WRI INDIA

TECHNICAL NOTE

Measuring and monitoring tree cover and plant canopy height in Pune city, India

Pulakesh Das, Sayantan Dey, Kanchana Balasubramanian, and Parth Sarathi Roy

CONTENTS

Technical notes document the research or analytical methodology underpinning a publication, interactive application, or tool.

Suggested Citation: P, Das, S, Dey, K. Balasubramanian, and P. S. Roy 2023. "Measuring and monitoring tree cover and plant canopy height in Pune city, India." Technical Note. Washington, DC: World Resources Institute. Available online at : https://doi.org/10.46830/writn.21.00111

ABSTRACT

Trees in urban landscapes (also termed *trees outside forests* [TOF]) are perennial woody plants designed to sustain biodiversity, improve environmental quality, beautify the landscape, mitigate urban heating, provide shade, and reduce pollution. In India, the Forest Survey of India (FSI) conducts the national assessment of TOF in urban areas by employing extensive ground data. This technical note highlights the potential of the latest remote sensing data and machine learning techniques to rapidly monitor TOF and their structural attributes with limited ground data and satellite-data-derived products. The current study was conducted in the Pune Municipal Corporation (PMC) area. Publicly available Sentinel-1 microwave and Sentinel-2 optical data were used to predict tree height using Global Ecosystem Dynamics Investigation (GEDI) Light Detection and Ranging (LiDAR) data as the ground truth. The Random Forest (RF) machine learning model was applied for image classification and regression analysis. Sentinel-2 optical data were also used for land use land cover (LULC) change mapping, which shows 620 ha of tree cover loss from 2016–17 to 2019–20. The regression analysis indicated reliable tree height estimates (R^2= 0.74 and RMSE = 2.85 m). This study, with an acceptable accuracy level for many city-level uses, represents a reliable methodology for rapid TOF change and canopy height assessment using publicly available data, which can be useful for city planning in Pune. The methodology used in the current study can be scaled up to other urban landscapes.

INTRODUCTION

Trees outside forests (TOF) are trees in cropland, urban areas, or other lands outside the Recorded Forest Area (RFA) boundary (FSI 2019). TOF consist of trees in block plantations (patches of tree cover > 0.1 ha), scattered in farmland and home gardens, and in linear plantations along boundaries, roads, streams, community lands, and so on. Increasing TOF can help meet India's commitments, such as restoring degraded land, its Nationally Determined Contribution, the goals of the sub-mission on agroforestry, and net zero carbon emissions by 2070. Similarly to forests, TOF provide various social, economic, environmental, and ecological benefits, such as soil enrichment, provisioning of fuelwood and fodder, non-timber forest products for local communities, and mitigating on of climate change (Jose 2009). TOF constitute a significant proportion of tree cover in urban areas and provide essential ecosystem services, such as sequestering carbon, improving air quality, regulating urban hydrology and micro-climate conditions, and enhancing recreational and cultural values (Bolund and Hunhammar 1999; Dobbs et al. 2011; Dwyer et al. 1992). The latest Intergovernmental Panel on Climate Change (IPCC) Sixth Assessment Report (AR6) has reiterated that tree cover loss and urbanization cause heatwave generation and regional temperature rise, creating the urban heat island effect. Assessing tree cover dynamics and their interaction with the other LULC categories is essential for highlighting the ongoing land-use planning activities and prescribing climate change adaptation measures. WRI India's Restoration Opportunities Atlas estimates a restoration potential of more than 80 million ha (Mha) in India through mosaic restoration, which primarily involves the integration of TOF in different land uses, including cropland and other lands (Chaturvedi et al. 2018).

Effective urban planning in accordance with green infrastructure development requires suitable data on existing green spaces and their changes in recent years. In India, the FSI conducts the national assessment of TOF. The FSI classifies tree cover patches with an area ≥1 ha as forest, and tree cover patches found outside the RFA with an area <1 ha are classified as TOF (FSI 2019). The total TOF recorded in India in 2019 covered 29.38 million ha (36.40 percent of the total forest and tree cover area) (FSI 2019). The FSI follows a grid-based inventory approach to mapping TOF and estimating their growing stock, in which sampling is carried out in selected grids for a particular year. The FSI digitally classifies Sentinel-2 multispectral data, followed by manual editing and refinement for TOF mapping in rural areas. In comparison, the tree cover mapping scheme used in urban areas entirely relies on ground-based observation. The FSI uses Urban Frame Survey blocks (comprising areas of 600 to 800 population or 120–160 households) with well-defined boundaries provided by the National Sample Survey Organization as the sampling frame for urban TOF assessment. The ground data on tree count, crown cover, and Culturable Non-Forest Area (CNFA) of the selected grids are used to estimate the total TOF area and growing stock in urban areas (FSI 2019).

Multi-temporal systematic tree cover mapping and change monitoring are useful for several reasons, such as assessing forest cover resources, biodiversity, and various ecosystem services, including carbon sequestration potential, contribution to regulating the hydrological cycle, micro-climatic conditions, and mitigation of urban heat islands; estimating tree cover loss and gain; and urban planning for improving green cover spaces. Tree canopy height estimation is essential for assessing the ecosystem structure and growth, estimating the aboveground plant biomass and carbon stock, and so on. Although several public national and global datasets are available, they have regional and local biases and are generally unsuitable for urban green space assessment and planning within a city (Ghosh et al. 2022; Nandy et al. 2021).

Advances in satellite remote sensing and data processing methods and platforms allow tree cover to be monitored and canopy height to be estimated. Although tree cover mapping and change assessment using remote sensing data are well established, tree canopy height estimation is challenging owing to the lack of suitable satellite data. At the same time, ground-based assessments are costly, labor intensive, and time consuming. The latest remote sensing data and advanced data processing approaches are employed for rapid tree cover mapping and structural attribute characterization. The publicly available Sentinel-1 C-band synthetic aperture radar (SAR) data and Sentinel-2 optical data enable tree height estimation at finer spatial and temporal scales (Ghosh et al. 2020; Fagua et al. 2019). The SAR backscatter integrated with optical remote sensing data enables improved tree stand height estimation using regression-based techniques (Roy et al. 2021; Lee and Lee 2018). Light Detection and Ranging (LiDAR) data enable tree height estimation at the plot level (representing circular plots) and is used as a surrogate for field observations (Potapov et al. 2021). Many studies have used the Global Ecosystem Dynamics Investigation (GEDI) LiDAR tree height data in combination with various multispectral and microwave data for tree canopy height estimation (Qi et al. 2019; Potapov et al. 2021). Several machine learning techniques are used in image classification and regression, among which the Random Forest (RF) has shown high accuracy in estimating tree canopy height by integrating Sentinel-1 SAR backscatter and Sentinel-2 data and derived proxies (Ghosh et al. 2020; Roy et al. 2021). Many studies have estimated tree canopy height from a landscape level to a global scale (Potapov et al. 2021). However, satellite-data-based canopy height estimation in urban landscapes is scarce, especially in India. The current study employs Sentinel-2 optical data for urban green space monitoring from 2016 to 2020 in the Pune Municipal Corporation (PMC) area, Maharashtra, India. Moreover, the study uses GEDI LiDAR data for tree height estimation by employing Sentinel-1 (microwave) and Sentinel-2 (optical) data as determinant variables.

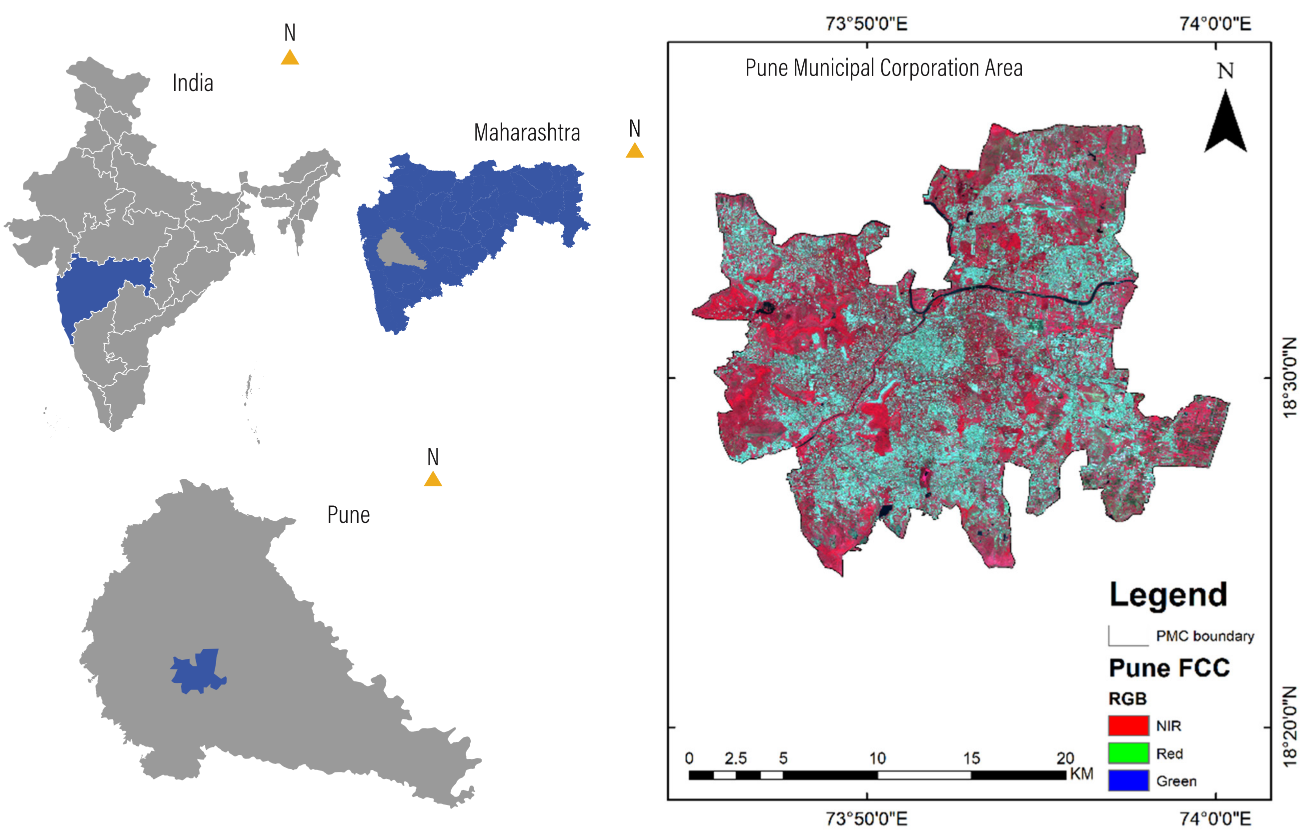
OBJECTIVES

- Monitor and summarize change patterns between tree cover and other LULC classes in the period 2016–17 to 2019–20 within the PMC boundaries.
- Estimate tree canopy height with Sentinel-1 and Sentinel-2 data in the PMC area for 2019–20 using machine learning techniques, with GEDI LiDAR data as the ground truth.

STUDY AREA

The study was conducted in Pune, the second-largest metropolitan city in Maharashtra, India (see Figure 1). The PMC area is 35,637 ha, and the average altitude is 560 m above the mean sea level. Pune, a tropical city situated in the Western Ghats region of India, is classified under the tropical climate zone. It is one of India's fastest-growing cities and has shown an overall increase in built-up area from 11,660 ha in 1990 to 16,690 ha in 2019 at the expense of cropland, scrubland, fallow land, and so on (Gohain et al. 2021). According to the 2011 Census of India (Census of India 2011), the population of Pune city is above 3 million. Maps of the PMC and its administrative boundaries were downloaded from the PMC GIS portal (PMC n.d.a). The deciduous broadleaf forest is the dominant forest type, which sheds leaves (leaf-fall period) from January to February.

Figure 1 | **Location map of the PMC and post-monsoon SFCC**



Notes: NIR = near infrared; PMC = Pune Municipality Corporation; SFCC = Standard False Color Composite.

Source: WRI India authors and PMC (n.d.a).

DATA AND METHODOLOGY

Publicly available optical Sentinel-2 satellite data were classified for LULC and change identification between two time periods (2016–17 and 2019–20). Sentinel-2 optical and Sentinel-1 microwave satellite data were also used as proxy variables to generate a wall-to-wall canopy height map using GEDI canopy height data as a reference. The tree canopy height was estimated for the tree cover area identified in the 2019–20 LULC map. The RF machine learning model was employed for the LULC classification and regression analysis for the tree canopy height estimation. The input data and methodology used in the study are described in the following sections.

Data used

Satellite data

LULC MAPPING

Multi-temporal Sentinel-2 multispectral images were used for LULC mapping. Sentinel-2 top of atmosphere (TOA) reflectance data (10 m and 20 m bands) were accessed from the Google Earth Engine (GEE) platform for three seasons (post-monsoon, leaf-fall, and driest season) for two periods: 2016–17 and 2019–20. The average TOA reflectance of all the cloud-free images in a season was used. Cloud-free images available from November to January were used to compute the average TOA reflectance image of 2016 after the monsoon; similarly, cloud-free images from February to April were used for the leaf-fall period, and those from May to June were used for the driest season.

CANOPY HEIGHT MAPPING

This technical note aims to present how well a combination of Sentinel-2 (optical) and Sentinel-1 (radar) data can estimate tree canopy height.

GEDI LiDAR tree height data were used as the ground truth for the height estimation. The GEDI LiDAR sensor provides the canopy height metric at a footprint level (circular plot) with a 25 m diameter. The distance between the two footprints is 60 m in the along-track direction and 600 m in the across-track direction. GEDI data were downloaded from the Earthdata Search portal (Earthdata n.d.). GEDI data points for the study area, collected from November 2019 to June 2020, were used in the current study. GEDI data represent the maximum tree canopy height within a circular plot (with a footprint of radius 25 m).

In addition to using the 2019–20 Sentinel-2 optical data for LULC classification, these data were also used to generate predictor variables for tree height estimation. The seasonal (post-monsoon, leaf-fall, and driest seasons) mean TOA reflectance bands were used in addition to the four vegetation indices derived from these average reflectance bands: Enhanced Vegetation Index (EVI), Fraction of Vegetation Cover (FVC), Leaf Area Index (LAI), and Normalized Difference Vegetation

Table 1 | **Data used in the study**

DATA	RESOLUTION	YEAR	SOURCE
Sentinel-2	10 m and 20 m	2016–17: Post-monsoon (November–January)	Google Earth Engine
		2017: Leaf-fall period (February–April), driest period (May–June)	
		2019–20: Post-monsoon (November–January)	
		2020: Leaf-fall period (February–April), driest season (May–June)	
Sentinel-1	20 m	November 2019–June 2020	
GEDI	Vector data (25 m footprint data)	November 2019–June 2020	Earth Data Search
SRTM DEM	30 m	2011	Google Earth Engine

Notes: GEDI = Global Ecosystem Dynamics Investigation; SRTM DEM = Shuttle Radar Topographic Mission-Digital Elevation Model.

Sources: WRI India authors.

Index (NDVI). In addition, the Sentinel-1 backscatter ($\sigma 0$) values in the vertical-vertical (VV) and vertical-horizontal (VH) polarization bands were used as input variables for tree height estimation. The pre-processed Sentinel-1 data were accessed from the GEE platform. The mean seasonal (post-monsoon, leaf-fall, and driest season) backscatter values were employed as additional input variables for tree height estimation. The Shuttle Radar Topographic Mission (SRTM)-derived Digital Elevation Model (DEM) was employed to derive topographic variables, which were also used as an input variable.

Methodology

LULC mapping

The RF machine learning algorithm was employed for LULC classifications performed in 2016–17 and 2019–20. A minimum of 100 reference data points for each LULC class (bare land, built-up, cropland, grassland, tree cover, and waterbodies) were generated using visual image interpretation of Sentinel-2 and high-resolution Google Earth imagery. The same reference data points (taken from their respective image dates) were used for training in image classification for 2016–17 and 2019–20 to avoid biases in training and data validation. This was achieved by carefully creating the training data points based on the multi-temporal satellite images and Google Earth images. Only those reference points were considered where no LULC change was observed in the images between 2016–17 and 2019–20.

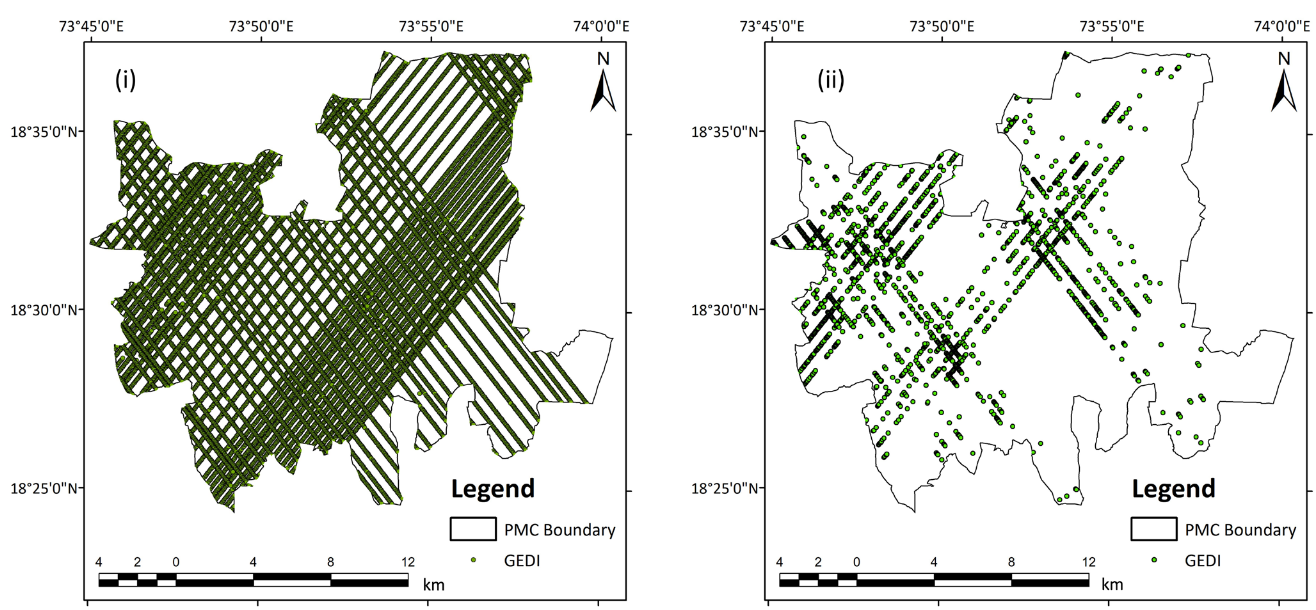
Tree height estimation

The tree canopy height was estimated using a regression-based approach. GEDI canopy height data were used as the observed or dependent variable. Among the various relative height (RH) indicators, RH95 of GEDI data was adopted here because it is considered the most accurate metric (GEDI n.d.; Potapov et al. 2020). The total number of available GEDI data points for 2019 and 2020 was 42,568. The GEDI data points were overlaid on the LULC map of 2019–20. We found 3,570 GEDI data points overlapping with the identified tree cover class in the LULC map of 2019–20 and considered them for further analysis (Figure 2). The predictor variables included seasonal mean leaf-off (January–February), pre-monsoon (March–May), and post-monsoon (October–December) Sentinel-1 backscatter images (VV and VH bands); seasonal mean Sentinel-2 data TOA reflectance bands (10 m and 20 m) and the derived vegetation indices (EVI, FVC, LAI, and NDVI); and the DEM (their relative importance in prediction in the final model is shown in Appendices A and D). The NDVI and EVI were calculated using Equations 1 and 2, in which the FVC and LAI were estimated using the SNAP tool (ESA n.d.) and employing Sentinel-2 TOA reflectance bands.

$$\mathbf{NDVI = (NIR\text{-}Red)/(NIR+Red) \ldots (Eq.\ 1)}$$

$$\mathbf{EVI = 2.5^*(NIR\text{-}Red)/(NIR+6^*Red\text{-}7.5^*Blue+1) \ldots (Eq.\ 2)}$$

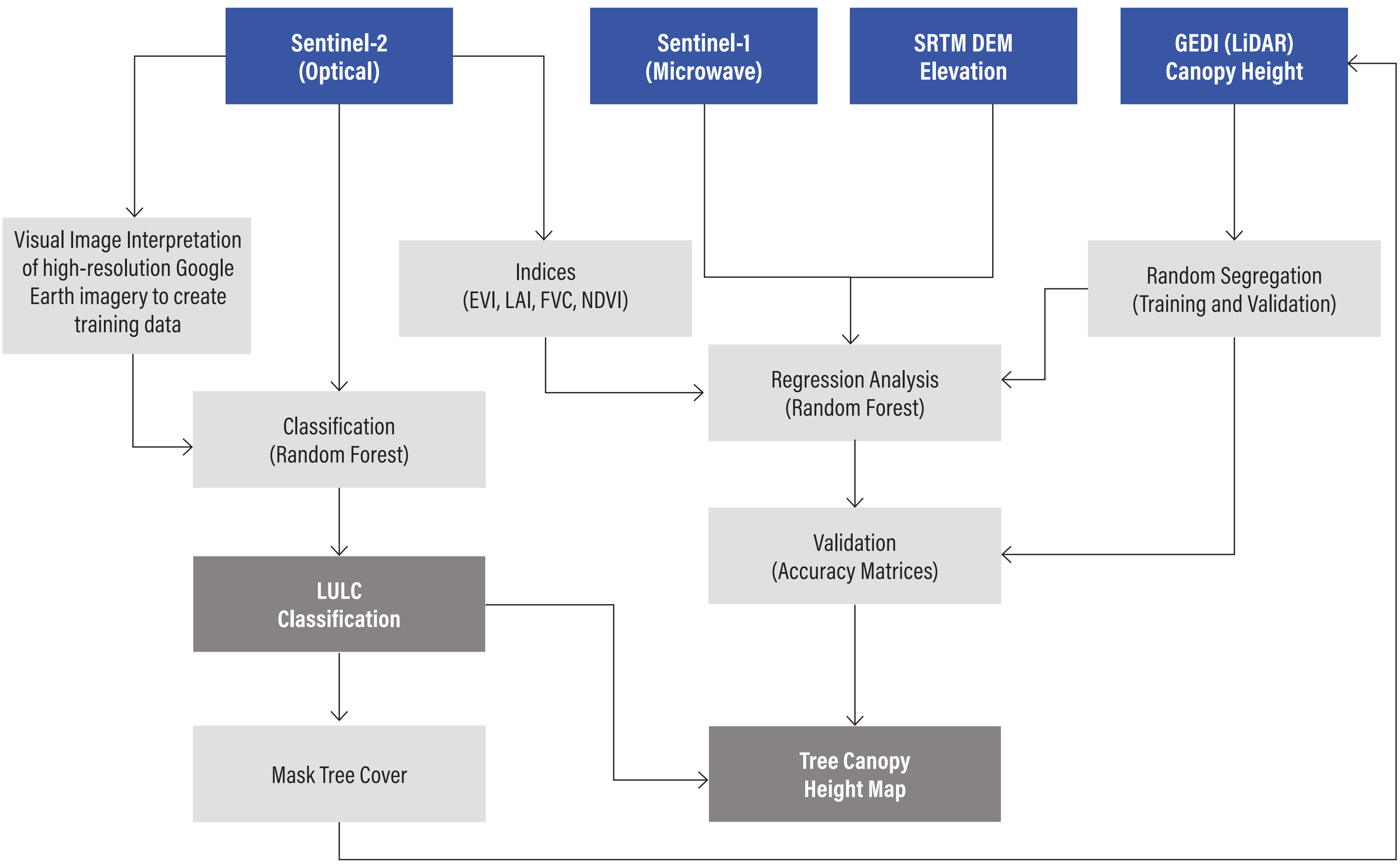
Figure 2 | **The GEDI data points: (i) original and (ii) 3,570 points (overlapped with tree cover) selected for canopy height estimation**



Notes: NIR = GEDI = Global Ecosystem Dynamics Investigation; km = kilometers; PMC = Pune Municipal Corporation.
Source: Earthdata n.d.

The RF machine learning algorithm was applied for regression analysis, in which a randomly selected 70 percent of the reference data was used for model building or training and the remaining 30 percent for validation (see Figures A1 and A2 in Appendix A). Two important input model tuning parameters are used in RF regression: mtry and ntree. The parameter mtry denotes the number of independent variables sampled to each predictor tree, and ntree denotes the number of regression trees grown. The maximum number of trees (ntree) was fixed as 1,000, and the model developed with all the predictor variables (mtry = 6) indicated the best accuracy (the lowest error). The developed RF model was then used to create tree canopy height maps of the study site based on the identified tree cover in the LULC map. The overall methodological flow diagram used in LULC mapping and tree canopy height estimation in the current study is shown in Figure 3.

Figure 3 | **Overall methodological flow diagram used in LULC mapping and tree height estimation**



Notes: EVI = Enhanced Vegetation Index; FVC = Fraction of Vegetation Cover; GEDI = Global Ecosystem Dynamics Investigation; LAI = Leaf Area Index; LiDAR = Light Detection and Ranging; LULC = land use land cover; NDVI = Normalized Difference Vegetation Index; SRTM DEM = Shuttle Radar Topographic Mission-Digital Elevation Model.

Source: WRI India authors.

RESULTS

Assessment of LULC changes with a focus on the tree cover class

The RF machine learning model achieved an overall classification accuracy of 89.22 percent (Kappa: 0.86) for 2016–17 and 92.65 percent (Kappa: 0.91) for 2019–20. A few misclassifications were found between a few classes, such as grassland, tree cover, cropland, and bare land (see Tables B1 and B2 in Appendix B). The lowest User's Accuracy was obtained for grassland (~73 percent), whereas it was more than 87 percent for the rest of the classes in 2016–17 and 2019–20. The lowest Producer's Accuracy in 2016–17 was obtained for grassland (76 percent), followed by bare land (76 percent), which was seen for bare land and cropland (79 percent) in 2019–20.

The built-up area was found to be the dominant class, covering 35.91 percent of the total study area, followed by tree cover (30.54 percent) and grassland (16.56 percent) in 2016–17 (see Figure 4 and Table 2). The maximum increase in area from 2016–17 to 2019–20 was observed for the built-up area (2.35 percent; 839.09 ha), followed by cropland (0.84 percent; 298.96 ha) (see Table 2). On the contrary, a decrease in the area was observed for tree cover (1.74 percent; 619.94 ha), and grassland (1.39 percent; 497.01 ha), and a minor decrease was observed for bare land (0.1 percent; 34.45 ha) (see Table 2 and Figure 4). In comparison, a minor increase in waterbody area (13.18 ha; 0.04 percent) was seen during the study period. Most of the observed built-up area expansion replaced tree cover, grassland, and bare land. A few examples of various LULC changes are shown, which indicated tree cover and grassland loss due to various developmental activities and tree cover management (see Appendix C). In Appendix C, Figure C1 shows the conversion of cropland, grassland, and tree cover to built-up areas (construction of Metro Range Hill Depot), Figure C2 shows the conversion of bare land to built-up areas (between Bharat Mata Road and Tingre Park Road), and Figure C3 shows tree cover loss on the Mhatoba Tekdi hill.

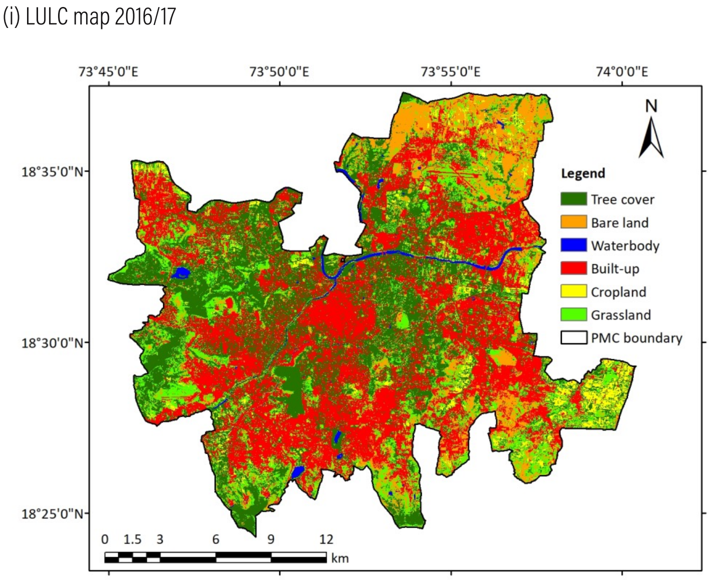
Table 2 | **Area in hectares (and percentage change) of LULC classes in 2016-17 and 2019-20**

LULC CLASS	2016-17	2019-20	CHANGE
Bare Land	3,985.98 (11.18)	3,951.53 (11.09)	–34.45 (–0.1)
Built-up	12,799.01 (35.91)	13,638.1 (38.27)	839.09 (2.35)
Cropland	1,660.13 (4.66)	1,959.09 (5.5)	298.96 (0.84)
Grassland	5,901.12 (16.56)	5,404.11 (15.16)	–497.01 (–1.39)
Tree Cover	10,882.72 (30.54)	10,262.78 (28.8)	–619.94 (–1.74)
Waterbody	408.66 (1.15)	421.84 (1.18)	13.18 (0.04)

Notes: LULC = land use land cover.

Sources: WRI India authors.

Figure 4 | **LULC maps of PMC: (i) 2016-17 and (ii) 2019-20**



Notes: LULC = land use land cover; PMC = Pune Municipal Corporation.

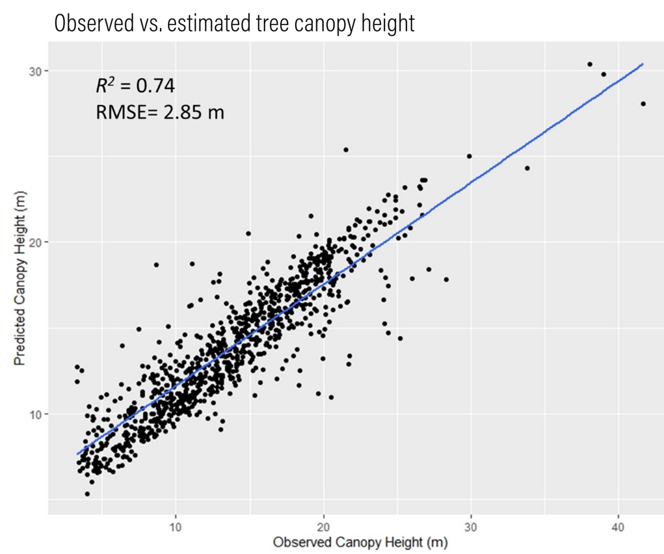
Source: WRI India authors.

Tree height estimation

The RF regression model was built using 70 percent of the total GEDI data points overlapping with the tree cover. A comparison with the validation data (30 percent of the total observations) indicated reliable accuracy (R^2 = 0.74; RMSE = 2.85 m) (Figure 5). The variable importance plot identified the relative importance of various input or determinant variables in canopy height estimation (see Figure D1 in Appendix D). The highest contribution was identified for TOA reflectance in the NIR bands (Sentinel-2 bands 8 and 8A), followed by the Red Edge spectral band (B5) and VV backscatter of the leaf-off season. Moderate importance was recorded for the EVI, microwave backscatter values, and elevation; lower importance was recorded for the rest of the spectral bands, vegetation indices, and microwave backscatter values.

The GEDI data-derived observed canopy height ranges between 1.98 m and 51.8 m. However, 90 percent and 99 percent of the total observations indicated a canopy height ≤ 18 m and ≤ 24 m, respectively (see Figure E1 in Appendix E). The estimated canopy height values varied between 2.75 m and 42.78 m, with a mean of 11.57 m (see Figure 6(i)). The result indicated that the estimated canopy height for most of the trees (98.92 percent) varied between 6 m and 18 m, whereas trees with a height < 6 m and > 18 m represent less than 0.11 percent and 0.97 percent of the total tree cover, respectively (see Figure 6(ii)). The spatial trend indicated a lower canopy height in larger tree cover patches than in scattered tree cover patches.

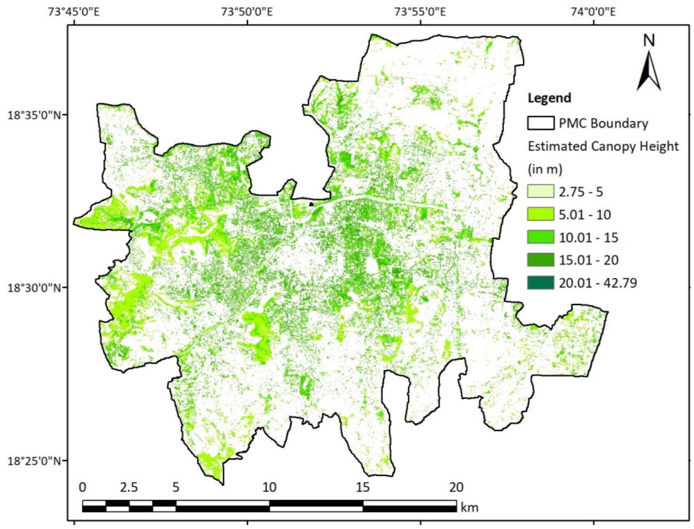
Figure 5 | **Accuracy (R^2 and RMSE) in estimating tree canopy height map**

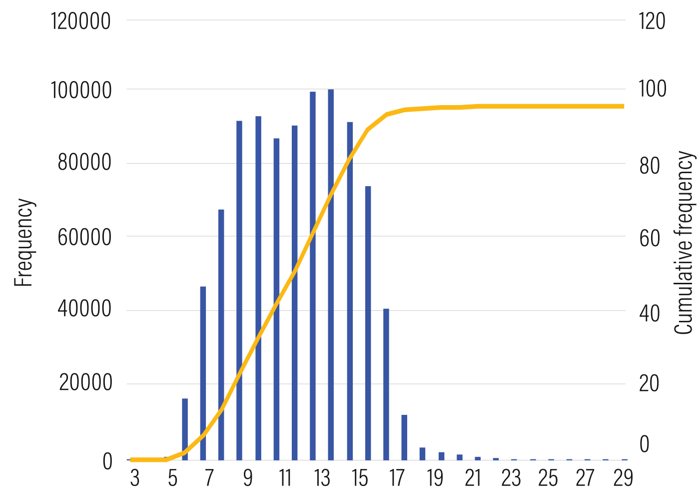


Notes: RMSE = root mean square error.

Source: WRI India authors.

Figure 6 | **(i) Estimated tree canopy height map and (ii) canopy height distribution**





Source: WRI India authors.

DISCUSSION

Previous studies have mostly employed moderate-resolution Landsat data (30 m) for urban green space monitoring at various temporal scales. However, the Sentinel-2 data with four important spectral bands at 10 m resolution enable more precise land resource mapping in heterogeneous landscapes, including urban areas. The RF machine learning algorithm was used for LULC mapping using multi-temporal optical data. The high classification accuracy obtained in the current study demonstrates the efficacy of the RF machine learning algorithm in LULC classification (Das and Pandey 2019; Talukdar et al. 2020; Mishra et al. 2021; Shetty et al. 2021). The tree cover area identified in the current study was compared with the latest Esri global LULC map generated using Sentinel-2 data and employing the artificial intelligence (AI) technique (Karra et al. 2021). A visual comparison of the Esri LULC map and Google Earth imagery shows that the former map underestimated the total tree cover, which was accurately captured in the current study. This could indicate the limitation of the global model or data at the local scale. The major LULC changes from 2016–17 to 2019–20 were observed as an expansion in built-up area (839.09 ha) and cropland (298.96 ha) and a reduction in bare land (34.45 ha), grassland (497.01 ha), and tree cover area (619.94 ha). Various factors are causing the rapid urban area enlargement of Pune city, including population expansion, rural-to-urban migration, inter-state migration, and agglomeration of peri-urban areas (Link et al. 2021; Singh and Basu 2020; Lad and Petkar 2022). In addition, infrastructure projects also contributed to the built-up area expansion (Appendix C). Bhaskar (2012) studied the urbanization of Pune city and reported a 4,300 ha growth from 1999 to 2009, with most of the urban area growth replacing barren and fallow land, followed by tree cover. Recently, Gohain et al. (2021) studied the LULC change in Pune city using Landsat data and reported a nearly 5,000 ha increase in built-up area from 1990–2019 at the expense of agriculture, scrubland, and fallow land. They also observed a corresponding increase in the land surface temperature of 1.4°C in summer and winter compared to the surrounding rural areas.

The optical and microwave backscatter (GRD) data do not provide direct canopy height estimates. Instead, these variables help estimate canopy height as proxy variables. The multi-temporal TOA reflectance, spectral indices, and microwave backscatter values are widely utilized in canopy height estimation. The RF machine learning algorithm was used to develop the nonlinear relationship between dependent (GEDI data) and independent variables (Sentinel-1 and Sentinel-2 bands, vegetation indices, and DEM) (Potapov et al. 2020). GEDI LiDAR tree canopy height data were used here owing to the unavailability of reliable ground data. The RF regression model indicated reliable accuracy (R^2 = 0.74; RMSE = 2.85 m) in estimating plant canopy height by integrating the optical and microwave determinants. Fagua et al. (2019) integrated ALOS-PALSAR (Phased Array type L-band SAR) and Landsat-8 data as determinant variables for canopy height estimation using Airborne LiDAR data as a ground reference.

They applied five regression models (RF, multivariate adaptive regression splines [MARS], linear regression [lm], Lasso and Elastic-Net Regularized Generalized Linear Models [GLM.net], and Support Vector Machine [SVM]) for three forest types and three predictor groups. They observed an error ranging from 1.2 to 3.4 m for dry forests and 5.1–7.4 m in the rainforest and reported the highest accuracy for the RF model. Nandy et al. (2021) employed Sentinel-1 backscatter bands and derived indicators to estimate canopy height using ICESat-2 LiDAR data in Dehradun district, which is located in the western Indian Himalayan foothills. They applied the RF regression model and reported a very high accuracy (R^2 = 0.89 and RMSE = 1.11 m). The global tree canopy height map generated by Potapov et al. (2020) for 2019 was compared with the current study outcome. A comparison with validation data (30 percent of the total GEDI observations) shows erroneous tree canopy height estimation (R^2 < 0.01; RMSE > 9 m) using the global model (see Figure F1 in Appendix F) (Potapov et al. 2020). This could indicate the limitations of global or continental models at a local scale.

We considered using the Pune Tree Census survey, which was conducted from 2016 to August 2019, as the ground truth (the exact date of the survey for each point is unavailable [*Hindustan Times* 2022; PMC n.d.b]). The tree census data contains about 4.01 million individual data points within the PMC area. These data contain every tree's geographic location (latitude and longitude) with species information (local and scientific name), tree height, crown cover, girth size, tree health, and so on. In the tree census survey, 40,890 dead trees were recorded within the PMC area. The raintree *(Samanea saman)*, giripushpa *(Gliricidia sepium)*, gulmohar *(Delonix regia)*, subabul *(Leucaena leucocephala)*, and banyan *(Ficus benghalensis)* are found to be the dominant tree species in the PMC area. From the remaining data (excluding dead trees), we randomly selected multiple tree cover patches for visual inspection with respect to the high-resolution QuickBird imagery (accessed from Google Earth). The comparison shows geolocation errors for many data points in the tree census data, such as false-positive observations in which many points represent trees in the tree census data but are absent in the high-resolution Google Earth imagery (see Figures G1 and G2 in Appendix G). Similarly, false-negative observations show the presence of tree cover in high-resolution Google Earth imagery, but this is not recorded in the tree census data (see Figures G1 and G2 in Appendix G). Moreover, the number of trees recorded in a patch needs to be rechecked and revalidated (see Figure G2 in Appendix G). Due to discrepancies in many geolocations and data records, the tree census survey data were not employed in the current study for tree cover mapping and canopy height estimation. We think the Pune tree census survey data need to be verified and cleaned.

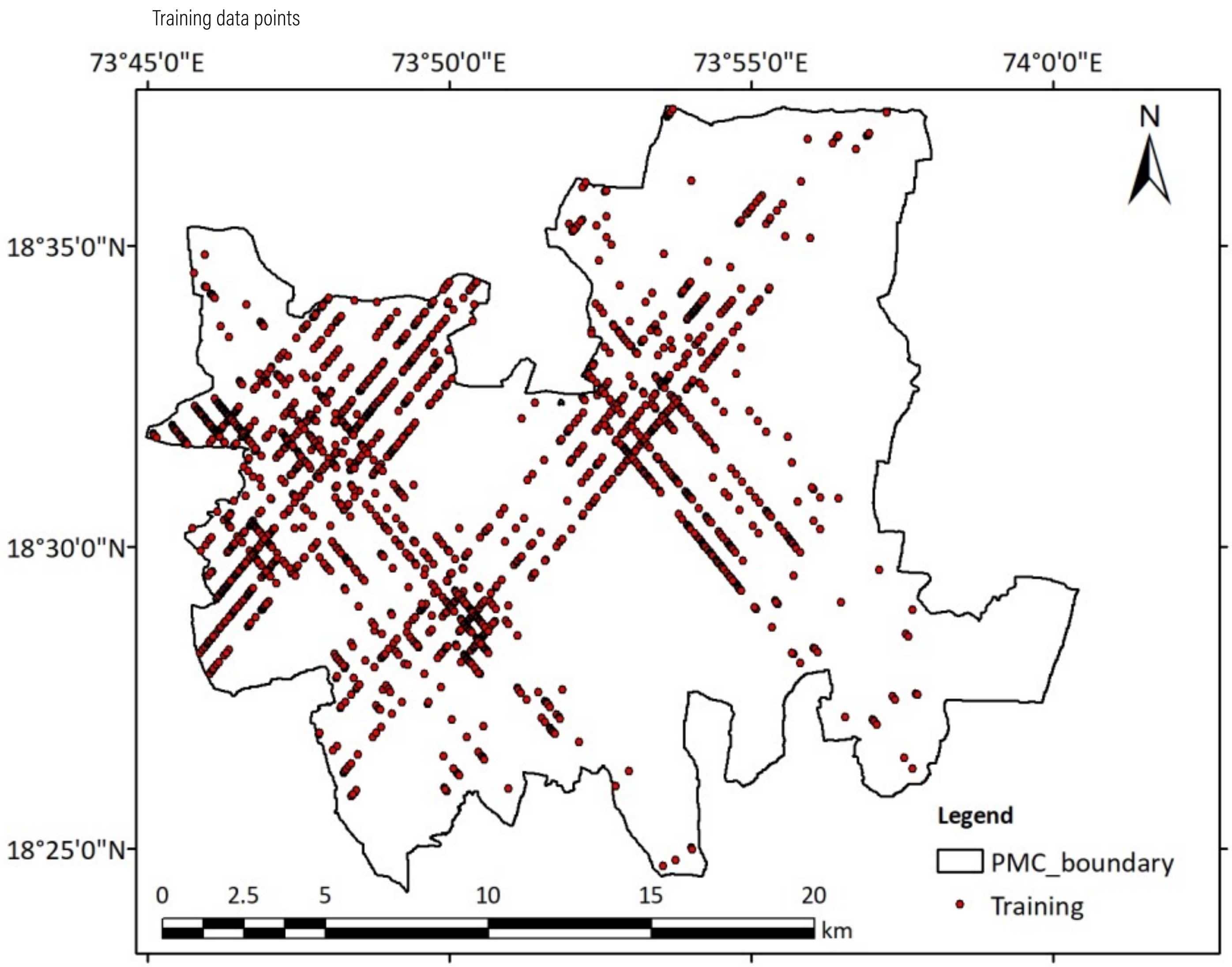
The approaches adopted in the current study represent cost-effective and robust methodologies based on publicly available satellite data. We could not collect field data due to COVID-related travel restrictions. A further study could focus on acquiring ground data and verifying the acquired accuracy in Pune city. The adopted approach can be tested in other landscapes (urban and rural) by employing suitable ground observations. However, future studies may include higher-resolution data for more accurate urban green space mapping. Continuous decreases in green cover and increases in impervious cover in cities are observed globally (Nowak and Greenfield 2020). Improving urban green space can fulfill important urban needs, including heat-risk mitigation, sequestering atmospheric CO_2, improving air quality, increasing climate resilience, and providing the benefits of multiple other ecosystem services. Understanding the LULC dynamics, especially the changes in tree cover and impervious cover, is essential for evaluating urban heating, prescribing suitable tree-based interventions, and selecting sites for such interventions. The spatially explicit existing tree cover resources and canopy height maps are important for evaluating the biodiversity, biomass, and carbon sequestration potential. A significant loss of green cover is observed in Pune city at the expense of urban area expansion within this three-year period. We recommend factoring in the conservation and protection of existing tree cover into planning. Alternatively, the bare land areas could be utilized to expand built-up areas. Recently, Balasubramanian et al. (2022) assessed the existing tree cover and interventions in Kochi city, Kerala, and suggested protecting and maintaining the existing tree cover, including mangroves. They have prescribed suitable tree-based restoration activities for urban landscapes, such as plantations in vacant areas, home gardens, boundary plantations, and avenue plantations. Such studies can be conducted in Pune and other Indian cities to explore the potential for and initiate a tree cover increase.

CONCLUSIONS

The GEE platform was used in most of the data processing and provided high computational power with minimal system requirements on the user side. The adopted methodologies were mostly automated, which enabled monitoring of land cover dynamics, tree cover and urban green space, and tree height at the desired temporal intervals. The adopted method can be extrapolated to other landscapes to rapidly assess TOF, biodiversity, biomass, and carbon sequestration potential. Moreover, these spatially explicit maps would help the urban planners of the PMC, the Maharashtra State Forest Department, policy developers, and decision-makers to prioritize the protection and maintenance of tree cover areas, identify areas and interventions for tree cover increase, build green infrastructure, increase public green spaces, develop suitable policies, and so on. Moreover, the spatially explicit LULC change and canopy height maps generated in this study can act as important inputs to numerous applications, including ecosystem service assessment, support for climate change mitigation, and socioeconomic and human well-being studies. The adopted approach will assist the national TOF monitoring carried out by the FSI. The report will be useful for various national and international institutes such as the Indian Council of Forestry Research and Education (ICFRE) and the International Centre for Research in Agroforestry (ICRAF) and help them strengthen tree cover monitoring, planning, and implementation activities.

APPENDIX A

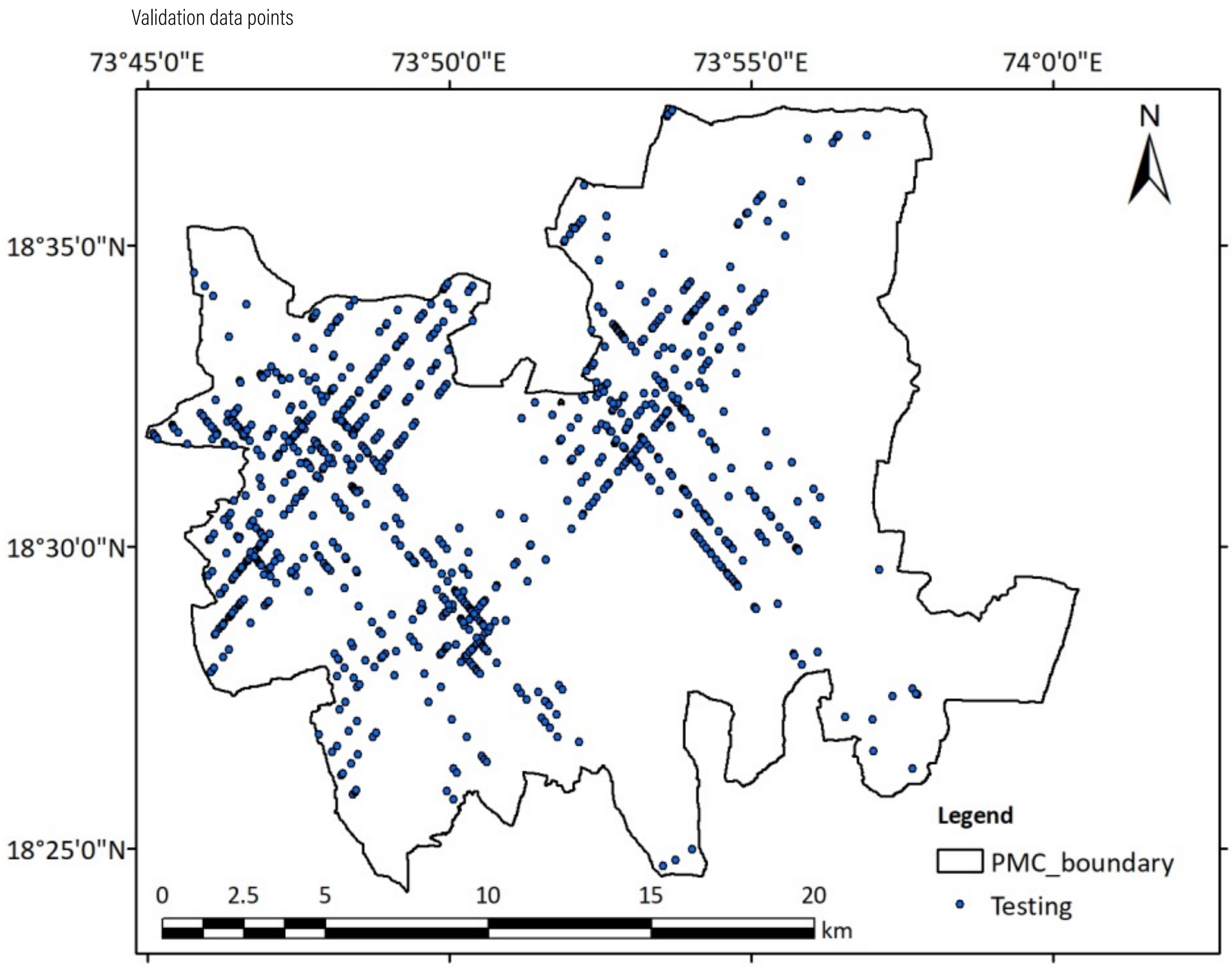
Figure A1 | **Training GEDI data points**



Notes: GEDI = Global Ecosystem Dynamics Investigation; PMC = Pune Municipal Corporation

Source: Earthdata n.d.; PMC n.d.a.

Figure A2 | **Testing GEDI data points**



Notes: GEDI = Global Ecosystem Dynamics Investigation; PMC = Pune Municipal Corporation

Source: Earthdata n.d.; PMC n.d.-a.

APPENDIX B

Table B1 | **LULC classification error matrix for 2016-2017**

1	REFERENCE DATA 2019-20 CHANGE								
	2016-2017	**Tree Cover**	**Bareland**	**Waterbody**	**Built-up**	**Cropland**	**Grassland**	**Row Total**	**User's Accuracy**
Classified data	Tree Cover	63	2	0	2	1	2	70	90.00
	Bareland	0	29	0	0	0	4	33	87.88
	Waterbody	0	0	16	0	0	0	16	100.00
	Built-Up	0	1	0	34	0	0	35	97.14
	Cropland	0	3	0	0	21	0	24	87.50
	Grassland	2	3	0	0	2	19	26	73.08
	Column Total	**65**	**38**	**16**	**36**	**24**	**25**	**204**	
	Producer's Accuracy	**96.92**	**76.32**	**100.00**	**94.44**	**87.50**	**76.00**		

Notes: Overall Classification Accuracy = 89.22%; Kappa: 0.86. LULC = land use land cover.

Source: WRI India authors.

Table B2 | **LULC classification error matrix for 2019-2020**

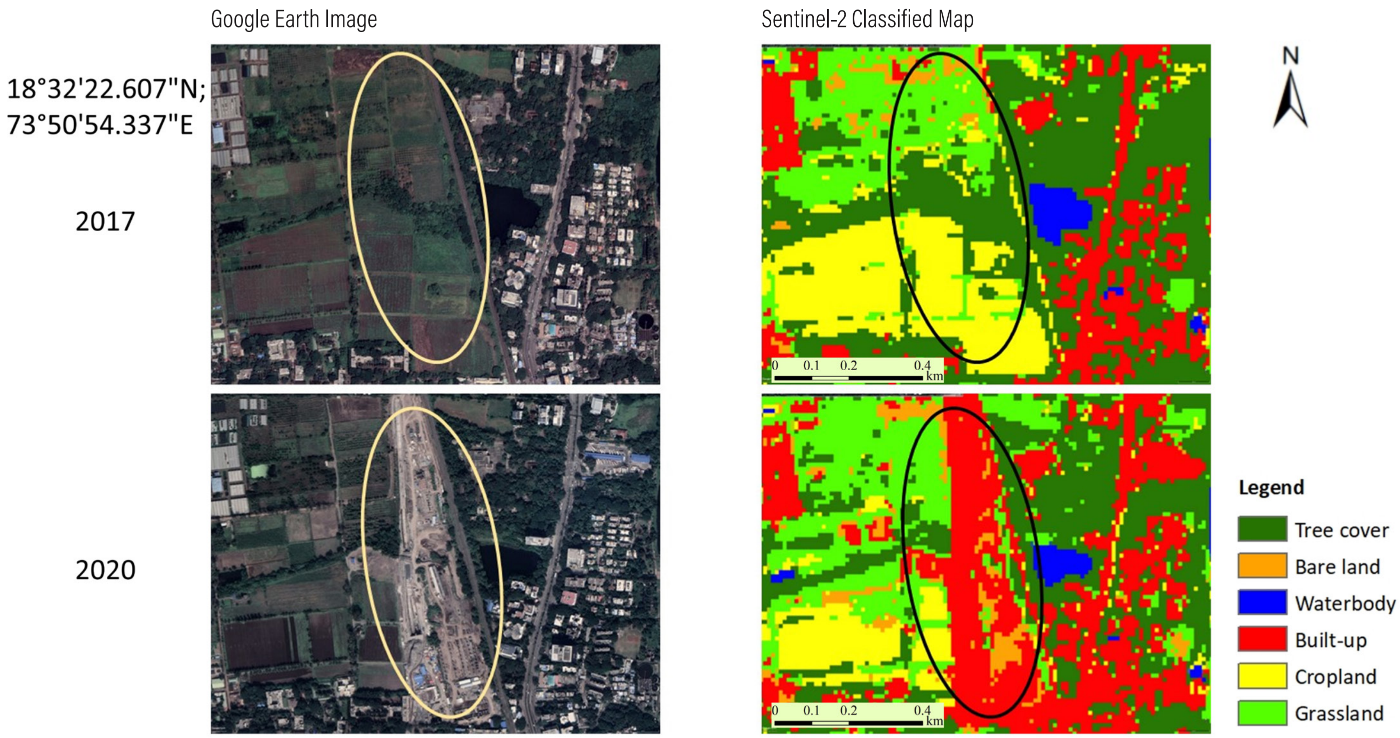
2	REFERENCE DATA 2019-20 CHANGE								
	2016-2017	**Tree Cover**	**Bareland**	**Waterbody**	**Built-up**	**Cropland**	**Grassland**	**Row Total**	**User's Accuracy**
Classified data	Tree Cover	65	0	0	1	1	0	67	97.01
	Bareland	0	30	0	0	0	1	31	96.77
	Waterbody	0	0	16	0	0	0	16	100.00
	Built-Up	0	2	0	35	0	0	37	94.59
	Cropland	0	1	0	0	19	0	20	95.00
	Grassland	0	5	0	0	4	24	33	72.73
	Column Total	**65**	**38**	**16**	**36**	**24**	**25**	**204**	
	Producer's Accuracy	**100.00**	**78.95**	**100.00**	**97.22**	**79.17**	**96.00**		

Notes: Overall Classification Accuracy = 92.65%; Kappa: 0.91. LULC = land use land cover.

Source: WRI India authors.

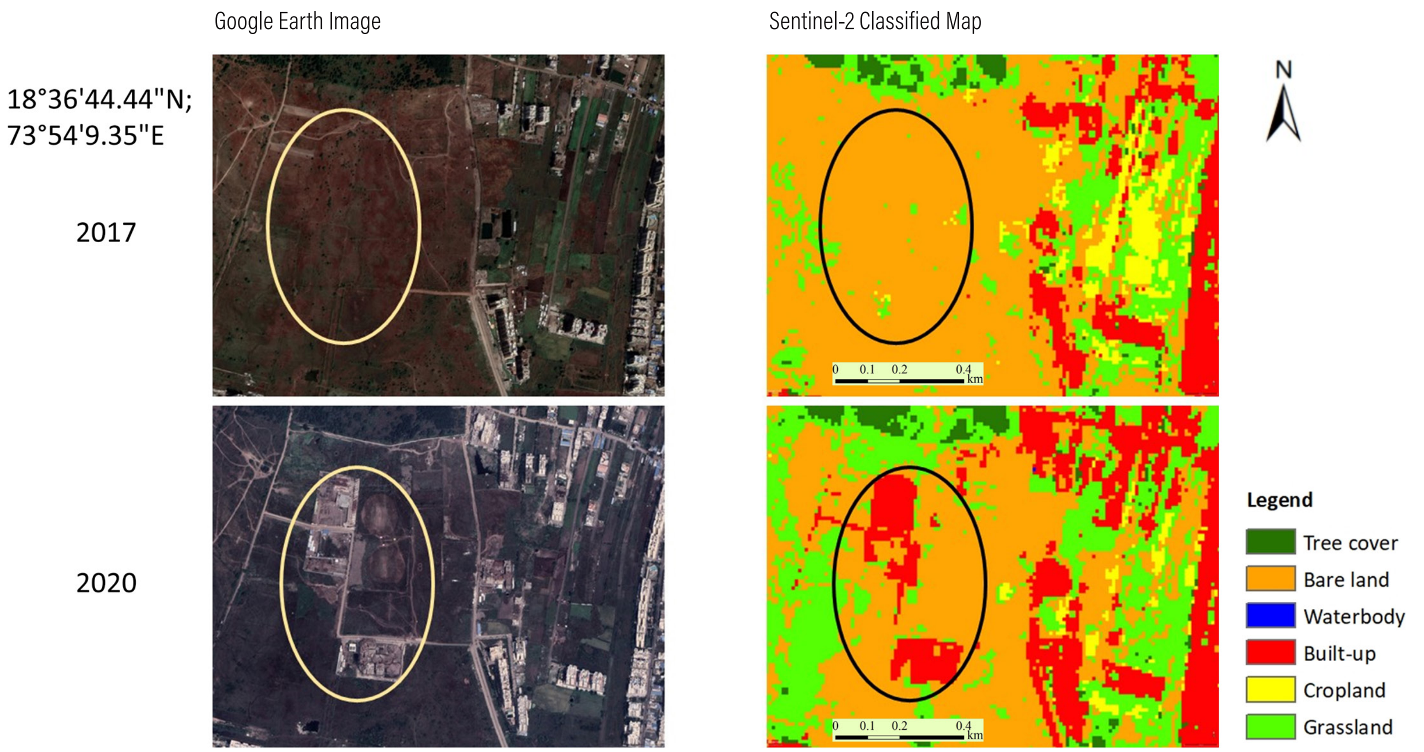
APPENDIX C

Figure C1 | **Cropland, grassland, and tree cover conversion to built-up area due to Metro Range Hill Depot construction**



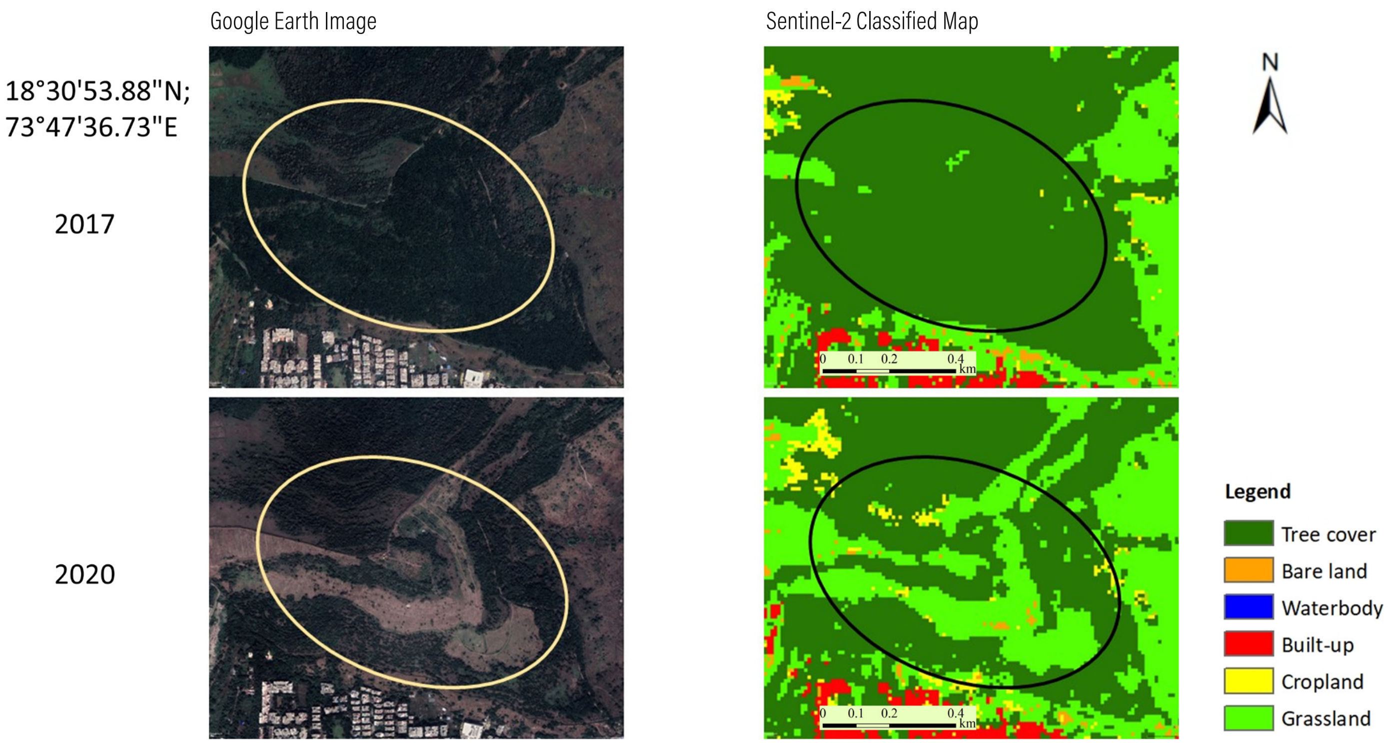
Source: WRI India authors.

Figure C2 | **Cropland, grassland, and tree cover conversion to the built-up area between Bharat Mata Road and Tingre Park Road**



Source: WRI India authors.

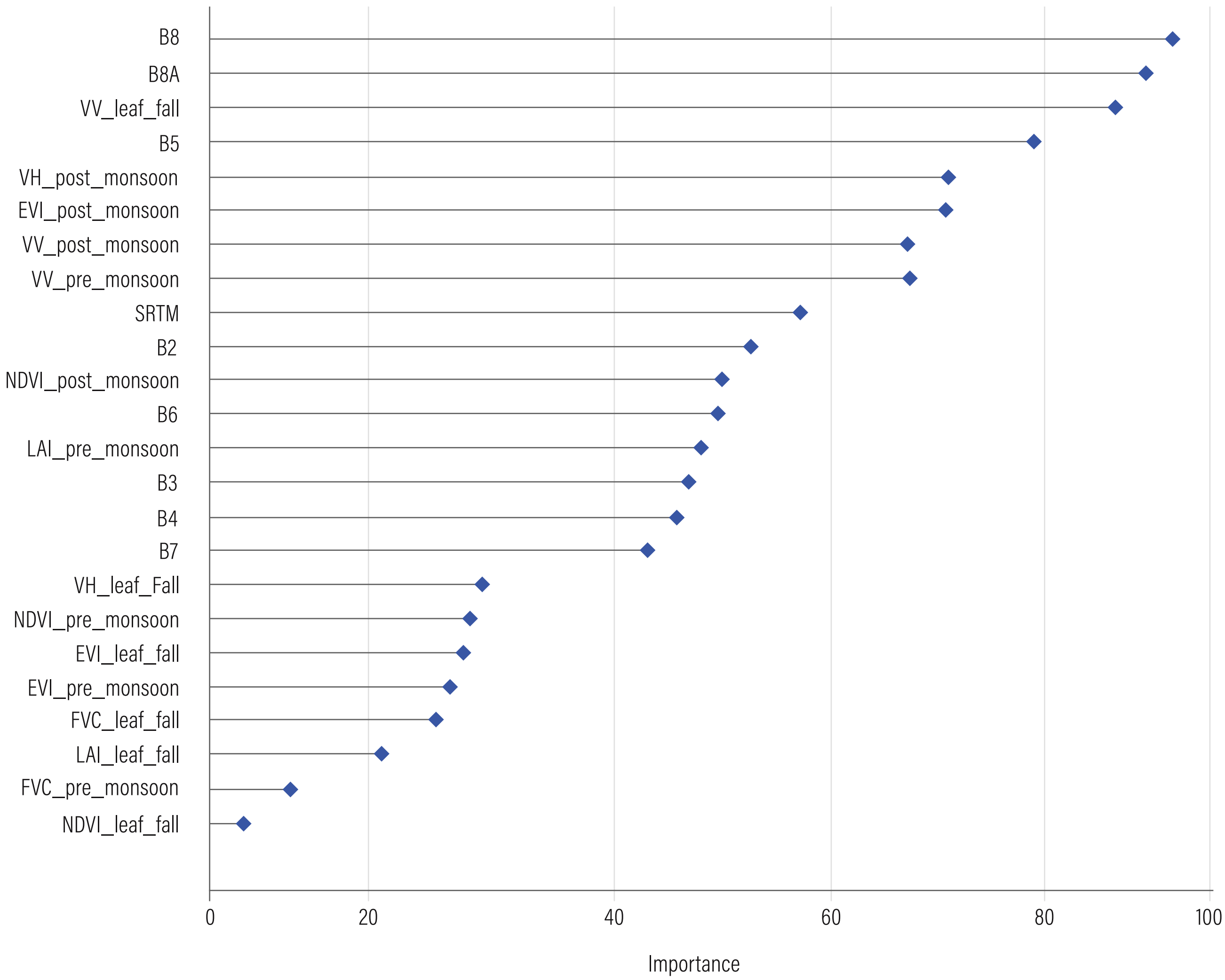
Figure C3 | **Tree cover loss and conversion to grassland on the Mhatoba Tekdi hill**



Source: WRI India authors.

APPENDIX D

Figure D1 | **Variable importance plot in canopy height regression analysis**

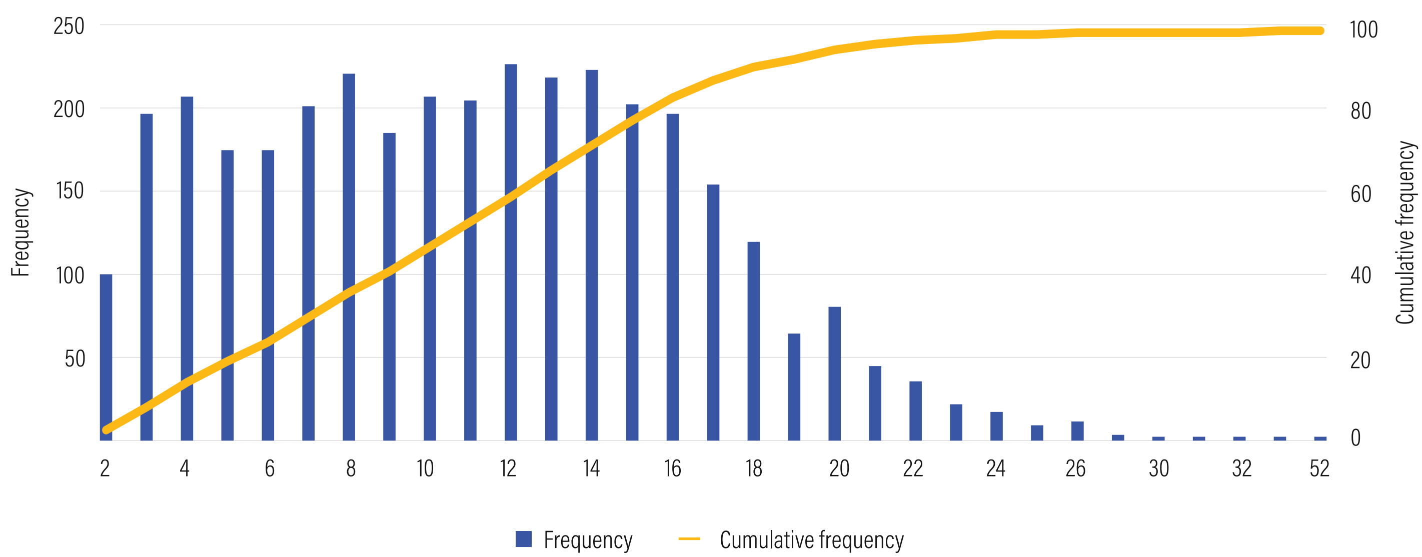


Notes: The X-axis indicates the variable importance ranges between 0 and 100.

Sources: WRI India authors.

APPENDIX E

Figure E1 | **GEDI-data-derived tree height distribution**

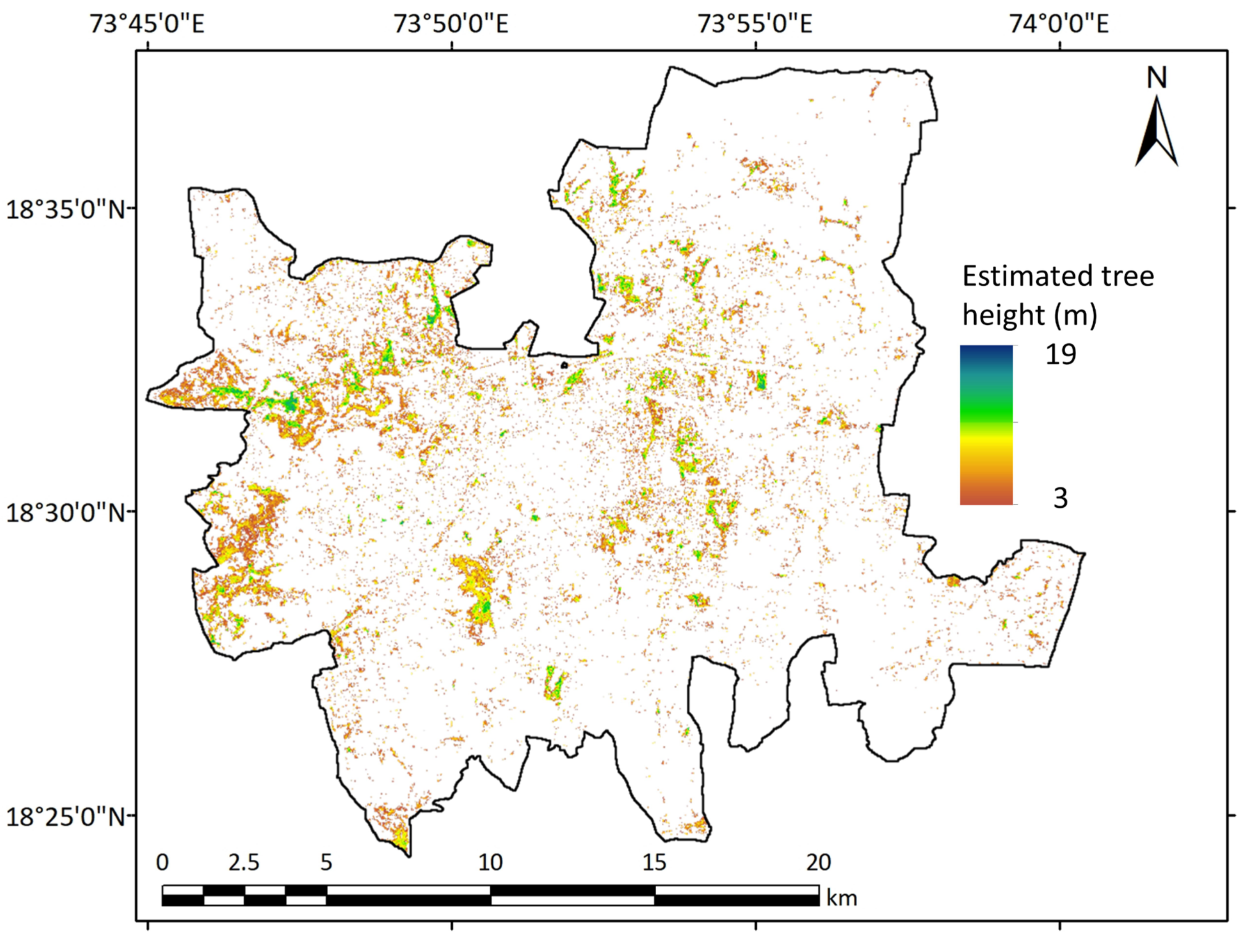


Notes: GEDI = Global Ecosystem Dynamics Investigation.

Sources: WRI India authors.

APPENDIX F

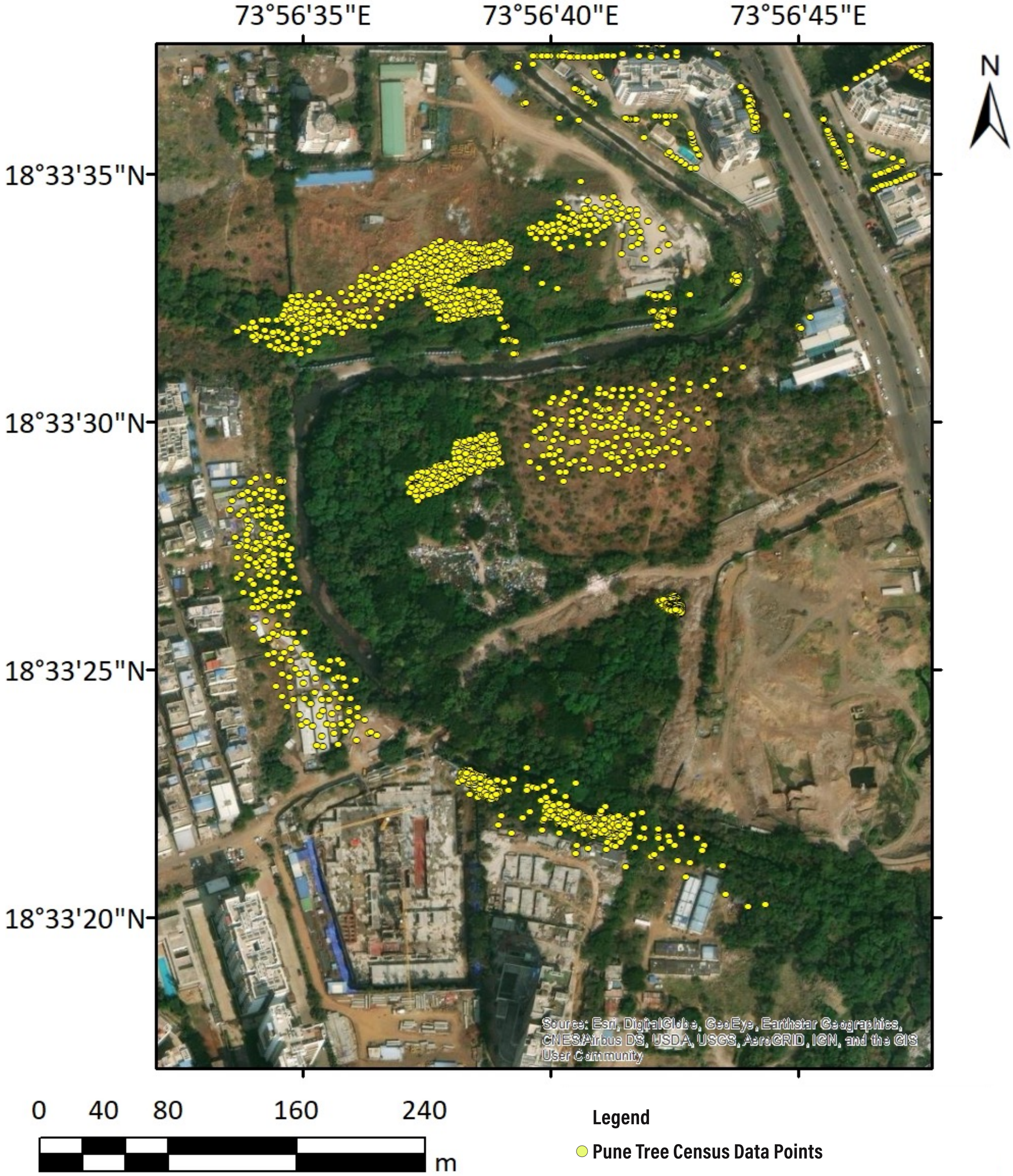
Figure F1 | **Estimated global canopy height map (accessed from Google Earth Engine)**



Sources: Potapov et al. 2020.

APPENDIX G

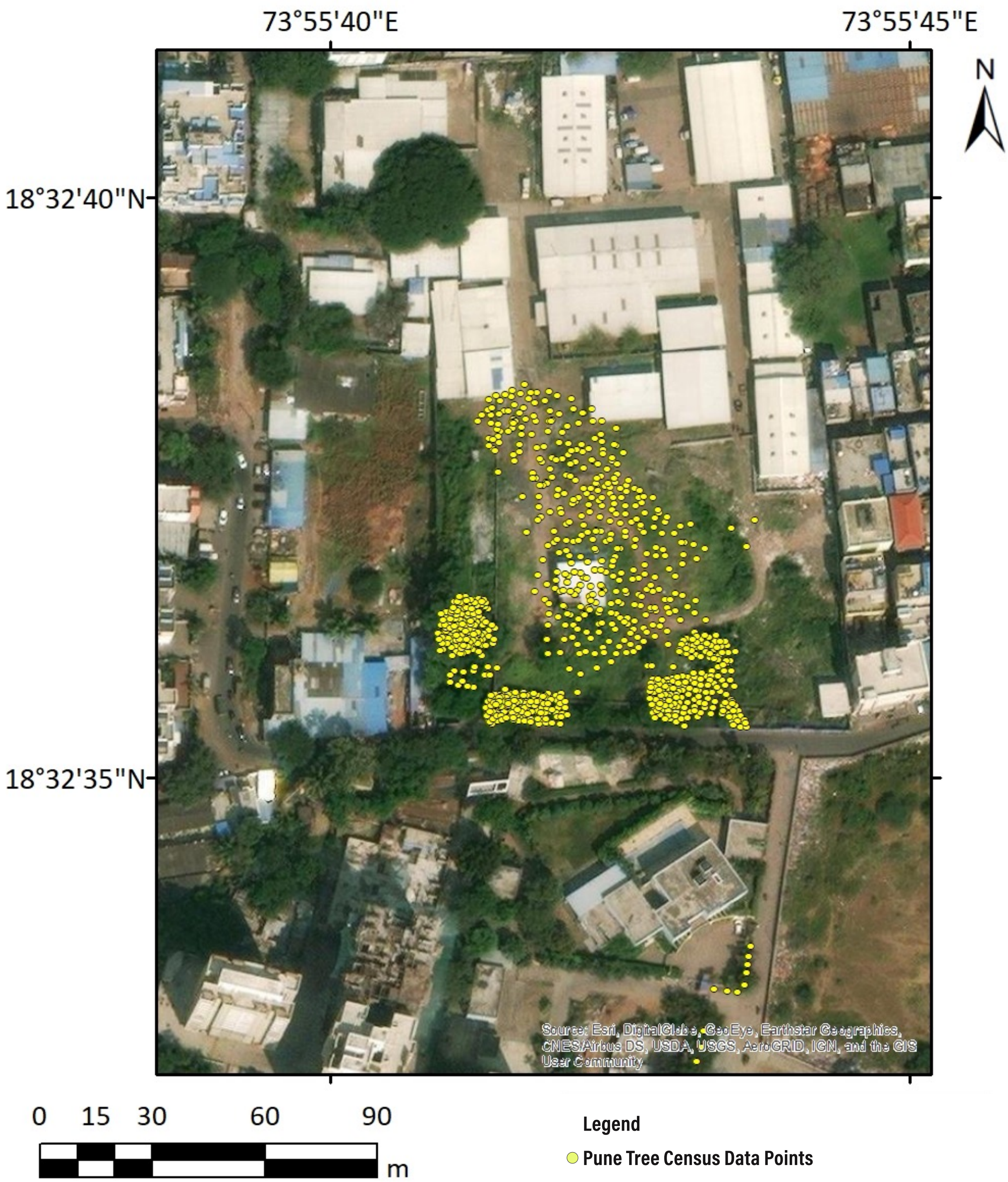
Figure G1 | **Comparison of Pune tree census survey data with TCC high-resolution Google Earth imagery**



Notes: TCC = True Color Composite. Green pixels represent tree cover, and yellow dots represent Pune tree census survey data.

Sources: WRI India authors.

Figure G2 | **Inaccurate geolocation, overestimation, and underestimation of trees in tree census data**



Notes: Green pixels represent tree cover, and yellow dots represent Pune tree census survey data.

Sources: WRI India authors.

REFERENCES

Balasubramanian, Kanchana, Sidhtharthan Segarin, Priya Narayanan, and Pulakesh Das. 2022. "Opportunities for Improving Urban Tree Cover: A Case Study in Kochi." In *Blue-Green Infrastructure Across Asian Countries: Improving Urban Resilience and Sustainability*, edited by Shalini Dhyani, Mrittika Basu, Harini Santhanam, and Rajarshi Dasgupta, 251–70. Springer. https://doi.org/10.1007/978-981-16-7128-9_11.

Bhaskar, Padigala. 2012. "Urbanization and Changing Green Spaces in Indian Cities (Case Study—City of Pune)." *International Journal of Geology, Earth and Environmental Sciences* 2 (2): 148–56.

Bolund, Per, and Sven Hunhammar. 1999. "Ecosystem Services in Urban Areas." *Ecological Economics* 29 (2): 293–301. https://doi.org/10.1016/S0921-8009(99)00013-0.

Census of India. 2011. *Census of India.* New Delhi: Office of the Registrar General and Census Commissioner, Ministry of Home Affairs, Government of India.

Chaturvedi, Rohini, Marie Duraisami, K.M. Jayahari, C.B. Kanchana, Ruchika Singh, Sidhtharthan Segarin, and Prabhakar Rajagopal. 2018. "Restoration Opportunities Atlas of India." Washington, DC: World Resources Institute. Available online at www.india.restorationatlas.org/methodology

Das, P., and V. Pandey. 2019. "Use of Logistic Regression in Land-Cover Classification with Moderate-Resolution Multispectral Data." *Journal of the Indian Society of Remote Sensing* 47 (8). https://doi.org/10.1007/s12524-019-00986-8.

Dobbs, Cynnamon, Francisco J. Escobedo, and Wayne C. Zipperer. 2011. "A Framework for Developing Urban Forest Ecosystem Services and Goods Indicators." *Landscape and Urban Planning* 99 (3–4): 196–206. https://doi.org/10.1016/j.landurbplan.2010.11.004.

Dwyer, John F., E. Gregory McPherson, Herbert W. Schroeder, and Rowan A. Rowntree. 1992. "Assessing the Benefits and Costs of the Urban Forest." *Journal of Arboriculture* 18: 227–34. https://doi.org/10.48044/jauf.1992.045.

Earthdata. n.d. "Earthdata Search." *Earthdata: Find a DAAC*. https://search.earthdata.nasa.gov/search. Accessed November 1, 2022.

ESA (European Space Agency). n.d. "SNAP—Earth Online." The European Space Agency. https://earth.esa.int/eogateway/tools/snap. Accessed November 1, 2022.

Fagua, J. Camilo, Patrick Jantz, Susana Rodriguez-Buritica, Laura Duncanson, and Scott J. Goetz. 2019. "Integrating LiDAR, Multispectral and SAR Data to Estimate and Map Canopy Height in Tropical Forests." *Remote Sensing* 11 (22): 2697. https://doi.org/10.3390/rs11222697.

GEDI (Global Ecosystem Dynamics Investigation). n.d. "Products." *GEDI Ecosystem Lidar.* https://gedi.umd.edu/data/products/. Accessed November 1, 2022.

Ghosh, Sujit Madhab, Mukunda Dev Behera, and Somnath Paramanik. 2020. "Canopy Height Estimation Using Sentinel Series Images through Machine Learning Models in a Mangrove Forest." *Remote Sensing* 12 (9): 1519. https://doi.org/10.3390/rs12091519.

Ghosh, Sujit Madhab, Mukunda Dev Behera, Subham Kumar, Pulakesh Das, Ambadipudi Jaya Prakash, Prasad Kumar Bhaskaran, Parth Sarathi Roy, Barik, Saroj Kanta, Jeganathan, Chockalingam, Srivastava, Prashant Kumar & Behera, Soumit Kumar. 2022. "Predicting the forest canopy height from LiDAR and multi-sensor data using machine learning over India" 23) :5968. https://doi.org/10.3390/rs14235968.

Gohain, Kashyap Jyoti, Pir Mohammad, and Ajanta Goswami. 2021. "Assessing the Impact of Land Use Land Cover Changes on Land Surface Temperature over Pune City, India." *Quaternary International* 575–576 (February): 259–69. https://doi.org/10.1016/j.quaint.2020.04.052.

Hindustan Times. 2019. "Geo-Enabled Tree Census Finds Pune Has More Green Cover in 2019." *Hindustan Times*, May 22. https://www.hindustantimes.com/pune-news/geo-enabled-tree-census-finds-pune-has-more-green-cover-in-2019/story-Q9BFM-rXnEeY8BfooNy5Y5K.html.

FSI. 2019."India State of Forest Report 2019". Dehradun, India: FSI.

Jose, Shibu. 2009. "Agroforestry for Ecosystem Services and Environmental Benefits: An Overview." *Agroforest Systems* 76 (1): 1–10. https://doi.org/10.1007/s10457-009-9229-7.

Karra, Krishna, Caitlin Kontgis, Zoe Statman-Weil, Joseph C. Mazzariello, Mark Mathis, and Steven P. Brumby. 2021. "Global Land Use/Land Cover with Sentinel 2 and Deep Learning." In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 4704. IEEE. DOI: 10.1109/IGARSS47720.2021.9553499.

Lad, Ketaki Suhas, and Arati Siddharth Petkar. 2022. "Assessing Transformations in Peri-urban Areas Using GIS: A Case of Pune City, India." *GeoScape* 16 (1): 80–92. https://doi.org/10.2478/geosc-2022-0007.

Lee, Won-Jin, and Chang-Wook Lee. 2018. "Forest Canopy Height Estimation Using Multiplatform Remote Sensing Dataset." *Journal of Sensors* 2018 (April): e1593129. https://doi.org/10.1155/2018/1593129.

Link, Ann-Christine, Yuanzao Zhu, and Raphael Karutz. 2021. "Quantification of Resilience Considering Different Migration Biographies: A Case Study of Pune, India." *Land* 10 (11): 1134.

Mishra, Arun Pratap, Ishwari Datt Rai, Divesh Pangtey, and Hitendra Padalia. 2021. "Vegetation Characterization at Community Level Using Sentinel-2 Satellite Data and Random Forest Classifier in Western Himalayan Foothills, Uttarakhand." *Journal of the Indian Society of Remote Sensing* 49 (4): 759–71. https://doi.org/10.1007/s12524-020-01253-x.

Nandy, Subrata, Ritika Srinet, and Hitendra Padalia. 2021. "Mapping Forest Height and Aboveground Biomass by Integrating ICESat-2, Sentinel-1 and Sentinel-2 Data Using Random Forest Algorithm in Northwest Himalayan Foothills of India." *Geophysical Research Letters* 48 (14): e2021GL093799. https://doi.org/10.1029/2021GL093799.

Nowak, David J., and Eric J. Greenfield. 2020. "The Increase of Impervious Cover and Decrease of Tree Cover within Urban Areas Globally (2012–2017)." *Urban Forestry & Urban Greening* 49: 126638. https://doi.org/10.1016/j.ufug.2020.126638.

Potapov, Peter, Xinyuan Li, Andres Hernandez-Serna, Alexandra Tyukavina, Matthew C Hansen, Anil Kommareddy, Amy Pickens, et al. 2021. "Mapping Global Forest Canopy Height through Integration of GEDI and Landsat Data." *Remote Sensing of Environment* 253 (February): 112165. https://doi.org/10.1016/j.rse.2020.112165.

PMC (Pune Municipal Council). n.d.a. "PMC Web GIS Portal." Pune Municipal Council. https://gis.pmc.gov.in. Accessed November 1, 2022.

PMC. n.d.b. "PMC Open Data Store." *Pune Municipal Council*. http://opendata.punecorporation.org/Citizen/CitizenDatasets/Index. Accessed November 1, 2022.

Qi, Wenlu, Seung-Kuk Lee, Steven Hancock, Scott Luthcke, Hao Tang, John Armston, and Ralph Dubayah. 2019. "Improved Forest Height Estimation by Fusion of Simulated GEDI Lidar Data and TanDEM-X InSAR Data." *Remote Sensing of Environment 221* (February): 621–34. https://doi.org/10.1016/j.rse.2018.11.035.

Roy, Sambhunath, Sujoy Mudi, Pulakesh Das, Sujit Ghosh, Pravat Kumar Shit, Gouri Shankar Bhunia, and John Kim. 2021. "Estimating Above Ground Biomass (AGB) and Tree Density Using Sentinel-1 Data." *In Spatial Modeling in Forest Resources Management: Rural Livelihood and Sustainable Development*, edited by Pravin Kumar Shit, Hamid Reza Pourghasemi, Pulakesh Das, and Gouri Shankar Bhunia. Springer. https://doi.org/10.1007/978-3-030-56542-8_11.

Shetty, Shobitha, Prasun Kumar Gupta, Mariana Belgiu, and S.K. Srivastav. 2021. "Assessing the Effect of Training Sampling Design on the Performance of Machine Learning Classifiers for Land Cover Mapping Using Multi-temporal Remote Sensing Data and Google Earth Engine." *Remote Sensing* 13 (8): 1433. https://doi.org/10.3390/rs13081433.

Singh, Chandni, and Ritwika Basu. 2020. "Moving in and out of Vulnerability: Interrogating Migration as an Adaptation Strategy along a Rural Urban Continuum in India."*Geographical Journal* 186 (1): 87–102. https://doi.org/10.1111/geoj.12328.

Talukdar, Swapan, Pankaj Singha, Susanta Mahato, Shahfahad, Swades Pal, Yuei-An Liou, and Atiqur Rahman. 2020. "Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—a Review." *Remote Sensing* 12 (7): 1135. https://doi.org/10.3390/rs12071135.

ACKNOWLEDGMENTS

The authors thank the Sustainable Landscape Restoration (SLR) team members, WRI India, for their continuous support. The authors are immensely grateful to Dr. Ruchika Singh, Dr. K.M. Jayahari, and Sidhtharthan Segarin, WRI India, for their continuous feedback, suggestions, and encouragement. The authors thank Gregory Taff, WRI DC, and the reviewers for their constructive comments and recommendations. The study was conducted as a part of WRI India's internal research grant, and this financial support is acknowledged here.

ABOUT THE AUTHORS

Pulakesh Das was a Program Manager, Sustainable Landscapes and Restoration, WRI India.

Sayantan Dey was an Early Professional Program Consultant, Sustainable Landscapes and Restoration, WRI India

Kanchana Balasubramanian was a Senior Research Associate, Sustainable Landscapes and Restoration, WRI India

Parth Sarathi Roy is a Senior Fellow, Sustainable Landscapes and Restoration, WRI India

DISCLAIMER: Maps are for illustrative purpose and does not imply the expression of any opinion on the part of WRI India, concerning the legal status of any country or territory or concerning the delimitation of frontiers or boundaries.

ABOUT WRI INDIA

WRI India, an independent charity legally registered as the India Resources Trust, provides objective information and practical proposals to foster environmentally sound and socially equitable development. Our work focuses on building sustainable and liveable cities and working towards a low carbon economy. Through research, analysis, and recommendations, WRI India puts ideas into action to build transformative solutions to protect the earth, promote livelihoods, and enhance human well-being. We are inspired by and associated with World Resources Institute (WRI), a global research organization. Know more: www.wri-india.org.

Our challenge

Natural resources are at the foundation of economic opportunity and human well-being. But today, we are depleting Earth's resources at rates that are not sustainable, endangering economies and people's lives. People depend on clean water, fertile land, healthy forests, and a stable climate. Livable cities and clean energy are essential for a sustainable planet. We must address these urgent, global challenges this decade.

Our vision

We envision an equitable and prosperous planet driven by the wise management of natural resources. We aspire to create a world where the actions of government, business, and communities combine to eliminate poverty and sustain the natural environment for all people.

Our approach

COUNT IT

We start with data. We conduct independent research and draw on the latest technology to develop new insights and recommendations. Our rigorous analysis identifies risks, unveils opportunities, and informs smart strategies. We focus our efforts on influential and emerging economies where the future of sustainability will be determined.

CHANGE IT

We use our research to inform government policies, business strategies, and civil society action. We test projects with communities, companies, and government agencies to build a strong evidence base. Then, we work with partners to deliver change on the ground that alleviates poverty and strengthens society. We hold ourselves accountable to ensure our outcomes will be bold and enduring.

SCALE IT

We don't think small. Once tested, we work with partners to adopt and expand our efforts regionally and globally. We engage with decision-makers to carry out our ideas and elevate our impact. We measure success through government and business actions that improve people's lives and sustain a healthy environment.