Modeling Worldwide Tree Biodiversity Using Canopy Structure Metrics from Global Ecosystem Dynamics Investigation (GEDI) data

Jin XU[1], Kjirsten Coleman[1], Volker C. Radeloff[2], Melissa Songer[1], Qiongyu Huang[1]

1 Conservation Biology Institute, National Zoological Park, Smithsonian Institution, Front Royal, VA 22630, USA
2 Forest and Wildlife Ecology Department, University of Wisconsin-Madison, 1630 Linden Drive, Madison, WI 53706

NASA NIP Project: The power of GEDI: Investigate the efficacy of spaceborne Lidar to model biodiversity and characterize habitat heterogeneity at the continental and global scales. PI: Qiongyu Huang. Proposal/Award Number: 80NSSC21K0936

Introduction

- ❑ Biodiversity of tree species within forest systems has an effect on productivity, ecosystem resilience and function (Wang and Gamon, 2019).
- ❑ It is essential to quantify tree species richness to understand and manage forest ecosystems over broad scales (Wang and Gamon, 2019), including and especially, at a global scale.
- ❑ The launch of NASA's Global Ecosystem Dynamics Investigation (GEDI) in December of 2018 provided new possibilities for exploring tree species richness at a global scale (Dubayah et al., 2020).
- ❑ In this study, we focused on exploring the capacity for using forests' and tree species' unique spectral and structural characteristics for predicting tree species richness.

Objective

- ❑ What is the efficacy of space-borne lidar metrics in predicting global tree species richness?
- ❑ What is the capacity of the GEDI-based model in predicting tree species richness in different climate zones?
- ❑ To what extent do GEDI metrics improve a tree species richness model based on spectral vegetation metrics alone?

Data Source

➢ ForestGEO dataset

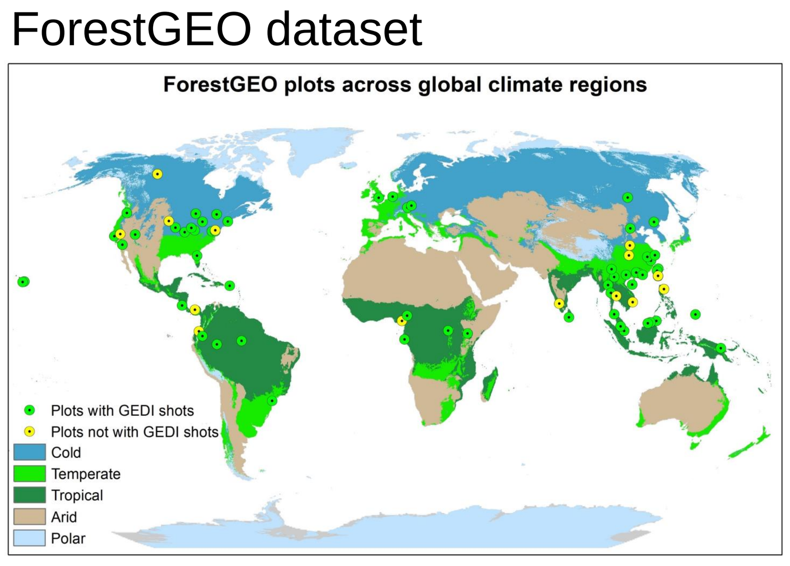

Figure 1. Distribution map of 74 ForestGEO plots across climate zones and global regions. Lighter green circles denote plots that have GEDI shots (N = 60; cold: n = 15, temperate: n = 21, tropical: n = 24) and yellow circles denote plots that do not have GEDI shots (N=14; cold: n = 3, temperate: n = 4, tropical: n = 7).

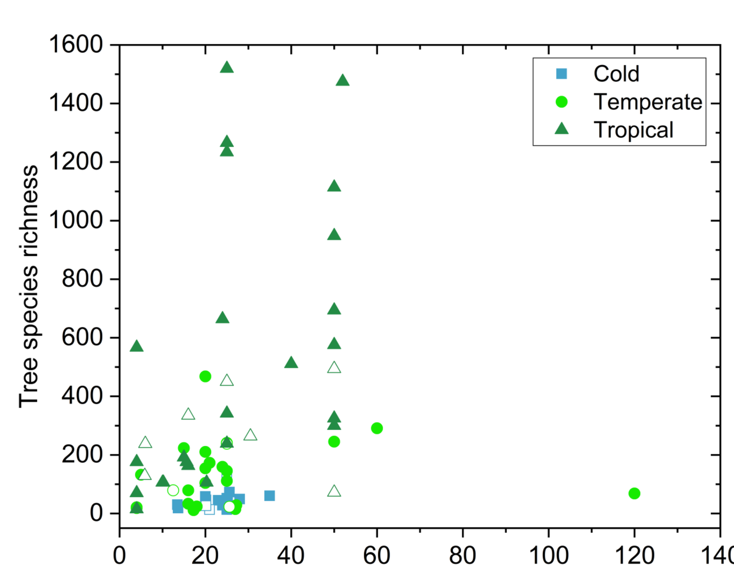

Figure 2. ForestGEO site plot area vs the corresponding number of species. The hollowed shapes show the plots not covered by GEDI shots.

➢ NEON dataset

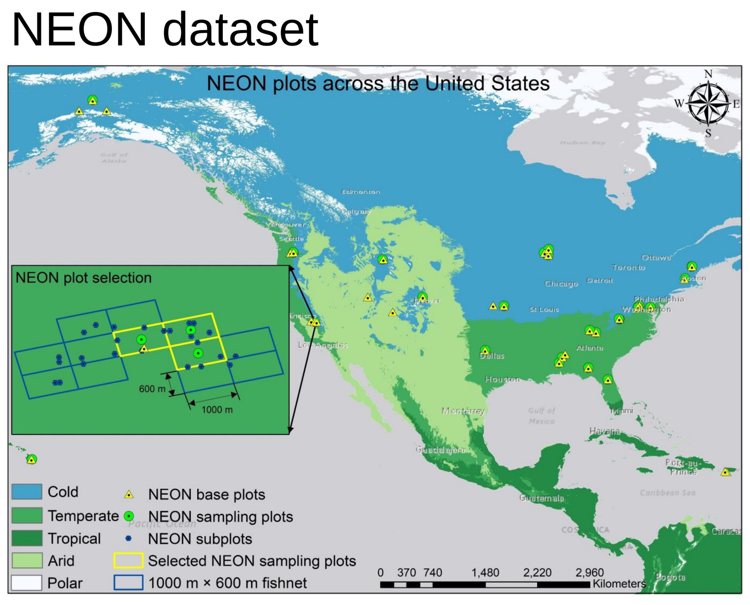

Figure 3. 35 NEON plot distribution map across climate zones and continental United States, Alaska, Hawaii, and Puerto Rico. Yellow circles represent base plots (N = 35), green circles represent sampling plots (N = 48), and blue circles represent subplots (N = 723). Yellow polygons are the selected NEON sampling plots and blue polygons illustrate the 1000 m × 600 m fishnet.

➢ GEDI shot filtration

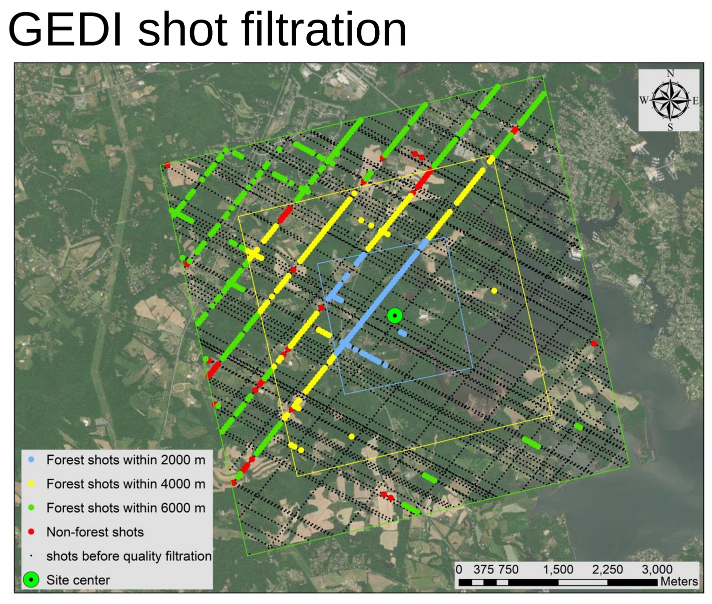

Figure 4. ForestGEO example plot located at Smithsonian Environmental Research Center (Edgewater, MD) with 2000 m (blue), 4000 m (yellow), and 6000 m (green) pixel sizes. GEDI shots have been filtered for quality and masked for landcover type. Filtered GEDI shots appear as a point in the associated pixel size color.

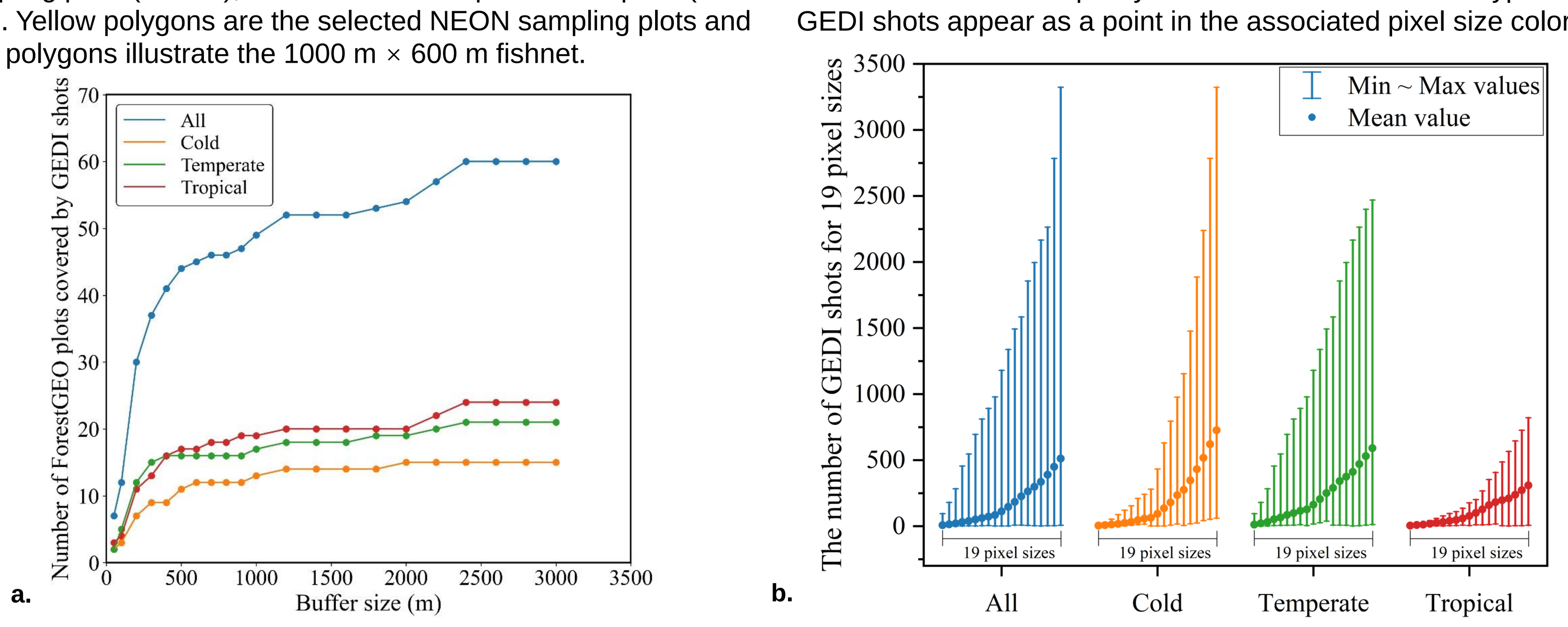

Figure 5. (a) The number of ForestGEO plots covered by GEDI shots across 19-pixel sizes (400, 600, 800, 1000, 1200, 1400, 1600, 1800, 2000, 2400, 2800, 3200, 3600, 4000, 4400, 4800, 5200, 5600, and 6000- resolution; meter), and (b) The minimum, maximum, and mean of the number of ForestGEO plots covered by GEDI shots across 19-pixel sizes and climate zones.

Table 1. List of metrics

Metric categories	Metric name
Fixed predictor (1)	Plot size (ha)
GEDI metrics (16)	$RH100_{mean}$, $RH100_{std}$, PAI_{mean}, PAI_{std}, $Cover_{mean}$, $Cover_{std}$, FHD_{mean}, FHD_{std}, N_layer_{mean}, N_layer_{std}, $PAVD\_ratio_{mean}$, $PAVD\_ratio_{std}$, PAI_ratio_{mean}, PAI_ratio_{std}, $Cover\_ratio_{mean}$, $Cover\_ratio_{std}$
Spectral vegetation metrics (3)	$DHIs\text{-}NDVI_{cum}$, $DHIs\text{-}NDVI_{min}$, $DHIs\text{-}NDVI_{var}$

* Reference for metric names: standard deviation (std), relative height (RH100), plant area index (PAI), total canopy cover (Cover), foliage height diversity (FHD), the number of canopy layers (N_layer), a vertical plant area volume density ratio (PAVD_ratio), a vertical PAI ratio (PAI_ratio), a vertical cover ratio (Cover_ratio), dynamic habitat indices (DHIs), normalized difference vegetation index (NDVI), cumulative (cum), minimum (min), variation (var).

Table 2. Models for predicting tree species richness using the ForestGEO dataset

Response variable	Models	Predictors
ForestGEO tree species richness	DHIs-only	Plot size + spectral vegetation metrics
	GEDI-only	Plot size + GEDI metrics
	GEDI-DHIs	Plot size + GEDI metrics + spectral vegetation metrics

Methods

➢ Workflow

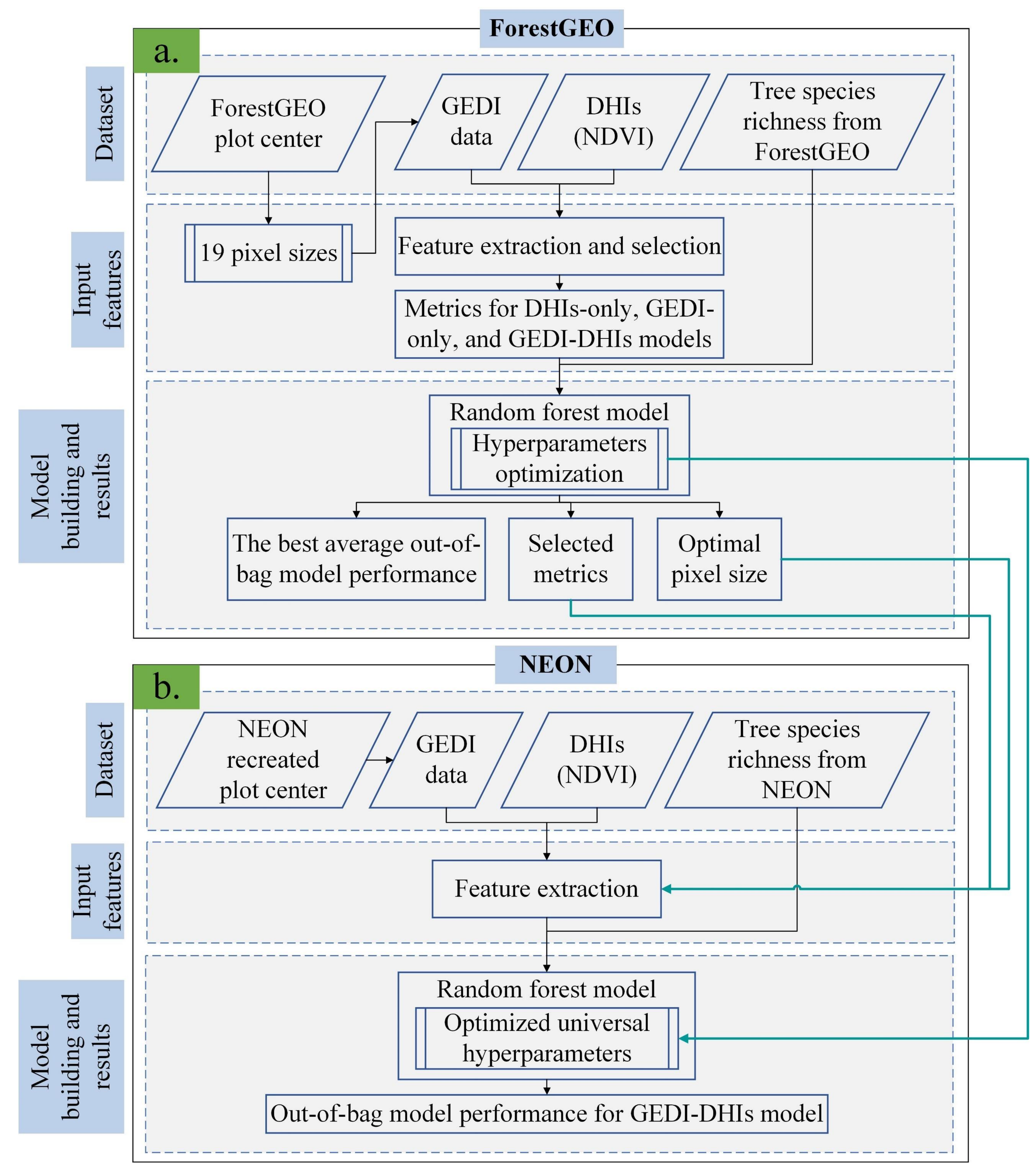

Figure 6. The study workflow. (a) Workflow demonstrating the analysis based on the ForestGEO dataset and GEDI metrics. The analyses included the feature extraction and selection for model building, the hyperparameter optimization, and optimal pixel size selection. (b) The efficacy of GEDI metrics was also evaluated using the NEON dataset based on the optimized universal hyperparameters and optimal pixel size from ForestGEO dataset.

Results

➢ Model performance across pixel sizes and feature importance

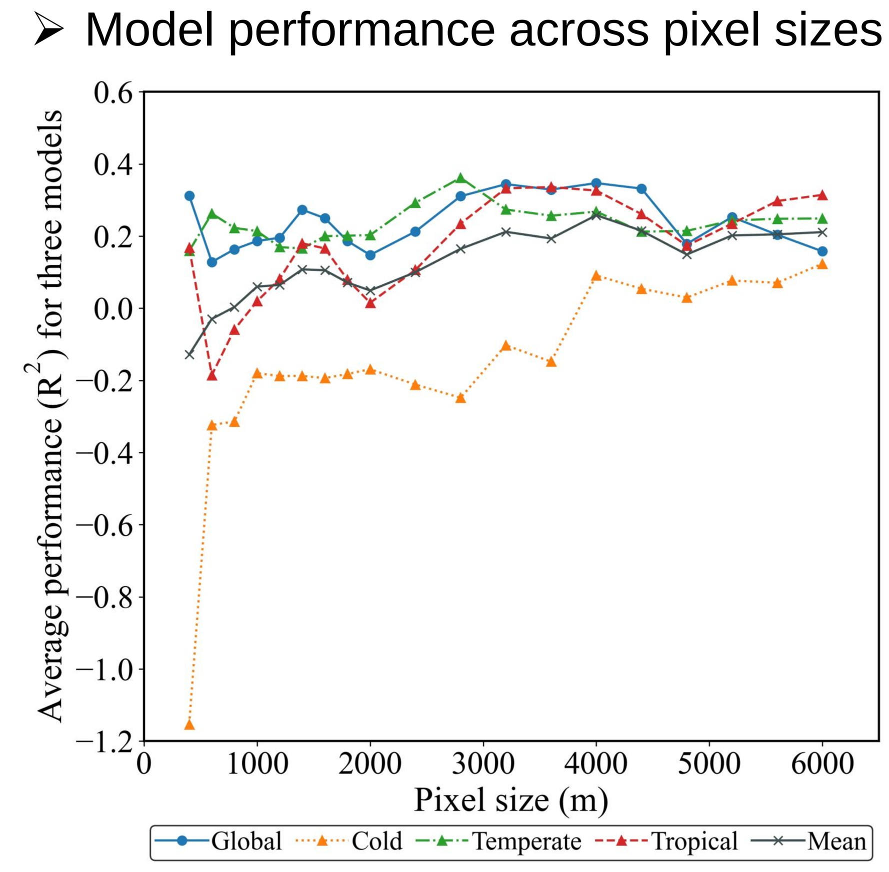

Figure 7. Averaged model performance of DHIs-only, GEDI-only, and GEDI-DHIs models for each climate zone and each pixel size based on optimized universal hyperparameters.

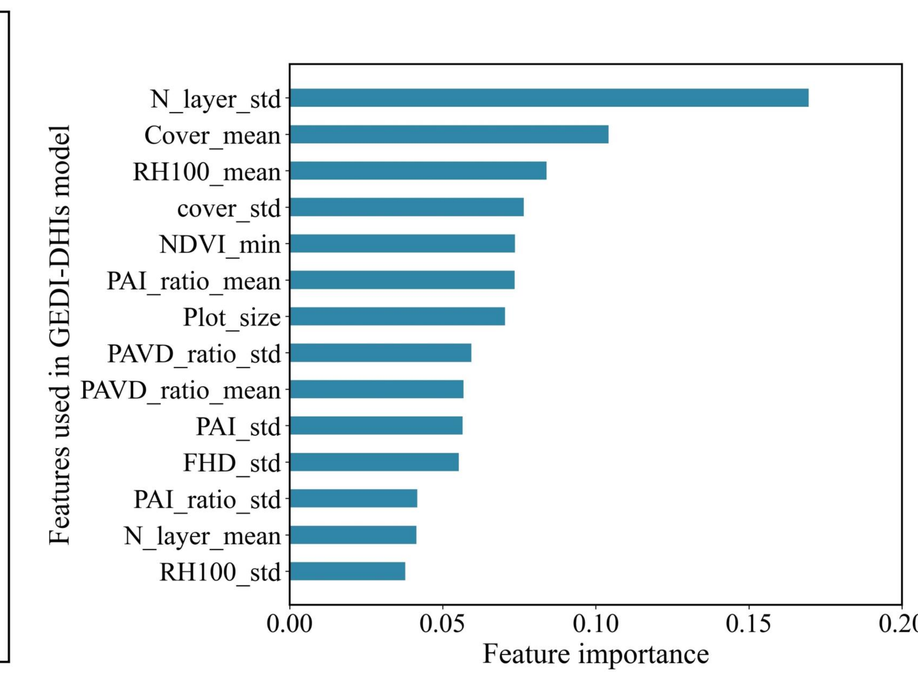

Figure 8. Feature importance from the model with the best performance (GEDI-DHIs). Reference for metric names: the number of canopy layers (N_layer), relative height (RH100), plant area index (PAI), plant area volume density (PAVD), foliage height diversity (FHD), standard deviation (std).

Table 3. Model performance based on optomized hyperparameters and pixel size (4000 m)

Model Performance	DHIs-only (ForestGEO)			GEDI-only (ForestGEO)			GEDI-DHIs (ForestGEO)			GEDI-DHIs (NEON)		
	R^2	RMSE	NRMSE	R^2	RMSE	NRMSE	R^2	RMSE	NRMSE	R^2	RMSE	NRMSE
Global	0.24	308.84	21%	0.39	275.17	19%	0.41	271.84	19%	**0.64**	**35.36**	**13%**
Cold	0.05	16.29	27%	0.10	15.81	26%	0.12	15.69	26%	**0.47**	**25.78**	**20%**
Temperate	0.09	111.84	24%	0.35	94.33	21%	0.37	92.85	20%	**0.77**	**37.02**	**14%**
Tropical	0.28	371.82	25%	0.37	348.15	24%	0.32	361.10	25%	-	-	-

Conclusions and future work

- ❑ GEDI-DHIs model using the ForestGEO dataset performed best for predicting tree species richness, followed by GEDI-only and DHIs-only models.
- ❑ A 4000 m pixel size that can be used in future was the optimal extent to quantify and aggregate GEDI metrics for predicting tree species richness.
- ❑ The method provides a new avenue to help introduce more metrics for tree species richness across climate zones in future research and support for forest conservation management.
- ❑ With the increasing availability of GEDI data, expanding field data range with normal distribution is becoming a critical factor for developing high accuracy forest richness models.
- ❑ Models using higher spectral resolution (e.g., hyperspectral images) metrics, which has numerous narrow bands, could also be used.

References

- Wang, R., Gamon, J.A., 2019. Remote sensing of terrestrial plant biodiversity. *Remote Sensing of Environment*, 231: 111218.
- Scherer-Lorenzen, M., Luis Bonilla, J., Potvin, C., 2007. Tree species richness affects litter production and decomposition rates in a tropical biodiversity experiment. *Oikos*, 116(12): 2108-2124.
- 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: 100002.