Leveraging Geospatial Foundation Model to estimate Aboveground Biomass and studying it's effect on forest temperature

Anonymous authors
Paper under double-blind review

1 Abstract

Accurate above-ground biomass (AGB) estimation is critical for assessing forest health and understanding its role in the global carbon cycle. This study shall leverage NASA/IBM's Prithvi 2.0, a geospatial foundation model trained on Harmonized Landsat-Sentinel 2 (HLS) data, to estimate AGB using Sentinel-1 radar and Sentinel-2 multispectral imagery. By training on data covering different forest biomes, we aim to develop a robust and transferable model. We shall also investigate the relationship between AGB and land surface temperature (LST) using ECOSTRESS sensor data to understand forest-climate interactions.

2 Introduction

Above-ground biomass (AGB), defined as the total mass of plant matter above the Earth's surface, is one of the most crucial indicators of forest health globally. Large-scale studies of forests and their impact on the global climate use AGB to quantify changes in landscapes caused by natural phenomena and human activities. As a vital mediator in the carbon cycle, a forest's ability to act as a carbon sink is directly linked to the amount of AGB it holds(Hailemariam Temesgen and Sessions, 2015).

The past century has witnessed the effects of declining forest cover in the form of accelerated global warming and climate change. Consequently, accurate estimation and monitoring of AGB is a key focus of foresters and climate scientists. Although accurate measurement of tree biomass traditionally requires destructive sampling, cutting down the tree, and weighing its components, this approach is environmentally harmful and not feasible for large-scale forest studies(Henry et al., 2010).

To avoid destructive sampling, foresters have developed allometric equations that estimate a tree's biomass using measurable attributes such as canopy height, diameter at breast height (DBH), species, and topographical location(Vorster et al., 2020). These equations have been the standard methodology for calculating and monitoring forest biomass at the forest level.

In recent decades, forest studies have undergone a profound transformation with the advent of various remote sensing technologies, including passive sensors such as optical and multispectral cameras and active sensors such as radar and LiDAR (Gleason and Im, 2011). When mounted on platforms such as UAVs, aircraft, and satellites, these sensors provide a suite of technologies to study the Earth at multiple scales. These advances have facilitated the development of new methodologies for analyzing various forest attributes, including above-ground biomass (AGB)(Fassnacht et al., 2023).

These methodologies typically involve collecting remotely sensed data, deriving key vegetative indices, and applying statistical and machine learning models such as linear regression, gradient boosting, and random forests(Li et al., 2020),(Luo et al., 2024), (Li et al., 2024). These approaches paved the way for more advanced deep learning architectures, including multi-layer perceptrons, convolutional neural networks, and stacked autoencoders (Parvez Rana and Tokola, 2023), (Ali et al., 2024), (Zhang et al., 2019).

While deep learning architectures have shown promising results, training them from scratch still requires substantial amounts of reference data and computational resources. Unfortunately, most Earth observation tasks, including AGB estimation, face significant challenges due to the scarcity of high-quality reference data(Nagai et al., 2020). To address this limitation, the remote sensing community has increasingly adopted geospatial foundation models in recent years. These models are not trained for specific tasks; instead, they ingest vast amounts of input data and learn latent encodings through self-supervised training mechanisms (Jakubik et al., 2023), (Cong et al., 2022), (Tseng et al., 2024) . Once trained, these models can be fine-tuned for various downstream tasks, such as AGB estimation(Muszynski et al., 2024).

In this study, we leverage the latest foundation model, Prithvi 2.0, developed by NASA and IBM, which has been trained on the global Harmonized Landsat-Sentinel 2 (HLS) dataset(Szwarcman et al., 2024). We propose a fine-tuning architecture (illustrated in Figure 1) that estimates forest biomass using Sentinel-1 radar data and Sentinel-2 multispectral data. The architecture and training pipeline are discussed in more detail in the next section.

3 Proposed Work

Task 1 - Estimation of above-ground biomass using satellite imagery
The biomass contained in a tree is influenced by several factors, ranging from structural attributes like canopy height and diameter at breast height (DBH) to environmental variables such as location and topography, as well as specific characteristics like species and tree health. Previous studies have demonstrated that spectral signatures captured by multispectral sensors provide information about tree species and health, while radar data can effectively infer tree structure and topography.

In this study, we aim to estimate above-ground biomass (AGB) using globally reaching wall-to-wall data from satellite sensors onboard Sentinel-1 and Sentinel-2. Our proposed architecture 1 leverages the geospatial foundation model, Prithvi 2.0, to process multispectral time-series data from Sentinel-2. Trained on 4.2 million global time-series samples from NASA/ESA's Harmonized Landsat-Sentinel 2 (HLS) dataset, Prithvi 2.0 incorporates both temporal and spatial embeddings for satellite imagery. The frozen encoder of Prithvi 2.0 extracts latent representations and features from Sentinel-2 imagery, which are sent through an untrained decoder.

Parallelly, we process Sentinel-1 time-series imagery using an untrained U-Net architecture with a temporal attention encoder (U-TAE)(Garnot and Landrieu, 2022), inspired by the winning model of the Biomassmasters Data Challenge(Nascetti et al., 2023). Then, the outputs from the two pipelines are concatenated and sent through a single 3X3 convolutional kernel to get the AGB map. The input for each sensor is limited to four time stamps, adhering to Prithvi 2.0's design specifications.

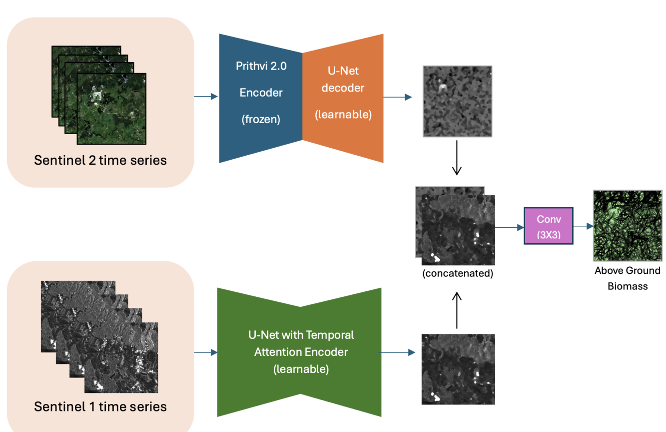

Figure 1: Our proposed architecture

The use of a foundation model like Prithvi 2.0 alleviates the data-intensive nature of training deep learning architectures from scratch. By freezing the weights of Prithvi 2.0's encoder, the learnable parameters in our model are restricted to the decoder for Prithvi, the U-TAE network for Sentinel 1and the final 3X3 kernel. However, careful selection of training data remains crucial. As AGB varies significantly with tree species and environmental biomes, building a transferable model poses a challenge.

For this study, we utilize three main datasets: the Finland forest benchmark dataset (Biomassmasters)(Nascetti et al., 2023), the Central Africa–South Asia biomass dataset by Rodda et al.(Rodda et al., 2024) and GEDI 4A data product for United States (Dubayah et al., 2022). The first two datasets were developed using allometric equations applied to tree inventory and LiDAR surveys. Together, they provide AGB maps across diverse environmental and geographical regions. Conversely, GEDI is a fullwaveform satellite LiDAR sensor that provides multiple tree metrices to study forests, including AGB at its footprint level. These datasets span three different forest biomes — boreal forests in Finland, tropical forests in Central Africa and South Asia and temperate forests in United States.

Task 2 - Studying the Correlation Between Above-Ground Biomass and Forest Temperature
Global temperatures have been steadily rising over the past two centuries, a trend strongly linked to increasing human activity and development. Concurrently, global forest coverage has been declining, primarily due to land conversion for human settlement and extensive deforestation for timber. The reduction in forest cover amplifies the concentration of greenhouse gases in the atmosphere, further accelerating the rise in global temperatures.

Interestingly, several studies have suggested that increasing temperatures also affect the health and growth of forests, often quantified by changes in above-ground biomass (AGB) (Zhao et al., 2019), (Fu and Sun, 2022). Some research have also attempted to model AGB using land surface temperature (LST) and other environmental factors(Rosas-Chavoya et al., 2023), (Jiang et al., 2021).

Building upon the AGB estimation model developed in this study, we aim to investigate the relationship between AGB and LST. Our hypothesis is that LST is influenced by AGB and visa versa and that this relationship can be utilized to study temperature patterns across forested regions. To test this hypothesis, we plan to use data collected by NASA JPL's ECOSTRESS sensor, which operates aboard the International Space Station (ISS). ECOSTRESS provides high-resolution measurements of LST and emissivity at a spatial resolution of 70 meters, offering valuable insights into localized temperature variations in forested areas.

By combining AGB estimates with LST data, this study aims to contribute to a deeper understanding of the interactions between forest biomass and temperature, potentially informing strategies for forest conservation and climate change mitigation.

4 Conclusion and Pathway to Climate Impact

In this study, we have proposed a methodology to accurately estimate AGB using open-source satellite imagery from Sentinel 1 and Sentinel 2. Embracing the rise of foundation models for geospatial data in last two years, we incorporate NASA/IBM's latest foundation model Prithvi 2.0 in our model, to train using the limitedly available biomass data. Our work here has a direct link to climate impact, providing a scalable and ubiquitously applicable model to estimate biomass, improving our understanding of forests. It has application in forestry, providing a way to timely monitor AGB, and can also be used by policy makers to quantify the impact of deforestation activity and restoration policy. Furthermore, our objective is to explore the potential correlation between AGB and forest temperature patterns using data from NASA JPL's ECOSTRESS sensor. Understanding this relationship can provide critical insights into how forests interact with changing climate conditions, further supporting strategies for climate change mitigation and adaptation.

References

Hassan Ali, Jahangir Mohammadi, and Shaban Shataee Jouibary. Deep and machine learning prediction of forest above-ground biomass using multi-source remote sensing data in coniferous planted forests in iran. *European Journal of Forest Research*, 143(6):1731–1745, 2024. ISSN 1612-4677. doi: 10.1007/s10342-024-01721-w. URL `https://doi.org/10.1007/s10342-024-01721-w`.

Yezhen Cong, Samar Khanna, Chenlin Meng, Patrick Liu, Erik Rozi, Yutong He, Marshall Burke, David B. Lobell, and Stefano Ermon. SatMAE: Pre-training transformers for temporal and multi-spectral satellite imagery. In Alice H. Oh, Alekh Agarwal, Danielle Belgrave, and Kyunghyun Cho, editors, *Advances in Neural Information Processing Systems*, 2022. URL `https://openreview.net/forum?id=WBhqzpF6KYH`.

R.O. Dubayah, J. Armston, J.R. Kellner, L. Duncanson, S.P. Healey, P.L. Patterson, S. Hancock, H. Tang, J. Bruening, M.A. Hofton, J.B. Blair, and S.B. Luthcke. Gedi l4a footprint level aboveground biomass density, version 2.1, 2022. URL `https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds`$_id = 2056$.

Fabian Ewald Fassnacht, Joanne C White, Michael A Wulder, and Erik Næsset. Remote sensing in forestry: current challenges, considerations and directions. *Forestry: An International Journal of Forest Research*, 97(1):11–37, 05 2023.

Gang Fu and Wei Sun. Temperature sensitivities of vegetation indices and aboveground biomass are primarily linked with warming magnitude in high-cold grasslands. *Science of The Total Environment*, 843:157002, 2022. ISSN 0048-9697. doi: https://doi.org/10.1016/j.scitotenv.2022.157002. URL `https://www.sciencedirect.com/science/article/pii/S0048969722040992`.

Vivien Sainte Fare Garnot and Loic Landrieu. Panoptic segmentation of satellite image time series with convolutional temporal attention networks, 2022. URL `https://arxiv.org/abs/2107.07933`.

Colin J. Gleason and Jungho Im. A review of remote sensing of forest biomass and biofuel: Options for small-area applications. *GIScience & Remote Sensing*, 48(2):141–170, 2011. doi: 10.2747/1548-1603.48.2.141.

Krishna Poudel Andrew Gray Hailemariam Temesgen, David Affleck and John Sessions. A review of the challenges and opportunities in estimating above ground forest biomass using tree-level models. *Scandinavian Journal of Forest Research*, 30(4):326–335, 2015. doi: 10.1080/02827581.2015.1012114.

M. Henry, A. Besnard, W.A. Asante, J. Eshun, S. Adu-Bredu, R. Valentini, M. Bernoux, and L. Saint-André. Wood density, phytomass variations within and among trees, and allometric equations in a tropical rainforest of africa. *Forest Ecology and Management*, 260 (8):1375–1388, 2010. ISSN 0378-1127. doi: https://doi.org/10.1016/j.foreco.2010.07.040. URL `https://www.sciencedirect.com/science/article/pii/S037811271000424X`.

Johannes Jakubik, Sujit Roy, C. E. Phillips, Paolo Fraccaro, Denys Godwin, Bianca Zadrozny, Daniela Szwarcman, Carlos Gomes, Gabby Nyirjesy, Blair Edwards, Daiki Kimura, Naomi Simumba, Linsong Chu, S. Karthik Mukkavilli, Devyani Lambhate, Kamal Das, Ranjini Bangalore, Dario Oliveira, Michal Muszynski, Kumar Ankur, Muthukumaran Ramasubramanian, Iksha Gurung, Sam Khallaghi, Hanxi, Li, Michael Cecil, Maryam Ahmadi, Fatemeh Kordi, Hamed Alemohammad, Manil Maskey, Raghu Ganti, Kommy Weldemariam, and Rahul Ramachandran. Foundation models for generalist geospatial artificial intelligence, 2023. URL `https://arxiv.org/abs/2310.18660`.

Fugen Jiang, Mykola Kutia, Kaisen Ma, Song Chen, Jiangping Long, and Hua Sun. Estimating the aboveground biomass of coniferous forest in northeast china using spectral variables, land surface temperature and soil moisture. *Science of The Total Environment*, 785:147335, 2021. ISSN 0048-9697. doi: https://doi.org/10.1016/j.scitotenv.2021.147335. URL `https://www.sciencedirect.com/science/article/pii/S0048969721024062`.

Huaqiang Li, Fei Li, Jingfeng Xiao, Jiquan Chen, Kejian Lin, Gang Bao, Aijun Liu, and Guo Wei. A machine learning scheme for estimating fine-resolution grassland aboveground biomass over china with sentinel-1/2 satellite images. *Remote Sensing of Environment*, 311:114317, 2024. ISSN 0034-4257. doi: https://doi.org/10.1016/j.rse.2024.114317. URL https://www.sciencedirect.com/science/article/pii/S0034425724003353.

Yingchang Li, Mingyang Li, Chao Li, and Zhenzhen Liu. Forest aboveground biomass estimation using landsat 8 and sentinel-1a data with machine learning algorithms. *Scientific Reports*, 10(1):9952, 2020. ISSN 2045-2322. doi: 10.1038/s41598-020-67024-3. URL https://doi.org/10.1038/s41598-020-67024-3.

Mi Luo, Shoaib Ahmad Anees, Qiuyan Huang, Xin Qin, Zhihao Qin, Jianlong Fan, Guangping Han, Liguo Zhang, and Helmi Zulhaidi Mohd Shafri. Improving forest above-ground biomass estimation by integrating individual machine learning models. *Forests*, 15(6), 2024. ISSN 1999-4907. URL https://www.mdpi.com/1999-4907/15/6/975.

Michal Muszynski, Levente Klein, Ademir Ferreira da Silva, Anjani Prasad Atluri, Carlos Gomes, Daniela Szwarcman, Gurkanwar Singh, Kewen Gu, Maciel Zortea, Naomi Simumba, Paolo Fraccaro, Shraddha Singh, Steve Meliksetian, Campbell Watson, Daiki Kimura, and Harini Srinivasan. Fine-tuning of geospatial foundation models for aboveground biomass estimation, 2024. URL https://arxiv.org/abs/2406.19888.

Shin Nagai, Kenlo Nishida Nasahara, Tomoko Kawaguchi Akitsu, Taku M. Saitoh, and Hiroyuki Muraoka. *Importance of the Collection of Abundant Ground-Truth Data for Accurate Detection of Spatial and Temporal Variability of Vegetation by Satellite Remote Sensing*, chapter 11, pages 223–244. American Geophysical Union (AGU), 2020. ISBN 9781119413332. doi: https://doi.org/10.1002/9781119413332.ch11. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/9781119413332.ch11.

Andrea Nascetti, RITU YADAV, Kirill Brodt, Qixun Qu, Hongwei Fan, Yuri Shendryk, Isha Shah, and Christine Chung. Biomassters: A benchmark dataset for forest biomass estimation using multi-modal satellite time-series. In *Thirty-seventh Conference on Neural Information Processing Systems Datasets and Benchmarks Track*, 2023. URL https://openreview.net/forum?id=hrWsIC4Cmz.

Anne Tolvanen Basanta Gautam Shruthi Srinivasan Parvez Rana, Sorin Popescu and Timo Tokola. Estimation of tropical forest aboveground biomass in nepal using multiple remotely sensed data and deep learning. *International Journal of Remote Sensing*, 44(17): 5147–5171, 2023. doi: 10.1080/01431161.2023.2240508.

Suraj Reddy Rodda, Rakesh Fararoda, Rajashekar Gopalakrishnan, Nidhi Jha, Maxime Réjou-Méchain, Pierre Couteron, Nicolas Barbier, Alonso Alfonso, Ousmane Bako, Patrick Bassama, Debabrata Behera, Pulcherie Bissiengou, Hervé Biyiha, Warren Y. Brockelman, Wirong Chanthorn, Prakash Chauhan, Vinay Kumar Dadhwal, Gilles Dauby, Vincent Deblauwe, Narcis Dongmo, Vincent Droissart, Selvaraj Jeyakumar, Chandra Shekar Jha, Narcisse G. Kandem, John Katembo, Ronald Kougue, Hugo Leblanc, Simon Lewis, Moses Libalah, Maya Manikandan, Olivier Martin-Ducup, Germain Mbock, Hervé Memiaghe, Gislain Mofack, Praveen Mutyala, Ayyappan Narayanan, Anuttara Nathalang, Gilbert Oum Ndjock, Fernandez Ngoula, Rama Rao Nidamanuri, Raphaël Pélissier, Sassan Saatchi, Le Bienfaiteur Sagang, Patrick Salla, Murielle Simo-Droissart, Thomas B. Smith, Bonaventure Sonké, Tariq Stevart, Danièle Tjomb, Donatien Zebaze, Lise Zemagho, and Pierre Ploton. Lidar-based reference aboveground biomass maps for tropical forests of south asia and central africa. *Scientific Data*, 11(1):334, 2024. ISSN 2052-4463. doi: 10.1038/s41597-024-03162-x. URL https://doi.org/10.1038/s41597-024-03162-x.

Marcela Rosas-Chavoya, Pablito Marcelo López-Serrano, Daniel José Vega-Nieva, José Ciro Hernández-Díaz, Christian Wehenkel, and José Javier Corral-Rivas. Estimating aboveground biomass from land surface temperature and evapotranspiration data at the temperate forests of durango, mexico. *Forests*, 14(2), 2023. ISSN 1999-4907. URL https://www.mdpi.com/1999-4907/14/2/299.

Daniela Szwarcman, Sujit Roy, Paolo Fraccaro, orsteinn Elí Gíslason, Benedikt Blumenstiel, Rinki Ghosal, Pedro Henrique de Oliveira, Joao Lucas de Sousa Almeida, Rocco Sedona, Yanghui Kang, Srija Chakraborty, Sizhe Wang, Ankur Kumar, Myscon Truong, Denys Godwin, Hyunho Lee, Chia-Yu Hsu, Ata Akbari Asanjan, Besart Mujeci, Trevor Keenan, Paulo Arevalo, Wenwen Li, Hamed Alemohammad, Pontus Olofsson, Christopher Hain, Robert Kennedy, Bianca Zadrozny, Gabriele Cavallaro, Campbell Watson, Manil Maskey, Rahul Ramachandran, and Juan Bernabe Moreno. Prithvi-eo-2.0: A versatile multi-temporal foundation model for earth observation applications, 2024. URL https://arxiv.org/abs/2412.02732.

Gabriel Tseng, Ruben Cartuyvels, Ivan Zvonkov, Mirali Purohit, David Rolnick, and Hannah Kerner. Lightweight, pre-trained transformers for remote sensing timeseries, 2024. URL https://arxiv.org/abs/2304.14065.

Anthony G. Vorster, Paul H. Evangelista, Atticus E. L. Stovall, and Seth Ex. Variability and uncertainty in forest biomass estimates from the tree to landscape scale: the role of allometric equations. *Carbon Balance and Management*, 15(1):8, 2020. ISSN 1750-0680. doi: 10.1186/s13021-020-00143-6. URL https://doi.org/10.1186/s13021-020-00143-6.

Linjing Zhang, Zhenfeng Shao, Jianchen Liu, and Qimin Cheng. Deep learning based retrieval of forest aboveground biomass from combined lidar and landsat 8 data. *Remote Sensing*, 11(12), 2019. ISSN 2072-4292. URL https://www.mdpi.com/2072-4292/11/12/1459.

Jingxue Zhao, Tianxiang Luo, Haixia Wei, Zhaoheng Deng, Xiang Li, Ruicheng Li, and Yanhong Tang. Increased precipitation offsets the negative effect of warming on plant biomass and ecosystem respiration in a tibetan alpine steppe. *Agricultural and Forest Meteorology*, 279:107761, 2019. ISSN 0168-1923. doi: https://doi.org/10.1016/j.agrformet.2019.107761. URL https://www.sciencedirect.com/science/article/pii/S0168192319303776.