Enhancing Forest Canopy Height Mapping in Kaziranga National Park, Assam, by Integrating LISS IV and SAR data with GEDI LiDAR data Using Machine Learning

Mallika Bhuyan[1], Chockalingam Jeganathan[1], Girish S. Pujar[2]

[1] Department of Remote Sensing, BIT-Mesra, Ranchi, Jharkhand, India – 835215
[2] Rural Development & Watershed Monitoring Division, RSA, National Remote Sensing Centre (NRSC), Balanagar, Hyderabad, India – 500018
phdrs10001.20@bitmesra.ac.in, jeganathanc@bitmesra.ac.in, pujar@nrsc.gov.in

Keywords: Canopy Height, Random Forest, Extreme Gradient Boosting, Support Vector Machine, K-Nearest Neighbour, LISS IV.

Abstract

This study investigates the integration of spaceborne Global Ecosystem Dynamics Investigation (GEDI) LiDAR data with optical imagery from Linear Imaging Self Scanning Sensor (LISS IV), Sentinel-1 Synthetic Aperture Radar (SAR) and PALSAR data for continuous forest canopy height mapping in Kaziranga National Park (KNP), Assam for the years 2018 and 2022. Four machine learning models were trained and evaluated to assess the predictive ability of LISS IV data in conjunction with SAR variables. The mean canopy height was measured at approximately 8.58 m in 2018, which increased to about 9.07m in 2022. Results reveal Extreme Gradient Boosting (XGB) as the top-performing model, achieving an RMSE of 5.47m and an R^2 of 0.55. In comparison, Random Forest (RF), Support Vector Machine (SVM), and k-Nearest Neighbors (kNN) achieved RMSE values of 5.49, 5.51, and 5.73, respectively. Analysis shows a prevalent occurrence of canopy heights below 5 meters in KNP (more than 35% of the area), while taller canopies beyond 20m can be found in less than 5% of the area. This finding underscores the importance of integrating satellite data and machine learning and highlights the novel application of LISS IV data in enhancing canopy height mapping. Furthermore, it represents the first comprehensive attempt to map canopy height in KNP, laying the groundwork for further research on biomass assessment and carbon sequestration in this vital biodiversity hotspot. Overall, the study highlights the potential of leveraging advanced remote sensing technologies and machine learning approaches for improved understanding and management of forest ecosystems.

1. Introduction

Forest canopy height (CH) serves as both a product and driver of ecosystem processes, influencing biomass allocation, carbon storage, productivity, and biodiversity (Zhang et al., 2016). Accurate estimates of canopy height are crucial for assessing parameters like above-ground biomass, carbon stock, canopy complexity, and habitat quality (Li et al., 2015). Large-scale monitoring of canopy height is especially essential in the protected areas as it helps in understanding the disturbances, deforestation, and degradation dynamics occurring in these areas, serving as a valuable input for policymakers (Gupta & Sharma, 2023; Li et al., 2020). Remote sensing, through platforms such as synthetic aperture radar, satellite and drone oblique photography, and LiDAR/laser scanning, has become a pivotal tool for forest canopy height mapping in recent decades (Lefsky et al., 2005). Among these, LiDAR stands out as the most effective tool for mapping large-scale forest canopy height, owing to its unique capability to directly observe forest canopy structure in the vertical plane (Li et al., 2020). Although terrestrial and airborne LiDAR offer superior detail and accuracy at a small scale, spaceborne LiDAR offers broader coverage and frequent observations at low cost, making it more suitable for large-scale forest monitoring (LaRue et al., 2020).

The GEDI LiDAR instrument, launched in late 2018 and mounted on the International Space Station (ISS), monitors forest ecosystems, providing data on canopy structure and biomass (Dubayah et al., 2020). It measures forest structural characteristics by capturing vertical distributions of forest canopies in waveforms, offering valuable data on surface topography, elevation, canopy height, relative height metrics, plant area index (PAI), and gridded above-ground biomass (Potapov et al., 2021). With a spatial resolution of 25 meters, GEDI data consists of sparsely distributed footprints, which can be supplemented using Machine Learning (ML) models that incorporate optical and SAR remote sensing data for wall-to-wall prediction of canopy height (Jiang et al., 2021).

ML models, widely employed in remote sensing data analysis, offer effective means for continuous mapping of sparsely distributed GEDI canopy height (Wang et al., 2021). Categorized as parametric and non-parametric, they respectively handle linear and complex, nonlinear relationships between variables of interest. Examples of non-parametric ML models include neural networks, support vector machines (SVM), K-nearest neighbors (KNN), boosting models, and random forest (RF) (Jiang et al., 2021). Many prior studies (Bhandari et al., 2023; Gupta & Sharma, 2022; Li et al., 2020; Luo et al., 2023) have explored these models for canopy height prediction with reasonable accuracy. However, the capability of high resolution Indian LISS IV satellite data has yet to be explored as a predictor for mapping forest canopy height.

Therefore, this study seeks to leverage robust ML models (RF, XGB, SVM, and kNN) to predict canopy height by combining GEDI LiDAR and continuous spectral information from high-resolution LISS IV, Sentinel-1 SAR and PALSAR data in the Kaziranga National Park, India. Additionally, it will evaluate canopy height changes between 2018 and 2022 to further understand the dynamics of this critical ecosystem.

2. Study area

Kaziranga National Park, located in the Nagaon and Golaghat districts of Assam, between 26° 30′ N to 26° 45′ N latitude and 93° 08′ E to 93° 36′ E longitude (Figure 1). The region sprawls in an area of 429.93 km^2 across the southern floodplains of the Brahmaputra River, gently sloping from east to west, framed by the foothills and snow-covered peaks of the eastern Himalayas

(Goswami et al., 2019). Its landscape is characterized by riverine lakes, wetlands, grasslands, and elevated woodlands, which offer vital refuge for wildlife, particularly during the annual flooding of the Brahmaputra.

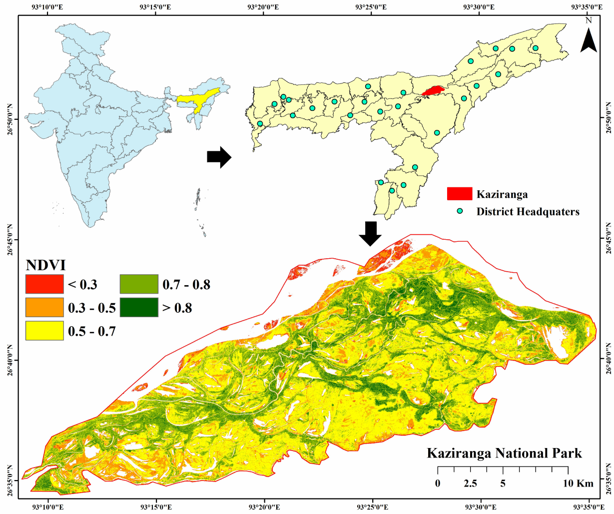

Figure 1. Location of the Study Area

As one of the largest protected areas in the Eastern Himalayan biodiversity hotspot region, Kaziranga is home to a diverse array of endemic and endangered species, including the iconic one-horned Rhinoceros, tigers, elephants, wild buffalos, swamp deer, and wild boars. The park's vegetative ecosystem, comprising grasslands, mixed deciduous, and semi-evergreen forests, owes its productivity and resilience to cyclical flooding, which replenishes nutrients and sustains biodiversity. Designated as a UNESCO World Heritage Site in 1985, Kaziranga remains a beacon of conservation and a testament to the rich natural heritage of the region (Borah et al., 2018).

However, in the contemporary years, Kaziranga has been facing increasing risks from extreme flooding and human activities like encroachment, poaching, and habitat fragmentation. Thereby, monitoring canopy height is of great importance as it provides crucial data about the forest's health, biodiversity, and resilience, guiding conservation efforts in this area.

3. Materials and methods

3.1. Datasets and pre-processing

3.1.1. GEDI LiDAR: NASA’s Global Ecosystem Dynamics Investigation (GEDI), a full-waveform LiDAR sensor deployed on the International Space Station (ISS) since April 2019, captures high-resolution vegetation structure measurements across temperate and tropical forests. Following the ISS orbit trajectory from approximately 52°N to 52°S, GEDI produces 25-meter footprints, resulting in eight ground tracks spaced about 600 meters across and 60 meters along the track. Its Level 2A Geolocated Elevation and Height Metrics Product (GEDI02_A) contains 100 relative height metrics, with rh_98 (the 98th percentile of relative height metrics) utilized in this study. The raster version of the original GEDI02_A product (25m footprint), (LARSE/GEDI/GEDI02_A_002_MONTHLY) has been used for analysis.

As a LiDAR sensor operating in the near-infrared wavelengths, GEDI's data quality is affected by atmospheric conditions. Thus, quality filtering, including removal of low-quality shots (value = 0) based on “quality flag”, degraded shots (value > 0) based on "degrade flag," as well as filtering out shots with low sensitivity (value < 0.95) using the "sensitivity" attribute, is conducted. Temporally, cloud-free LISS IV data is used for filtering GEDI data. Quality and temporal filtering of GEDI data was conducted using the Google Earth Engine (GEE) platform.

3.1.2. LISS IV: Linear Imaging Self Scanning Sensor (LISS IV) data (5.8 m spatial resolution) from the Resourcesat-2 satellite, downloaded from the Bhoonidhi portal of NRSC for 10th December 2022, has been utilized in this study. LISS IV operates in multi-spectral mode, capturing data across three spectral bands: green, red, and near-infrared (NIR), with a swath width of 70 kilometres (Verma et al., 2017). With a 10-bit radiometric quantization, LISS IV provides high radiometric accuracy, translating radiation intensity into 1024 grey levels. However, to optimize this accuracy, atmospheric effects must be corrected. Hence, atmospheric correction has been done using the MODTRAN-based QUAC module within the ENVI software package to derive Surface Reflectance (SR) LISS IV bands. Previous studies have highlighted the effectiveness of QUAC, particularly in vegetated areas, making it a suitable choice for this analysis (Saini et al., 2016).

3.1.3. Sentinel-1 SAR: Sentinel-1, comprising two satellites (Sentinel-1A and Sentinel-1B), operates in a near-polar sun-synchronous orbit at an altitude of 697 kilometres. Equipped with C-band synthetic aperture radar (SAR) sensors, the Sentinel-1 constellation can cover the entire Earth's surface in six days. Operating at a central frequency of 5.405 GHz, Sentinel-1A and Sentinel-1B utilize a dual polarization mode containing either 1 or 2 out of 4 possible polarization bands (VV or HH, and dual-band VV+VH and HH+HV). Sentinel-1 level 1 dual polarization (VV/VH) Ground Range Detected (GRD) scenes, acquired in interferometric wide-swath (IW) mode, were downloaded from the Copernicus Browser for December 2018 and 2022. The scenes were pre-processed using the Sentinel-1 Toolbox in SNAP, resulting in calibrated, terrain-corrected, and speckle-filtered VV and VH backscattering coefficient images. The backscattering coefficients were up-sampled to a 5-meter resolution to align with the pixel size of the LISS IV data, utilizing bilinear interpolation.

3.1.4. PALSAR: The yearly global 25 m PALSAR/PALSAR-2 data available in GEE were downloaded for the years 2018 and 2022. These images are ortho-rectified and slope-corrected. The HV/HH polarization data, provided as 16-bit digital numbers, were converted to gamma nought (γ0) values in decibels (dB) using the following equation:

$$\gamma 0 = 10 \times \log_{10}(DN^2) + CF \qquad (1)$$

where $\gamma 0$ is the backscattering coefficient, DN is the raw pixel value, and CF (-83) is the calibration factor (Ghosh et al., 2022).

3.1.5. SRTM: NASA’s Shuttle Radar Topography Mission (SRTM) data was used in this study to prepare an elevation map, as well as slope and aspect maps. These maps, initially generated at a spatial resolution of 30 meters, were enhanced to a 5-meter resolution through bilinear interpolation to match the other datasets in the study.

3.2. Methods

3.2.1. Satellite Data-Derived Proxies for Input Predictor Variables: A total of 37 predictor variables from LISS IV, SAR, PALSAR and SRTM data were prepared for input into the ML models (Table 1). The predictor variables underwent resampling to a spatial resolution of 5 meters using bilinear interpolation. This adjustment was made to standardize the data resolution and reduce the overall data volume. The canopy height map was similarly generated at a 5-meter resolution.

Data	Predictor Variables
LISS IV	Difference Vegetation Index (DVI), Global Environmental Monitoring Index (GEMI), Green Difference Vegetation Index (GDVI), Green Normalized Difference Vegetation Index (GNDVI), Green Ratio vegetation Index (GRVI), Infrared Percentage Vegetation Index (IPVI), Modified Chlorophyll Absorption ration Index – Improved (MCARI-I), Modified Non-Linear Index (MNLI), Modified Simple Ratio (MSR), Modified Triangular Vegetation Index (MTVI), Modified triangular vegetation Index – Improved (MTVI-I), Non-Linear Index (NLI), Normalized Difference Vegetation Index (NDVI), Optimized Soil Adjusted Vegetation Index (OSAVI), Red Green Ratio Index (RGRI), Renormalized Difference Vegetation Index (RDVI), Simple Ratio (SR), Soil Adjusted Vegetation Index (SAVI), Sum Green Index (SGI), Transformed Difference Vegetation Index (TDVI)
SENTINEL – 1 SAR	VV, VH, VV + VH, VV/VH, Normalized Difference Index (NDI), Radar Vegetation Index (RVI), Modified Radar Vegetation Index (mRVI)
PALSAR	HV, HH, HV + HH (Palsar_SUM), HV/HH (Palsar_Ratio)
SRTM	Elevation, Slope, Aspect

Table 1. Predictor variables for machine learning models derived from diverse data sources

3.2.2. Recursive Feature Elimination (RFE): RFE is an effective feature selection method that ranks and removes the least important features by recursively building models. Starting with all features, RFE iteratively eliminates those with the lowest importance based on model metrics until the optimal subset is found. In this study, RFE was used with Random Forest via the "caret" package in R, employing 10-fold cross-validation to ensure stable and reliable results. This process refined the feature set, ensuring model accuracy and interpretability. The final model, thus, incorporated 17 key variables, viz., Elevation, HV, VH, GRVI, VV, SGI, Sen_SUM, RGRI, B2, Slope, B4, GDVI, Palsar_SUM, GNDVI, HH, Sen_Ratio, and DVI.

3.2.3. Machine learning: The GEDI canopy height data points, distributed sparsely and uniformly, underwent outlier filtering, removing observations with canopy heights exceeding 35 meters or falling below 2.5 meters. Subsequently, a fishnet matching the GEDI pixel size was generated to extract mean pixel values from all predictor rasters. Rows containing NA values for any predictor variable were then eliminated, resulting in the retention of 11,521 observations. These observations were then randomly split into training and testing sets in an 80:20 ratio for input into the machine-learning models (Figure 2).

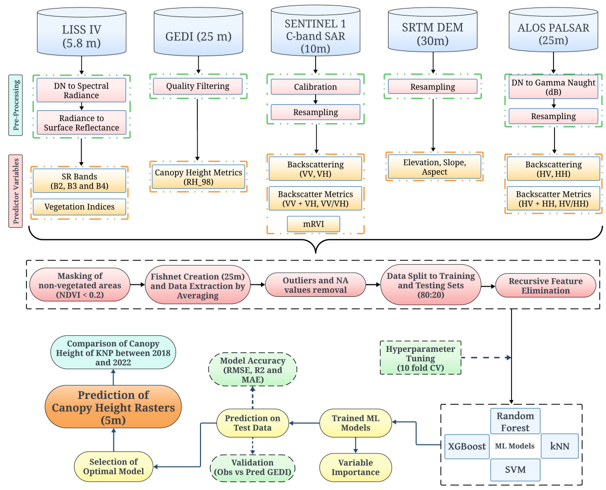

Figure 2. Methodology Flowchart

Four machine learning models were employed in this study utilizing the CARET package in RStudio (Kuhn, 2023), viz., Random Forest (RF), Extreme Gradient Boosting (XGB), Support vector Machine (SVM) and K-Nearest Neighbor (kNN) (Gupta & Sharma, 2022). SVMs stand out for their robustness and accuracy in machine learning, adept at discerning optimal hyperplanes to classify data across various dimensions, whether linear or nonlinear. kNN, while simple in concept, relies on memorized data to classify new points based on proximity, with sensitivity to distance metrics and k values. RF offers a versatile ensemble approach, resilient against overfitting and noise, making them ideal for both classification and regression tasks. Key parameters like randomly selected variables (*mtry*) and tree quantity (*ntree*), enable RFs to excel in handling high-dimensional datasets. XGB emerges as a leading machine learning library, known for its scalability and efficiency in implementing gradient-boosted decision trees. It employs regularization to curb overfitting and often outperforms other algorithms in speed and accuracy.

The model performance was assessed through R-squared (R^2) and Root Mean Square Error (RMSE). The robustness of the model's fit was explained by the Mean Absolute Error (MAE). The following formula were used to calculate these statistics:

$$R^2 = 1 - \left[\frac{\sum_{i=1}^{n}\left(P_{obs} - P_{pred}\right)^2}{\sum_{i=1}^{n}(P_{obs} - \bar{P})^2}\right] \tag{2}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(P_{obs} - P_{pred}\right)^2}{n}} \tag{3}$$

$$MAE = \sum_{i=1}^{n}\left|\frac{P_{obs} - P_{pred}}{n}\right| \tag{4}$$

where P_{obs} is the observed canopy height values, P_{pred} is the predicted canopy height values from the ML models, $\bar{P}$ is the mean observed canopy height values of the selected test samples and n is the sum of all selected test samples. Each model underwent 10-fold cross-validation to fine-tune their hyperparameters. Subsequently, the hyperparameters yielding the lowest RMSE were chosen as the optimal configuration for predicting canopy height.

4. Results

4.1 Canopy Height and Model Accuracy

Among the four machine learning models trained in this study, XGB emerged as the top performer, achieving an RMSE of 5.499 m and an R² of 0.55 in the region (Figure 4). In contrast, kNN had the lowest performance, with an RMSE of 5.73 m. The models were ranked by performance as follows: XGB > Random Forest > Support Vector Machine > kNN for KNP (Table 2).

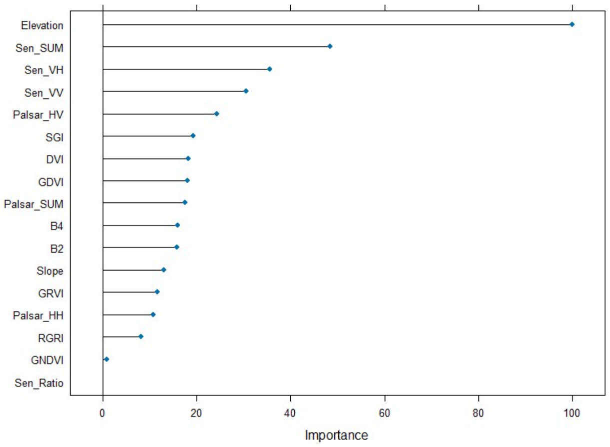

Figure 3. Variable importance for Forest Canopy height prediction in KNP

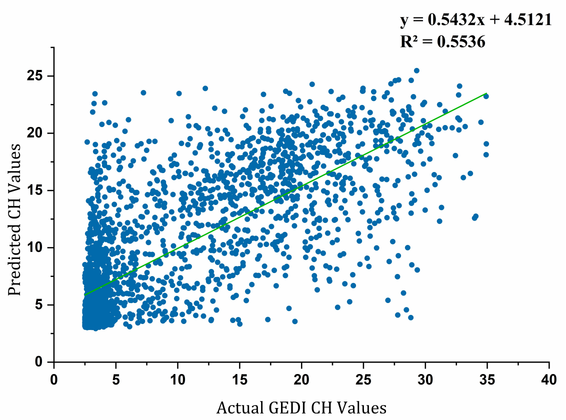

Figure 4. XGB based Actual vs Predicted Canopy Height in KNP

Based on its superior performance, XGB was selected for the wall-to-wall canopy height prediction in Kaziranga National Park for the years 2018 and 2022 in this study.

Tuned Hyperparameters		**RMSE**	**R²**	**MAE**
RF	mtry = 6, ntree = 800	5.499	0.553	3.77
XGB	nrounds = 500, eta = 0.015, max depth = 3, gamma = 0, col sample = 0.6, min child = 2, sub-sample = 0.75	**5.479**	0.557	3.75
SVM	Sigma = 0.03125, C = 0.7894737	5.51	0.559	3.47
kNN	k = 25	5.73	0.515	3.92

Table 2. Performance metrics and tuned hyperparameters of machine learning models for canopy height prediction

Notably, elevation, Sentinel-1 derived VV+VH, VH and VV, along with Palsar HV, emerged as the top 5 significant predictor variables for estimating canopy height according to the random forest based variable importance analysis (Figure 3).

4.2 Spatial Variation of Canopy Height in Kaziranga

KNP, predominantly covered by grasslands, generally exhibits canopy heights under 5 meters (Figures 5 and 6). However, in the southwestern extremity and eastern region of the park, forests feature canopies exceeding 20 meters. Along the riverbanks, canopy heights also surpass 15 meters. The eastern part of the park can be seen to have taller canopies compared to the western and southern areas.

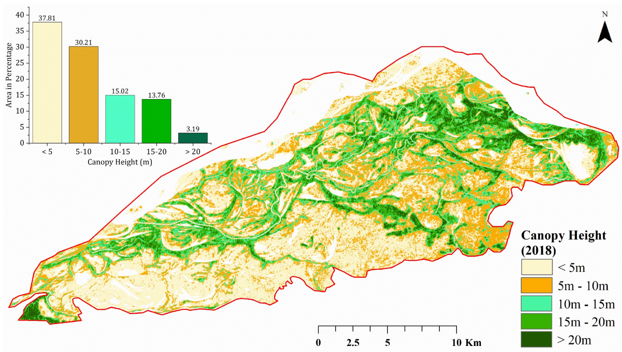

Figure 5. Canopy Height Map of KNP for 2018

In the two years considered for the study, the mean canopy height in Kaziranga National Park was approximately 8.58 ± 5.24 meters in 2018, rising to about 9.07 ± 5.61 meters in 2022. The overall canopy height across the park ranged from 2 meters to 25 meters.

The predicted canopy height in KNP has been categorized into five classes, as shown in Table 3. In both years, the largest portion of the park had a canopy height of less than 5 meters, covering 130.54 km² in 2018 and expanding slightly to 135.43 km² in 2022. Trees with heights ranging from 5 to 10 meters were more prevalent in 2018, spanning 104.29 km², but this decreased to 79.73 km² in 2022. Notably, the area with canopy heights exceeding 10 meters grew significantly, from 110.38 km² in 2018 to 130.04 km² in 2022. Thus, it can be seen that significant changes have taken place in canopy height across KNP between 2018 and 2022, with a notable increase in areas with taller canopies.

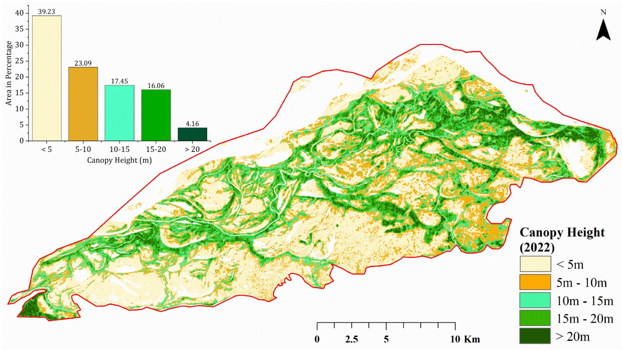

Figure 6. Canopy Height Map of KNP for 2022

Canopy Height (m)	2018	2022
< 5	130.54	135.43
5-10	104.29	79.73
10-15	51.86	60.23
15-20	47.50	55.45
> 20	11.02	14.36

Table 3. Area in Percentage (%) in different Canopy height classes as predicted by different machine learning models

5. Discussion

This study examines the integration of spaceborne GEDI LiDAR data with LISS IV optical imagery, Sentinel-1 SAR, and PALSAR data to improve the precision of high-resolution canopy height models and achieve comprehensive, wall-to-wall mapping of forest canopy height in Kaziranga National Park, Assam. Employing four distinct machine learning models, the research aims to enhance the accuracy of canopy height predictions. The results reveal that the models' predictions of canopy height fall within an acceptable accuracy range. Spaceborne LiDAR's moderate prediction accuracy in this study could stem from factors like uneven terrain, diverse species, and a lack of complex predictors, hindering precise canopy height estimation. Similar results were obtained in the study done by (Gupta & Sharma, 2022) in a mixed tropical forest of Gujarat. Also, the prediction accuracies are comparable with the study done by (Ghosh et al., 2022) in the forested areas of India. XGB emerges as the top-performing model in this study, attaining the highest accuracy metrics with an R-squared value of 0.557 and an RMSE of 5.479, followed closely by RF, SVM, and kNN.

The analysis of canopy height distributions in Kaziranga National Park reveals that nearly 40% of the area has a canopy height below 5 meters, largely due to the park's alluvial grasslands, which are maintained by annual flooding and burning, allowing it to support many threatened species. However, a notable shift has occurred between 2018 and 2022, with a significant increase in taller canopies exceeding 10 meters, while the area covered by vegetation between 5 to 10 meters has decreased. This pattern reflects the region's flood dynamics, where the heavy floods of 2017 led to primary succession, resulting in shorter canopies in 2018. While, the absence of major floods in 2021 and 2022 allowed vegetation to regrow, increasing canopy heights above 10 meters (Karmakar, 2024).

The results provide valuable insights into the region's vegetation resilience and regrowth in response to its annual flood cycles. Moreover, the incorporation of LISS IV data in this study marks a significant advancement in high-resolution canopy height mapping, especially within complex ecosystems like Kaziranga National Park. Despite being used for the first time in this context, LISS IV imagery demonstrated substantial potential in enhancing the accuracy of canopy height models when integrated with other satellite data sources, such as GEDI LiDAR and SAR. The study underscores the importance of utilizing multi-source data to capture finer details of forest structure, which is crucial for informed conservation efforts and better management of protected areas.

6. Conclusion

This study highlights the potential of integrating GEDI LiDAR data with optical (LISS IV) and microwave (SAR) imagery through machine learning for continuous forest canopy height mapping in Kaziranga National Park. By training and evaluating four machine learning models, the research emphasizes the importance of LISS IV and SAR data, with elevation and SAR variables being identified as key predictors. XGB emerged as the most accurate model. The analysis revealed that KNP is predominantly covered by grasslands with canopy heights below 5 meters, while taller canopies exceeding 20 meters are rare.

The study provides valuable insights into vegetation response to flood dynamics, offering a foundation for improved forest management and biodiversity conservation in KNP and similar ecosystems. As the first comprehensive attempt to map canopy height in KNP, these findings are critical for assessing above-ground biomass, CO^2 sequestration, and mitigating forest fire risks, contributing to better conservation strategies and forest loss prevention.

Acknowledgements

The authors thank NASA for providing the freely available GEDI and SRTM data. NRSC is also acknowledged for openly providing the LISS IV dataset. Furthermore, the authors would like to appreciate ESA and JAXA for freely providing the Sentinel and PALSAR datasets. The first author thanks the DST-INSPIRE program (Reg. No. – IF190591), for providing the required funding for conducting this research through the Sanction Order number (DST/INSPIRE Fellowship/2019/IF190591).

References

Bhandari, K., Srinet, R., Nandy, S., 2023. Forest Height and Aboveground Biomass Mapping by synergistic use of GEDI and Sentinel Data using Random Forest Algorithm in the Indian Himalayan Region. *Journal of the Indian Society of Remote Sensing*. https://doi.org/10.1007/s12524-023-01792-z

Borah, S. B., Sivasankar, T., Ramya, M. N. S., Raju, P. L. N., 2018. Flood inundation mapping and monitoring in Kaziranga National Park, Assam using Sentinel-1 SAR data. *Environmental Monitoring and Assessment*, *190*(9). https://doi.org/10.1007/s10661-018-6893-y

Dubayah, R., Blair, J. B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P. L., Qi, W., Silva, C. 2020. The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth’s forests and topography. *Science of Remote Sensing*, *1*. https://doi.org/10.1016/j.srs.2020.100002

Ghosh, S. M., Behera, M. D., Kumar, S., Das, P., Prakash, A. J., Bhaskaran, P. K., Roy, P. S., Barik, S. K., Jeganathan, C., Srivastava, P. K., Behera, S. K., 2022. Predicting the Forest Canopy Height from LiDAR and Multi-Sensor Data Using Machine Learning over India. *Remote Sensing*, 14(23). https://doi.org/10.3390/rs14235968

Goswami, V. R., Yadava, M. K., Vasudev, D., Prasad, P. K., Sharma, P., Jathanna, D., 2019. Towards a reliable assessment of Asian elephant population parameters: the application of photographic spatial capture–recapture sampling in a priority floodplain ecosystem. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-44795-y

Gupta, R., Sharma, L. K., 2022. Mixed tropical forests canopy height mapping from spaceborne LiDAR GEDI and multisensor imagery using machine learning models. *Remote Sensing Applications: Society and Environment*, *27*. https://doi.org/10.1016/j.rsase.2022.100817

Gupta, R., Sharma, L. K., 2023. Mapping Canopy Height from ICESat-2 and Landsat-9 using Machine Learning in the Himalayan Corbett Tiger Reserve, India. *2023 International Conference on Machine Intelligence for GeoAnalytics and Remote Sensing, MIGARS 2023*. https://doi.org/10.1109/MIGARS57353.2023.10064540

Jiang, F., Zhao, F., Ma, K., Li, D., Sun, H., 2021. Mapping the Forest Canopy Height in Northern China by Synergizing ICESat-2 with Sentinel-2 Using a Stacking Algorithm. *Remote Sensing*, *13*(8), 1535. https://doi.org/10.3390/rs13081535

Karmakar, S., 2024. Assam floods: Kaziranga facing worst deluge in decade, 174 animal deaths so far. *The Deccan Herald*. https://www.deccanherald.com/india/assam/assam-floods-kaziranga-facing-worst-deluge-in-decade-174-animal-deaths-so-far-3102048

Kuhn, M., 2023. *Classification and Regression Training [R package caret version 6.0-94]* (6.0-94). Comprehensive R Archive Network (CRAN). https://CRAN.R-project.org/package=caret

LaRue, E. A., Wagner, F. W., Fei, S., Atkins, J. W., Fahey, R. T., Gough, C. M., Hardiman, B. S., 2020. Compatibility of aerial and terrestrial LiDAR for quantifying forest structural diversity. *Remote Sensing*, *12*(9). https://doi.org/10.3390/RS12091407

Lefsky, M. A., Harding, D. J., Keller, M., Cohen, W. B., Carabajal, C. C., Del Bom Espirito-Santo, F., Hunter, M. O., de Oliveira, R., 2005. Estimates of forest canopy height and aboveground biomass using ICESat. *Geophysical Research Letters*, *32*(22), 1–4. https://doi.org/10.1029/2005GL023971

Li, W., Niu, Z., Liang, X., Li, Z., Huang, N., Gao, S., Wang, C., Muhammad, S., 2015. Geostatistical modeling using LiDAR-derived prior knowledge with SPOT-6 data to estimate temperate forest canopy cover and above-ground biomass via stratified random sampling. *International Journal of Applied Earth Observation and Geoinformation*, *41*, 88–98. https://doi.org/10.1016/j.jag.2015.04.020

Li, W., Niu, Z., Shang, R., Qin, Y., Wang, L., Chen, H., 2020. High-resolution mapping of forest canopy height using machine learning by coupling ICESat-2 LiDAR with Sentinel-1, Sentinel-2 and Landsat-8 data. *International Journal of Applied Earth Observation and Geoinformation*, *92*. https://doi.org/10.1016/j.jag.2020.102163

Luo, Y., Qi, S., Liao, K., Zhang, S., Hu, B., Tian, Y., 2023. Mapping the Forest Height by Fusion of ICESat-2 and Multi-Source Remote Sensing Imagery and Topographic Information: A Case Study in Jiangxi Province, China. *Forests*, *14*(3). https://doi.org/10.3390/f14030454

Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M. C., Kommareddy, A., Pickens, A., Turubanova, S., Tang, H., Silva, C. E., Armston, J., Dubayah, R., Blair, J. B., Hofton, M., 2021. Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote Sensing of Environment*, *253*, 112165. https://doi.org/10.1016/j.rse.2020.112165

Saini, V., Tiwari, R. K., Gupta, R. P., 2016. Comparison of FLAASH and QUAC atmospheric correction methods for Resourcesat-2 LISS-IV data. *Earth Observing Missions and Sensors: Development, Implementation, and Characterization IV*, *9881*, 98811V. https://doi.org/10.1117/12.2228097

Verma, A. K., Garg, P. K., Hari Prasad, K. S., 2017. Sugarcane crop identification from LISS IV data using ISODATA, MLC, and indices based decision tree approach. *Arabian Journal of Geosciences*, *10*(1). https://doi.org/10.1007/s12517-016-2815-x

Wang, H., Seaborn, T., Wang, Z., Caudill, C. C., Link, T. E., 2021. Modeling tree canopy height using machine learning over mixed vegetation landscapes. *International Journal of Applied Earth Observation and Geoinformation*, *101*, 102353. https://doi.org/10.1016/j.jag.2021.102353

Zhang, J., Nielsen, S. E., Mao, L., Chen, S., Svenning, J., 2016. Regional and historical factors supplement current climate in shaping global forest canopy height. *Journal of Ecology*, *104*(2), 469–478. https://doi.org/10.1111/1365-2745.12510