classifiers. The system does well at mapping mangroves on a regional scale and at handling data that contains unclassified pixels. One of the most well-liked techniques for non-parametric ensemble machine learning and high-quality mangrove categorization and environmental modelling is the random forest algorithm. Regression and classification trees are combined in it (CART) [55]. Random forest (RF), a classifier, has consistently shown to be a successful approach for mapping mangroves [56–58]. For example, when [59] tested four regularly used non-parametric classifiers for mapping the mangrove ecosystem, the RF classifier came out on top along with the support vector machine (SVM) with linear and outspread basis function kernels and regularised discriminant analysis. The tremendous potential of the RF classifier for mapping the mangrove environment led to the implementation of a pixel-based RF classifier within GEE for this project. In this case, an RF classifier was used to classify the photos that made up the seasonal Sentinel-1 data. The RF classifier was trained using half of the reference samples in the meantime. Numerous tuning parameters for the RF classifier have an impact on the classification step's training stage, directly affecting the classification outcomes. The number of trees and the variables at each node are the parameters that have the most influence.

Mangrove Above Ground Biomass Estimation

Hypothetically, stand density, DBH, and species all have a direct impact on forest AGB. An important topic is how to utilise multisource remote sensing datasets to their maximum potential. We have the biomass and associated imagery data over a set of sites inside a research region following the field campaign, remote sensing data acquisition, and processing. Assume, B ij = 1,2,3,...,N is the biomass and X is assumed to be the data vector. The AGB estimation is to discover the prediction model P:

$$B\hat{} = P(X) \qquad (5)$$

The purpose of developing this model to minimize the error of estimation:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(\hat{B}_i - B_i)}{N}} \qquad (6)$$

Machine learning techniques such neural networks, K nearest neighbor, regression trees (RT) like Extra Tree Regressor, XGB Regressor, AdaBoost Regressor, Random Forest, and MaxEnt are some linear and nonlinear regression machine learning methods that can be used as the prediction model. When creating prediction models using a parametric method like regression analysis, factors such as the spectral responses at optical data, backscattering, and attributes obtained from PolSAR and PolInSAR data, various indices from lidar data, and picture textures can be employed directly. The majority of machine learning algorithms, including MaxEnt, are regarded as nonparametric techniques that automatically pick up on intricate patterns and base judgments on data. It is possible to use data from various sources and in various formats, and the method for estimating biomass is known as spatial modelling of ABG [60]. Within the community of remote sensing research and applications, classification and random trees have grown in popularity [61, 62]. Random tree clearly outperforms traditional statistical techniques. It provides for the possibility of interactions and non-linearities across variables because no a priori assumptions are made regarding the nature of the relationships among the response (for example, biomass) and predictor factors (for example, remote sensing data). However, the precise selection of the training dataset accounts for a portion of the output error in a single RT. Random Forests are made to generate precise forecasts without over-fitting the data. Random forests refer to the process of building multiple trees using a randomised subset of variables using bootstrap samples. A ”forest” of trees is created through the growth of a vast number of trees (500-2000). A randomly selected subset of the entire number of predictors is utilised to determine the optimal split at each node [63–66]. Regression models for calculating AGB might be simple or multi-linear [67–69] or can be in-depth machine learning (ML) techniques [70–72]. Non-parametric methods utilising a variety of ML algorithms have shown to be more successful than parametric techniques using linear models for mapping and predicting forest AGBs. Lot of research is done in field of mangrove AGB mapping using non-parametric regression methods such as artificial neural network(ANN), random forest regression (RFR), support vector regression (SVM) and some recent studies have experimented with gradient boosting decision trees (GBDT) and extra gradient boost regression (XGBR) techniques [68, 70, 73, 74]. Particularly, a quantitative comparison of cuttingedge ML approaches for predicting AGBs in various forest ecosystems appears to be lacking in the existing research.

Data Normalization

It is common practise to use the process of normalisation in order to get data ready for machine learning. The objective of normalisation is to convert the values of the numeric columns in the dataset to a common scale without introducing any distortion into the variations in the value ranges. When it comes to machine learning, not every dataset requires the normalisation step. It is only required in situations where the feature has multiple ranges. Gradients may end up taking a long time, can fluctuate back and forth, and take a long time before they can eventually find their way to the global/local minimum because various features do not have identical ranges of values. Variables evaluated at varying scales have a differential impact on the model-fitting and modellearning functions; in some cases, these differences can even result in bias. In order to resolve this potential problem, feature-wise normalisation is typically carried out prior to the fitting of the model. In the current study, normalizing the input variables was observed to be important as the input features, SAR backscattered coefficients (VV, VH) and the ground truth AGB calculated through the allometric equations were not falling in the column range. Henceforth, the input features were normalized through the min-max normalization technique. The first step in the process involves linearly transforming the data using min-max normalisation, which is also known as feature scaling. All of the scaled data that fall within the range can be obtained by using this method (0, 1). The following is the formula for the min-max normalisation:

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{7}$$

The associations between the original data values are preserved by minmax normalization. We will have reduced standard deviations as a result of this restricted range, which can reduce the impact of outliers. The objective to use min-max normalization was to preserve the relationship among the original data values. This study is extensive research on Guneri mangrove forest, employing the usage of advance non-parametric machine learning algorithms. Where in the backscatter parameters (VV,VH) of the site locations are modelled against the ground truth values of the study site and the normalized field AGB using Extra Tree Regressor, XGB Regressor, Random Forest Regressor,
Bagging Regressor, Ada Boost regressor, K-Neighbors Regressor, SVR and MLP regressor. Comparative study of all the methods is shown in the study.

Results and Discussion

The section on results and discussion is divided into two subsections. The first subsection will go over the relationship between the derived biomass and the back-scattered values, as well as how to generate a biomass map from the predicted biomass values. The second subsection will go over land use land cover (LULC) mapping, and how to identify mangrove forests using random forest classification.

Above Ground Biomass (AGB) Estimation

The linear correlation between AGB and the backscattered values (VH & VV polarisation) was initially checked using the Multiple Linear Regression (MLR) technique on the collected 123 sample points, and 52.17% validation accuracy was achieved on 50 test dataset with RMSE of 0.16. Generative Adversarial Network (GAN) was used to generate synthetic data from the provided sample because it appeared that there were not enough data points to train and generate any promising models. With the use of GAN, data was doubled, totaling 246 sample points, coordinates for the generated samples were also generated using the same technique and cross validated by overlaying it on the Google Earth image, ensuring that all of the generated data points were within the study area. The produced dataset was once more subjected to MLR, however this time there was only a very slight improvement in the validation accuracy of 53.81% with RMSE of 0.16. To address this, we applied maximum normalisation to the AGB values, which divides each row by its highest absolute value rather than its average. This technique increased accuracy by 12%. Other normalisation techniques, such as mean, min-max, and range, were also used, however the outcome was marginally lower to that of maximum normalisation.

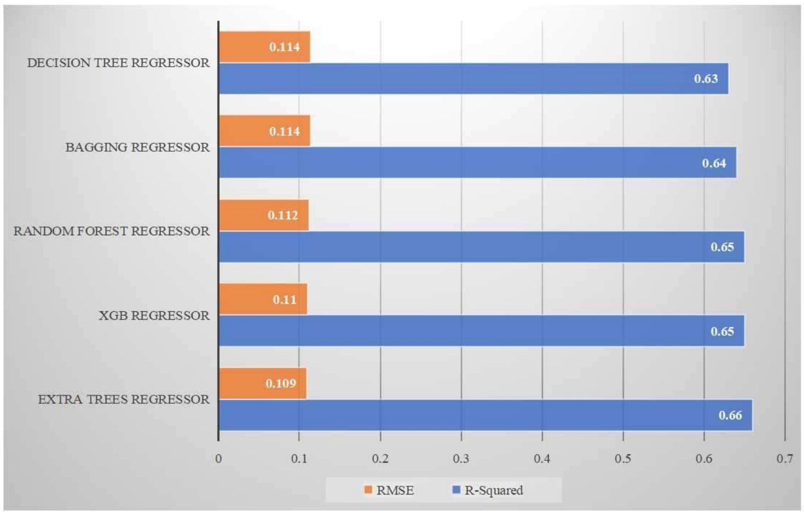

Fig. 7 Results of Synthetic data generation

The generated data was then tested using several regression methods such as Extra Trees Regressor, XGB Regressor, Random Forest Regressor, Bagging Regressor, and Decision Tree Regressor. The comparative study of the selected models is shown in the figure 7, and Extra Trees Regressor demonstrated a good validation accuracy of 66% with 0.10 RMSE, followed by XGB Regressor (65%), and Random Forest Regressor (65%).

We have considered evaluating the individual polarised bands (VV and VH) using the extra tree regressor model as well as the xgb regressor model based on the low RMSE achieved when compared with the other three regressor model. The predicted and actual AGB from individual VV and VH bands, as well as combinations of the two, are compared in figures 8 & 9 , respectively. Table 4 displays the performances of the models on different SAR band combinations. The mean cross validation score of 0.37 was high on the VH band when the extra tree model was used, whereas the score was only 0.35 for the same VH band when the XGB regressor was used. From this we can conclude that VH band contributes more in estimating above ground biomass.

After taking the VH band into consideration, we used an extra tree regressor to make a prediction about the AGB of the entire Guneri area. The heat map shown in figure 10 was generated from the predicted AGB by using the VH SAR band. The area with an AGB ranging from 171-193 kg/m2 had the highest biomass, followed by the area with an AGB ranging from 128-171 kg/m2. Land that has a biomass anywhere between 0 and 107 kg/m2 can be referred to as barren land.

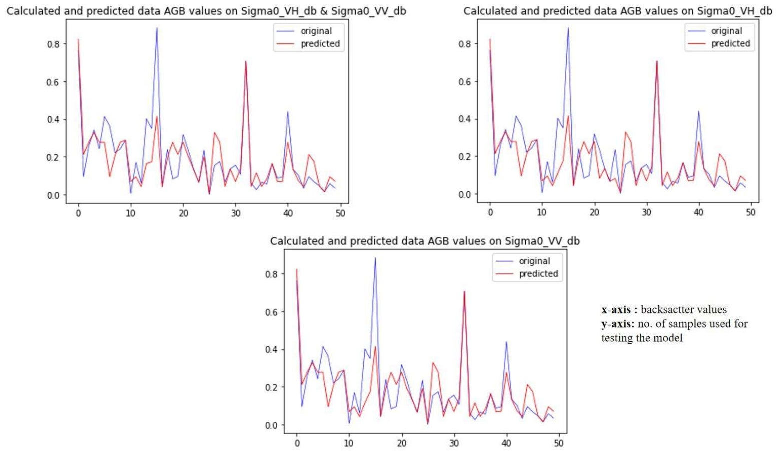

Fig. 8 Predicted AGB values using Extra Tree Regressor

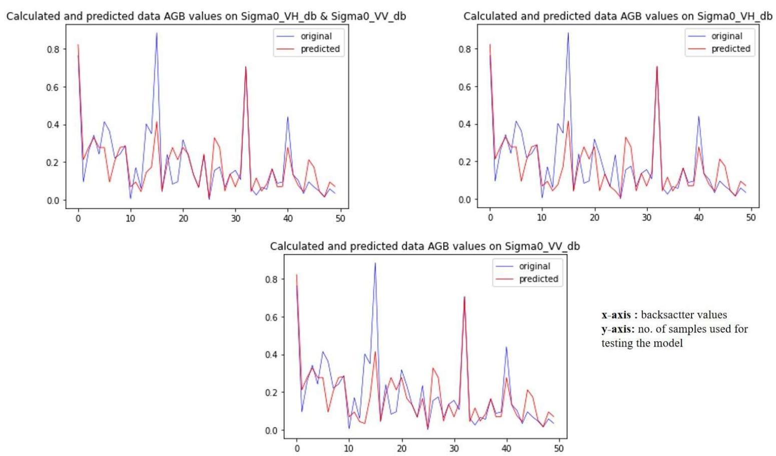

Fig. 9 Predicted AGB values using XGB Regressor

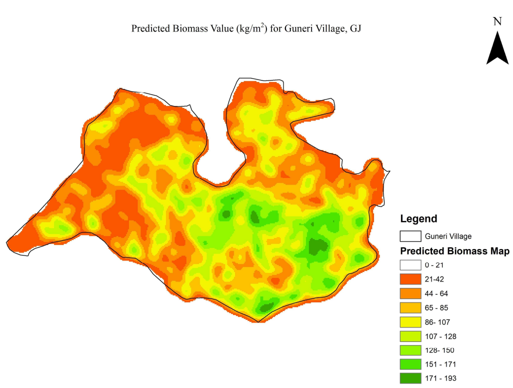

Fig. 10 Heat map generated from predicted AGB using VH band **Table 4** Mangrove discrimination from other vegetation using Sentinel imaging and machine learning

Models	Band Used	R2	RMSE	MSE	Mean CV Scores
Extra Trees Regressor	Sigma0 VH db	0.65	0.12	0.01	0.37
Extra Trees Regressor	Sigma0 VV db	0.65	0.11	0.01	0.29
Extra Trees Regressor	Sigma0 VH db & Sigma0 VV db	0.65	0.11	0.01	0.34
XGB Regressor	Sigma0 VH db	0.65	0.12	0.01	0.35
XGB Regressor	Sigma0 VV db	0.65	0.12	0.01	0.27
XGB Regressor	Sigma0 VH db & Sigma0 VV db	0.65	0.11	0.01	0.31

LULC mapping of the study area

If broadly categorized the land cover features present in the study area then it contains only three types of features namely barren land, mangrove forest and grass/Non-mangrove. We will be using supervised classification techniques to classify the study area into three mentioned categorise. Random Forest machine learning algorithm was used for LULC classification. We first assigned each pixel in the Sentinel-1 SAR image to a specific land use or land cover class by referring Sentinel-2 multi-spectral and google earth image. The RF algorithm uses multiple decision trees to make predictions, and the final prediction is determined by a majority vote among the individual trees. Random Forest is known for its high accuracy, ability to handle large datasets, and ability to handle missing data. It can also be used to identify important features in the data, which can be useful for understanding the factors that influence land use and land cover patterns. Figure 11 (a) shows the classified map generated using RF, (b) optical image of the study area, and (c) geo-referenced classification map overlaye on google earth image. The achieved accuracy with K fold RF classification is 91%, the RMSE is 0.526, the contribution of VH spectrum is 0.506, and the contribution of VV spectrum is 0.339. According to the classified map, the overall land distribution of the study area is as follows: 28% barren land, 18% mangroves, and 54% grass/non-mangrove.

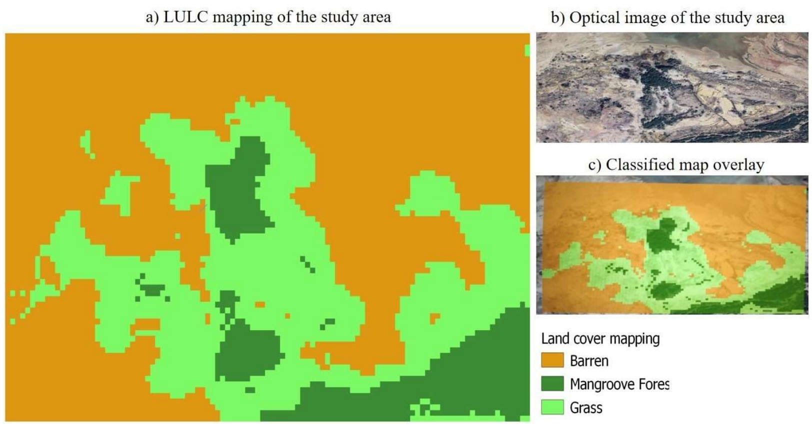

Fig. 11 Classified LULC map using random forest ML techniques

Conclusion

The landlocked inland Guneri mangrove forest is one of the oldest mangrove forests in India, which is absolutely disconnected from the sea and makes it one the very few places in the world to sustain such a unique ecological feature. Besides having such uniqueness, these mangroves are overlooked. Hence, this research is an initiative that can help preserve this bio-diversified heritage site. In this study a workflow is proposed to produce a mangrove ecosystem map and mangrove AGB estimation resulting in respectable accuracy. For mangrove and non-mangrove classification, a simple but robust Random Forest classifier is used, which produces an average accuracy of 91% and RMSE of 0.506. For AGB model generation machine learning techniques are applied to the datasets. The generated data were then tested using several regression methods such as Extra Trees Regressor, XGB Regressor, Random Forest Regressor, Bagging Regressor, and Decision Tree Regressor. The comparative study of the selected models is shown in the figure above, and Extra Trees Regressor demonstrated a good validation accuracy of 66% with 0.10 RMSE, followed by XGB Regressor (65%), and Random Forest Regressor (65%). This work validates the applicability of Random Forest (RF) and Extra Trees Regressor algorithms for mapping and estimating AGB for a unique landlocked mangrove site of Guneri, and it is observed that the results and robustness of the model are highly affected by the usage of a larger dataset and the geographical parameters of the study site.

References

[1] Upadhyay, V., Mishra, P.: Population status of mangrove species in estuarine regions of orissa coast, india. Tropical Ecology **49**(2), 183 (2008)

[2] Snedaker, S.C.: Mangrove species zonation: why? In: Contributions to the Ecology of Halophytes, pp. 111–125. Springer, ??? (1982)

[3] Bennett, E.L., Reynolds, C.J.: The value of a mangrove area in sarawak. Biodiversity & Conservation **2**(4), 359–375 (1993)

[4] Kuenzer, C., Bluemel, A., Gebhardt, S., Quoc, T.V., Dech, S.: Remote sensing of mangrove ecosystems: A review. Remote Sensing **3**(5), 878–928 (2011)

[5] Donato, D.C., Kauffman, J.B., Murdiyarso, D., Kurnianto, S., Stidham, M., Kanninen, M.: Mangroves among the most carbon-rich forests in the tropics. Nature geoscience **4**(5), 293–297 (2011)

[6] Le Toan, T., Quegan, S., Davidson, M., Balzter, H., Paillou, P., Papathanassiou, K., Plummer, S., Rocca, F., Saatchi, S., Shugart, H., *et al.*: The biomass mission: Mapping global forest biomass to better understand the terrestrial carbon cycle. Remote sensing of environment **115**(11), 2850–2860 (2011)

[7] Molto, Q., Rossi, V., Blanc, L.: Error propagation in biomass estimation in tropical forests. Methods in Ecology and Evolution **4**(2), 175–183 (2013)

[8] Mumby, P., Green, E., Edwards, A., Clark, C.: The cost-effectiveness of remote sensing for tropical coastal resources assessment and management. Journal of Environmental management **55**(3), 157–166 (1999)

[9] Nandy, S., Kushwaha, S.: Study on the utility of irs 1d liss-iii data and the classification techniques for mapping of sunderban mangroves. Journal of Coastal Conservation **15**(1), 123–137 (2011)

[10] Alatorre, L.C., S´anchez-Andr´es, R., Cirujano, S., Beguer´ıa, S., S´anchezCarrillo, S.: Identification of mangrove areas by remote sensing: The roc curve technique applied to the northwestern mexico coastal zone using landsat imagery. Remote Sensing **3**(8), 1568–1583 (2011)

[11] Giri, C., Pengra, B., Zhu, Z., Singh, A., Tieszen, L.L.: Monitoring mangrove forest dynamics of the sundarbans in bangladesh and india using multi-temporal satellite data from 1973 to 2000. Estuarine, coastal and shelf science **73**(1-2), 91–100 (2007)

[12] Roy, P.: Mangrove vegetation stratification using salyut 7 photographs. Geocarto International **4**(3), 31–47 (1989)

[13] Dwivedi, R., Rao, B., Bhattacharya, S.: Mapping wetlands of the sundaban delta and it's environs using ers-1 sar data. International Journal of Remote Sensing **20**(11), 2235–2247 (1999)

[14] Kushwaha, S., Dwivedi, R., Rao, B.: Evaluation of various digital image processing techniques for detection of coastal wetlands using ers-1 sar data. International Journal of Remote Sensing **21**(3), 565–579 (2000)

[15] Medeiros, T., Sampaio, E.: Allometry of aboveground biomasses in mangrove species in itamarac´a, pernambuco, brazil. Wetlands Ecology and Management **16**(4), 323–330 (2008)

[16] Tam, N., Wong, Y.S., Lan, C., Chen, G.: Community structure and standing crop biomass of a mangrove forest in futian nature reserve, shenzhen, china. Hydrobiologia **295**(1), 193–201 (1995)

[17] Neumann, M., Saatchi, S.S., Ulander, L.M., Fransson, J.E.: Assessing performance of l- and p-band polarimetric interferometric sar data in estimating boreal forest above-ground biomass. IEEE Transactions on Geoscience and Remote Sensing **50**(3), 714–726 (2012)

[18] Sinha, S., Jeganathan, C., Sharma, L.K., Nathawat, M.S.: A review of radar remote sensing for biomass estimation. International Journal of Environmental Science and Technology **12**(5), 1779–1792 (2015)

[19] Chen, B., Xiao, X., Li, X., Pan, L., Doughty, R., Ma, J., Dong, J., Qin, Y., Zhao, B., Wu, Z., *et al.*: A mangrove forest map of china in 2015: Analysis of time series landsat 7/8 and sentinel-1a imagery in google earth engine cloud computing platform. ISPRS Journal of Photogrammetry and Remote Sensing **131**, 104–120 (2017)

[20] Hu, L., Li, W., Xu, B.: Monitoring mangrove forest change in china from 1990 to 2015 using landsat-derived spectral-temporal variability metrics. International journal of applied earth observation and geoinformation **73**, 88–98 (2018)

[21] Jia, M., Wang, Z., Wang, C., Mao, D., Zhang, Y.: A new vegetation index to detect periodically submerged mangrove forest using single-tide sentinel-2 imagery. Remote Sensing **11**(17), 2043 (2019)

[22] Zhao, C., Qin, C.-Z.: 10-m-resolution mangrove maps of china derived from multi-source and multi-temporal satellite observations. ISPRS Journal of Photogrammetry and Remote Sensing **169**, 389–405 (2020)

[23] Jiang, Y., Zhang, L., Yan, M., Qi, J., Fu, T., Fan, S., Chen, B.: Highresolution mangrove forests classification with machine learning using worldview and uav hyperspectral data. Remote Sensing **13**(8), 1529 (2021)

[24] Cao, J., Liu, K., Zhuo, L., Liu, L., Zhu, Y., Peng, L.: Combining uav-based hyperspectral and lidar data for mangrove species classification using

the rotation forest algorithm. International Journal of Applied Earth Observation and Geoinformation **102**, 102414 (2021)

[25] Kovacs, J.M., Wang, J., Flores-Verdugo, F.: Mapping mangrove leaf area index at the species level using ikonos and lai-2000 sensors for the agua brava lagoon, mexican pacific. Estuarine, Coastal and Shelf Science **62**(12), 377–384 (2005)

[26] Su, X., Geng, J., Ma, R., Wang, H., Wang, X.: Mangrove species classification based on multiple vegetation index extraction and joint sparse representation. Mar. Environ. Sci **36**, 114–120 (2017)

[27] Xia, Q., Qin, C.-Z., Li, H., Huang, C., Su, F.-Z.: Mapping mangrove forests based on multi-tidal high-resolution satellite imagery. Remote Sensing **10**(9), 1343 (2018)

[28] Kumar, L., Schmidt, K., Dury, S., Skidmore, A.: Imaging spectrometry and vegetation science. In: Imaging Spectrometry, pp. 111–155. Springer, ??? (2002)

[29] McDonald, M.S.: Photobiology of Higher Plants. John Wiley & Sons, ??? (2003)

[30] Schmidt, K., Skidmore, A.: Spectral discrimination of vegetation types in a coastal wetland. Remote sensing of Environment **85**(1), 92–108 (2003)

[31] Huang, K., Yang, G., Yuan, Y., Sun, W., Meng, X., Ge, Y.: Optical and sar images combined mangrove index based on multi-feature fusion. Science of Remote Sensing **5**, 100040 (2022)

[32] Ghorbanian, A., Zaghian, S., Asiyabi, R.M., Amani, M., Mohammadzadeh, A., Jamali, S.: Mangrove ecosystem mapping using sentinel-1 and sentinel-2 satellite images and random forest algorithm in google earth engine. Remote Sensing **13**(13), 2565 (2021)

[33] Gupta, K., Mukhopadhyay, A., Giri, S., Chanda, A., Majumdar, S.D., Samanta, S., Mitra, D., Samal, R.N., Pattnaik, A.K., Hazra, S.: An index for discrimination of mangroves from non-mangroves using landsat 8 oli imagery. MethodsX **5**, 1129–1139 (2018)

[34] Kumar, T., Mandal, A., Dutta, D., Nagaraja, R., Dadhwal, V.K.: Discrimination and classification of mangrove forests using eo-1 hyperion data: A case study of indian sundarbans. Geocarto International **34**(4), 415–442 (2019)

[35] Ghosh, S., Behera, M., Jagadish, B., Das, A., Mishra, D.: A novel approach for estimation of aboveground biomass of a carbon-rich mangrove site in india. Journal of Environmental Management **292**, 112816 (2021)

[36] Vaghela, B., Chirakkal, S., Putrevu, D., Solanki, H.: Modelling above ground biomass of indian mangrove forest using dual-pol sar data. Remote Sensing Applications: Society and Environment **21**, 100457 (2021)

[37] Nesha, M.K., Hussin, Y.A., van Leeuwen, L.M., Sulistioadi, Y.B.: Modeling and mapping aboveground biomass of the restored mangroves using alos-2 palsar-2 in east kalimantan, indonesia. International Journal of Applied Earth Observation and Geoinformation **91**, 102158 (2020)

[38] Barford, D., Das, A.K., Egloff, M.-P.: The structure and mechanism of protein phosphatases: insights into catalysis and regulation. Annual review of biophysics and biomolecular structure **27**(1), 133–164 (1998)

[39] Lister, T., Wright, P.A., Chappell, P.H.: Optical properties of human skin. Journal of biomedical optics **17**(9), 090901 (2012)

[40] Vinod, K., Anasu Koya, A., Kunhikoya, V., Shilpa, P., Asokan, P., Zacharia, P., Joshi, K.: Biomass and carbon stocks in mangrove stands of kadalundi estuarine wetland, south-west coast of india. Indian Journal of Fisheries **65**(2), 89–99 (2018)

[41] Sahu, S., Kumar, M., Ravindranath, N.: Carbon stocks in natural and planted mangrove forests of mahanadi mangrove wetland, east coast of india. Current Science, 2253–2260 (2016)

[42] Hutchings, P., Saenger, P.: Ecology of mangroves. Ecology of mangroves (1987)

[43] Rahman Azera, P.: It takes a village – and an inspired leader – to protect kutch's unique mangroves. The Wire Science. Accessed 2020-10-02

[44] Kauffman, J.B., Donato, D.C., *et al.*: Protocols for the Measurement, Monitoring and Reporting of Structure, Biomass and Carbon Stocks in Mangrove Forests vol. 86. Cifor Bogor, Indonesia, ??? (2012)

[45] Ray, R., Ganguly, D., Chowdhury, C., Dey, M., Das, S., Dutta, M., Mandal, S.K., Majumder, N., De, T., Mukhopadhyay, S., *et al.*: Carbon sequestration and annual increase of carbon stock in a mangrove forest. Atmospheric Environment **45**(28), 5016–5024 (2011)

[46] Phillips, O.L., Arag˜ao, L.E., Lewis, S.L., Fisher, J.B., Lloyd, J., L´opezGonz´alez, G., Malhi, Y., Monteagudo, A., Peacock, J., Quesada, C.A., *et al.*: Drought sensitivity of the amazon rainforest. Science **323**(5919), 1344–1347 (2009)

[47] Chave, J., Condit, R., Aguilar, S., Hernandez, A., Lao, S., Perez, R.: Error propagation and scaling for tropical forest biomass estimates. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences **359**(1443), 409–420 (2004)

[48] Lambin, E.F.: Modelling and monitoring land-cover change processes in tropical regions. Progress in physical geography **21**(3), 375–393 (1997)

[49] Lohr, S.L.: Sampling: Design and Analysis. Chapman and Hall/CRC, ??? (2021)

[50] Geiß, C., Pelizari, P.A., Schrade, H., Brenning, A., Taubenb¨ock, H.: On the effect of spatially non-disjoint training and test samples on estimated model generalization capabilities in supervised classification with spatial features. IEEE Geoscience and Remote Sensing Letters **14**(11), 2008–2012 (2017)

[51] Ulaby, F.T., Held, D., Donson, M.C., McDonald, K.C., Senior, T.B.: Relating polaization phase difference of sar signals to scene properties.
IEEE Transactions on Geoscience and Remote Sensing (1), 83–92 (1987)

[52] Yoon, J., Jordon, J., Schaar, M.: Gain: Missing data imputation using generative adversarial nets. In: International Conference on Machine Learning, pp. 5689–5698 (2018). PMLR

[53] Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial networks. arXiv preprint arXiv:1406.2661 (2014)

[54] Shimada, M., Isoguchi, O., Tadono, T., Isono, K.: Palsar radiometric and geometric calibration. IEEE Transactions on Geoscience and Remote Sensing **47**(12), 3915–3932 (2009)

[55] Elmahdy, S.I., Ali, T.A., Mohamed, M.M., Howari, F.M., Abouleish, M., Simonet, D.: 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. Frontiers in Environmental Science **8**, 102 (2020)

[56] Manna, S., Raychaudhuri, B.: Mapping distribution of sundarban mangroves using sentinel-2 data and new spectral metric for detecting their health condition. Geocarto International **35**(4), 434–452 (2020)

[57] Mondal, P., Liu, X., Fatoyinbo, T.E., Lagomasino, D.: Evaluating combinations of sentinel-2 data and machine-learning algorithms for mangrove mapping in west africa. Remote Sensing **11**(24), 2928 (2019)

[58] Quang, N.H., Quinn, C.H., Stringer, L.C., Carrie, R., Hackney, C.R., Van Hue, L.T., Van Tan, D., Nga, P.T.T.: Multi-decadal changes in mangrove extent, age and species in the red river estuaries of viet nam. Remote Sensing **12**(14), 2289 (2020)

[59] Toosi, N.B., Soffianian, A.R., Fakheran, S., Pourmanafi, S., Ginzler, C., Waser, L.T.: Comparing different classification algorithms for monitoring mangrove cover changes in southern iran. Global Ecology and Conservation **19**, 00662 (2019)

[60] Saatchi, S.S., Harris, N.L., Brown, S., Lefsky, M., Mitchard, E.T., Salas, W., Zutta, B.R., Buermann, W., Lewis, S.L., Hagen, S., *et al.*: Benchmark map of forest carbon stocks in tropical

regions across three continents. Proceedings of the national academy of sciences **108**(24), 9899–9904 (2011)

[61] Goetz, S.J., Bunn, A.G., Fiske, G.J., Houghton, R.A.: Satellite-observed photosynthetic trends across boreal north america associated with climate and fire disturbance. Proceedings of the National Academy of Sciences **102**(38), 13521–13525 (2005)

[62] Hansen, M.C., Stehman, S.V., Potapov, P.V., Loveland, T.R., Townshend, J.R., DeFries, R.S., Pittman, K.W., Arunarwati, B., Stolle, F., Steininger, M.K., *et al.*: Humid tropical forest clearing from 2000 to 2005 quantified by using multitemporal and multiresolution remotely sensed data. Proceedings of the National Academy of Sciences **105**(27), 9439–9444 (2008)

[63] Breiman, L.: Random forest, vol. 45. Mach Learn **1** (2001)

[64] Breiman, L.: Using models to infer mechanisms. IMS Wald Lecture **2**, 59–71 (2002)

[65] Liaw, A., Wiener, M., *et al.*: Classification and regression by randomforest. R news **2**(3), 18–22 (2002)

[66] Prasad, A., Iverson, L., Liaw, A., Ecosystems, S., Mar, N.: Newer tree classification and techniques: Forests random prediction bagging for ecological regression. Ecosystems **9**, 181–199 (2006)

[67] Hamdan, O., Aziz, H.K., Hasmadi, I.M.: L-band alos palsar for biomass estimation of matang mangroves, malaysia. Remote Sensing of Environment **155**, 69–78 (2014)

[68] Pham, T.D., Yoshino, K.: Aboveground biomass estimation of mangrove species using alos-2 palsar imagery in hai phong city, vietnam. Journal of Applied Remote Sensing **11**(2), 026010 (2017)

[69] Wang, D., Wan, B., Liu, J., Su, Y., Guo, Q., Qiu, P., Wu, X.: Estimating aboveground biomass of the mangrove forests on northeast hainan island in china using an upscaling method from field plots, uav-lidar data and sentinel-2 imagery. International Journal of Applied Earth Observation and Geoinformation **85**, 101986 (2020)

[70] Pham, L.T., Brabyn, L.: Monitoring mangrove biomass change in vietnam using spot images and an object-based approach combined with machine learning algorithms. ISPRS Journal of Photogrammetry and Remote Sensing **128**, 86–97 (2017)

[71] Jachowski, N.R., Quak, M.S., Friess, D.A., Duangnamon, D., Webb, E.L., Ziegler, A.D.: Mangrove biomass estimation in southwest thailand using machine learning. Applied Geography **45**, 311–321 (2013)

[72] Zhu, Y., Liu, K., Liu, L., Wang, S., Liu, H.: Retrieval of mangrove aboveground biomass at the individual species level with worldview-2 images. Remote Sensing **7**(9), 12192–12214 (2015)

[73] Huang, G., Wu, L., Ma, X., Zhang, W., Fan, J., Yu, X., Zeng, W., Zhou, H.: Evaluation of catboost method for prediction of reference evapotranspiration in humid regions. Journal of Hydrology **574**, 1029–1041 (2019)

[74] Wu, J., Miu, F., Li, T.: Daily crude oil price forecasting based on improved ceemdan, sca, and rvfl: a case study in wti oil market. Energies **13**(7), 1852 (2020)