IMPROVING REMOTE MONITORING OF CARBON STOCK IN TROPICAL FORESTS WITH MACHINE LEARNING, A CASE STUDY IN INDONESIAN BORNEO

Anonymous authors
Paper under double-blind review

ABSTRACT

One of the most effective approaches to mitigating climate change is to monitor the carbon stock in the tropical rainforests. However, biomonitoring of carbon in the forests is expensive and challenging due to inaccessibility. To improve carbon stock monitoring and the evaluation of fine-scale forest loss, we established a rapid, automatic, and cost-efficient generalized machine learning framework [1] that uses diverse remote sensing data and satellite imagery to accurately estimate aboveground carbon density, at fine-grained resolution (tens of meters), in remote tropical rainforests. The study area first focused on rainforests in Indonesian Borneo. In our preliminary tests on 80 sites in Indonesian Borneo, our machine learning models were capable of producing ACD estimates with R^2 of 0.7-0.8, which is a significant improvement from the comparable works (0.5-0.6 at best). This machine learning framework will be used to facilitate further carbon stock modeling in other forest regions (e.g. Brazil) as well as for the general purpose of climate change mitigation.

1 INTRODUCTION

Biomonitoring of carbon in tropical rainforests can be expensive and challenging due to inaccessibility. Traditional fieldwork of estimating aboveground carbon is a labor-intensive and time-consuming process, requiring at least 5-10 field workers at a time and taking weeks to complete. Furthermore, even with this expansive effort and cost, it yields information restricted to one site and often that is largely untranslatable across other sites.

Existing platforms that describe and map forest loss, including Hansen Global Forest Watch (Hansen et al., 2013) (`https://data.globalforestwatch.org`), are highly useful data sources and tools for forest management and monitoring. Yet, they don't currently provide robust spatially gridded estimates of aboveground carbon density (ACD) and cannot be used for rapid and targeted scale change monitoring. The Hansen Global Forest Watch datasets are designed for global wall-to-wall coverage and aggregate data annually. Furthermore, as Hansen Global Forest Watch attests (Hansen et al., 2013), they do not adequately deal with uncertainty due to cloud cover in the underlying imagery. Therefore, novel datasets and novel data pipelines are required in order to produce high-resolution, accurate time series data on ACD changes, both for statistical analyses, as well as for detailed impact monitoring.

As a solution, in this study, we develop an open-source, rapid, automatic, scalable, and cost-efficient generalizable framework to accurately estimate ACD, including to accurately estimate uncertainty due to cloud cover. By stacking multiple sources of coarse-resolution satellite imagery and remotely-sensed earth data atop the high-resolution canopy tree height Light Detection and Ranging (LiDAR) data, we have trained machine learning (ML) and deep learning (DL) algorithms to estimate actual canopy tree height values over areas that lack high-quality LiDAR data. Our work demonstrates a solution for increased accessibility of carbon estimation of forests from freely-available, open-source remote sensing data.

[1] https://github.com/deleo-lab/carbon-remote-sensing-ml

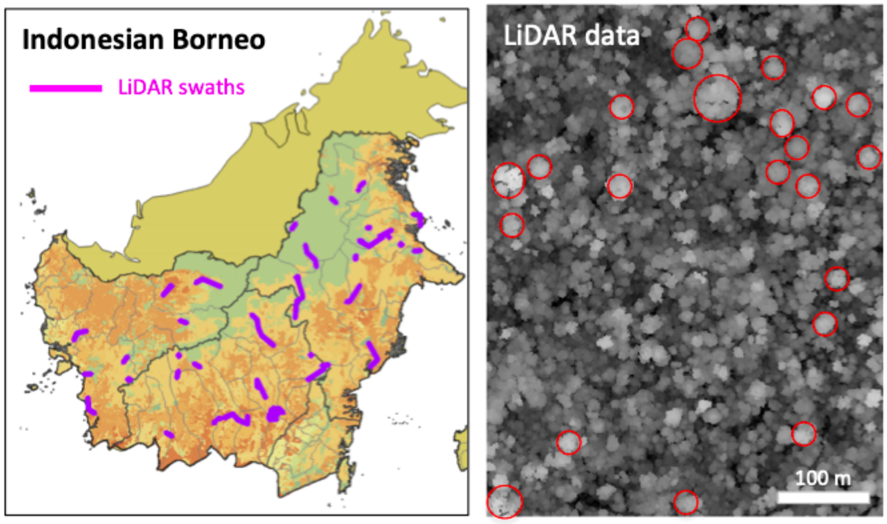

Figure 1: (Left) NASA CMS LiDAR data overview, indicating LiDAR swaths (purple color) in 90 sites in Indonesia Borneo. (Right) Raw LiDAR data points, red color circle indicates the tree canopy height > 30 m.

2 METHODOLOGY

2.1 DATA PIPELINE

Borneo, a 750,000 km^2 island, contains some of the world's most biodiverse and carbon-dense tropical forests. However, Borneo has lost about 62% of its forests within the last 40 years (Gaveau et al., 2019). In 2014, the NASA Carbon Monitoring System (CMS) surveyed 90 sites in Indonesian Borneo using plane-based LiDAR (Melendy et al., 2014), therefore, providing scientists with digital surface models (DSM) for 100,000 hectares of tropical forest tree canopy height (Figure 1). We included this canopy tree height data into our data pipeline and utilized it to calculate ACD at the pixel-level, as our ML regression target (see Section 2.2).

Similar to previous studies (Mascaro et al., 2014; Pham et al., 2014), we integrated several sources of common remote sensing data in our data pipeline, including LANDSAT-8, Sentinel-1, land cover fractional coverages from the Copernicus satellite mission, digital elevation model (DEM), and NASA CMS airborne LiDAR. In addition to these common data, we also integrated other data sources, such as vegetation indices commonly used to assess vegetation (e.g. NDVI, NDWI, NDII, EVI), Gray-Level Co-occurrence Matrix (GLCM) texture metrics (Haralick et al., 1973) for all spectral data, as well as the climatic data from BIOCLIM.

All of these data sources (listed in Table 1) were compiled into an uniform data processing pipeline which produced our dataset of over 300 features for each LiDAR value, resulting in over 10 million observations. This diverse remote sensing dataset was critical to providing the foundation on which we produced a robust, predictive ML framework.

2.2 ACD ESTIMATION

In this study, we followed the empirical relation, built by Jucker et al. (2018), which only requires tree canopy height (TCH) from airborne LiDAR as the input to calculating ACD. The empirical equations are as follows:

$$\mathrm{ACD} = 0.567 \times \mathrm{TCH}^{0.553} \times \mathrm{BA}^{1.081} \times \mathrm{WD}^{0.186}$$

$$\mathrm{BA} = 1.112 \times \mathrm{TCH}$$

$$\mathrm{WD} = 0.385 \times \mathrm{TCH}^{0.097}$$

Where ACD is in Mg C ha^{-1}, TCH, in m, BA, basal area, in m^2 ha^{-1}, and WD, wood density, in g cm^{-3}. Jucker et al. (2018) combined field data of tree height measurements with airborne laser scanning and hyperspectral imaging to characterize how topography shapes the vertical structure, wood density, diversity and ACD of nearly 15 km^2 of old-growth forest in Borneo. This general form of empirical relation was first developed by Asner & Mascaro (2014) and the allometric scaling coefficients in the equations are region-specific and are based on the field work that took place in that region.

2.3 ML ALGORITHMS

In our framework, we experimented with various ML and DL algorithms and techniques, including neural network models (artificial neural network, convolutional neural network), random forest (RF) (Breiman, 2001), XGBoost (Chen & Guestrin, 2016), LightGBM (Ke et al., 2017), and transfer learning with pre-trained models. From our data pipeline, we ran each ML algorithm with our dataset that contains 300 features and 10 million TCH and ACD observations. The regression target for each ML model is the ACD value in the pixel level. The goal is to develop a generalized ML model that is capable of predicting ACD solely based on the open-source remote sensing data in the regions that the LiDAR TCH is generally not available.

Data	**Type**	**Temporal Range**	**Spatial Resolution**	**Source**
LANDSAT-8	OLI/TIRS sensors	2013-2021	30 m	USGS
Vegetation Indices (NDVI, NDWI, NDII, EVI, calculated for Landsat-8)	Various measures associated with vegetation properties	Same as input	Same as input	Same as input
Sentinel-1	Synthetic Aperture Radar (SAR) instrument	2014-2021	10 m	ESA
Gray-Level Co-Occurrence Matrix (GLCM), derived from Landsat-8 and Sentinel-1	Textural image features derived from pixel spatial relationships	Same as input	Same as input	Same as input
Planet Labs/ PlanetScope	RGB + NIR high resolution data	2018-2021	3-5 m	Planet Labs, Inc. (available via API access)
ICESat	Spaced-based LiDAR	2018-2020	Fixed 100-m segments	NASA
CMS	Plane-mounted LiDAR	2014	1 m	NASA
NASA SRTM	Digital Elevation Model (elevation, slope, aspect)	2000	30 m	NASA
Bioclim	Climate	1970-2000	927.67 m	WorldClim
CopCover	Land Cover - Copernicus	2015	100 m	ESA

Table 1: Data Sources

3 RESULTS

In our initial experiments, to replicate the data and ML approach from the comparable studies (e.g. Mascaro et al. 2014, Pham et al. 2018) to serve as our benchmark, only multispectral satellite channels and land cover features were used in ML models (total of 30 features and 10,000 data points), RF consistently outperformed other models, with the out-of-sample R^2 = 0.45. For XGBoost, we performed extensive hyperparameter tuning using gridsearch, a ML tuning technique that attempts to compute the optimum values of the model's hyperparameters. We used XGBoost to perform feature importance ranking, which was later used to filter down the number of features to enhance the speed of the hyperparameter search. XGBoost had longer training and inference times, but it did surpass RF accuracy with some hyperparameter tuning, achieving R^2 = 0.52. Finally, we explored deep learning methods, including a Multi-layer Perceptron (MLP) model with regression as the final layer and convolutional neural networks (CNNs), however, the fine-tuned XGBoost model was still the best performing model with 30 data features. Note that these results are consistent with the comparable studies (e.g. Mascaro et al. 2014, Pham et al. 2018).

Next, we implemented the regression in XGBoost, RF and LightGBM, with our full dataset that consists of over 300 features for each LiDAR pixel value. We obtained the preliminary out-of-sample R^2 = 0.7-0.8, a dramatic improvement over the benchmark.

4 Discussion

In our benchmark experiment with 30 common remote sensing data features, although extensive model fine-tuning was performed, it was still not able to match the performance of simple ML models trained with the more diverse dataset with full 300 features. This resonates with the recent data-centric AI approach which aims to improve the data quality and richness instead of focus on improving ML code (Ng, 2021).

Note that the Landsat imagery used in our data pipeline is 30 x 30 m, so the sub-meter LiDAR data was resampled to derive mean height in each 30 x 30 m plot. The new generation of high temporal and spatial resolution satellite data, e.g. PlanetScope, which began observations over Borneo in 2018, consists of pixel cell sizes at 3-5 m^2 and would thus reduce the amount of data resampling, potentially providing a better estimation of TCH. PlanetScope sensors only capture four spectral bands (R, G, B NIR) vs. the eight from Landsat, but in our models features derived from NIR spectra consistently were ranked the most important, suggesting that PlanetScope sensors would still capture the most useful information.

The use of PlanetScope imagery is particularly appealing as PlanetScope satellites now cover the entire globe on a daily basis, allowing high temporal frequency imagery of the area of interest in Borneo, which greatly helps with dealing with cloud cover, particularly important in Borneo. Our next step involves combining PlanetScope high resolution satellite data with LiDAR canopy height data to be able to do daily monitoring of forest states at a spatial scale more relevant to individual trees and smaller patches of forest.

5 Conclusion

In this study, we developed an improved method for evaluating fine-scale carbon loss due to deforestation using ML and satellite data. Future efforts will employ Google Cloud to host this trained model on Google Earth Engine, which will create a rapid and easy way to accurately estimate how carbon loss changed during the pandemic and how much carbon loss has been averted by the integrated livelihood, health, and conservation programs. This will also provide insights for the necessary actions required to mitigate the effects of climate change.

References

G.P. Asner and J. Mascaro. Mapping tropical forest carbon: Calibrating plot estimates to a simple lidar metric. *Remote Sensing of Environment*, 140:614–624, 2014.

L. Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

T. Chen and C. Guestrin. Xgboost: A scalable tree boosting system. *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 1:785–794, 2016.

D.L. Gaveau, B. Locatelli, M.A. Salim, H. Yaen, P. Pacheco, and D. Sheil. Rise and fall of forest loss and industrial plantations in borneo (2000–2017). *Conservation Letters*, 12(3):e12622, 2019.

M.C. Hansen, P.V. Potapov, R. Moore, M. Hancher, S.A. Turubanova, A. Tyukavina, D. Thau, S.V. Stehman, S.J. Goetz, T.R. Loveland, and A. Kommareddy. High-resolution global maps of 21st-century forest cover change. *Science*, 342(6160)):850–853, 2013.

R.M. Haralick, K. Shanmugam, and I.H. Dinstein. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, 6:610–621, 1973.

T. Jucker, G.P. Asner, M. Dalponte, P.G. Brodrick, C.D. Philipson, N.R. Vaughn, Y.A. Teh, C. Brelsford, D.F. Burslem, N.J. Deere, and R.M. Ewers. Estimating aboveground carbon density and its uncertainty in borneo's structurally complex tropical forests using airborne laser scanning. *Biogeosciences*, 15(12):3811–3830, 2018.

G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.Y. Liu. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30, 2017.

J. Mascaro, G.P. Asner, D.E. Knapp, T. Kennedy-Bowdoin, R.E. Martin, C. Anderson, M. Higgins, and K.D. Chadwick. A tale of two "forests": Random forest machine learning aids tropical forest carbon mapping. *PloS one*, 9(1):e85993, 2014.

L. Melendy, S. Hagen, F.B. Sullivan, T. Pearson, S.M. Walker, P. Ellis, K.A. Kustiyo, O.R. Sambodo, M. Hanson, A.W. Klassen, and M.W. Palace. Cms: Lidar-derived canopy height, elevation for sites in kalimantan, indonesia, 2014. *ORNL DAAC, Oak Ridge, Tennessee, USA*, 2014.

Andrew Ng. A chat with andrew on mlops: From model-centric to data-centric ai. *https://youtu.be/06-AZXmwHjo*, 2021.

T.D. Pham, K. Yoshino, N.N. Le, and D.T. Bui. Estimating aboveground biomass of a mangrove plantation on the northern coast of vietnam using machine learning techniques with an integration of alos-2 palsar-2 and sentinel-2a data. *International Journal of Remote Sensing*, 39(22):7761–7788, 2014.