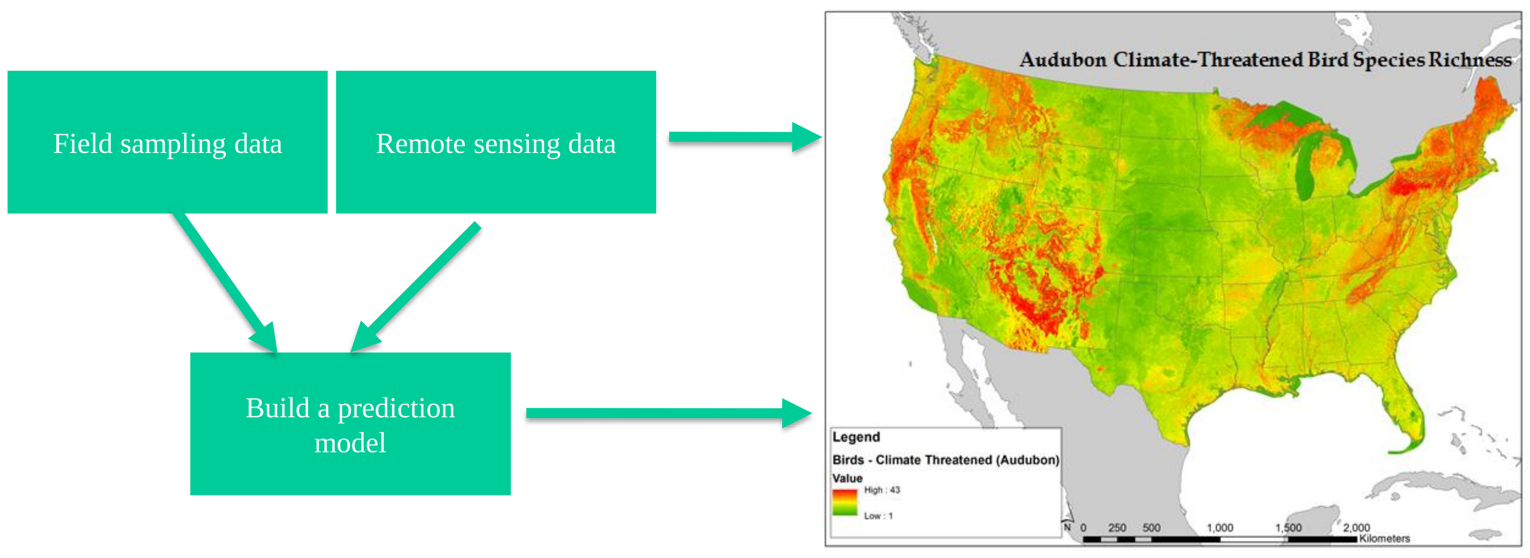
Spatial prediction of bird species richness in eastern USA using satellite imagery and ecological observation data

小幡　進午 (Shingo Obata)　　リスク解析戦略研究センター　特任助教

Introduction

Bird Species Richness: Brid species richness, the number of bird species in a given spatial extent, is one of the most widely used measures of biodiversity. Vegetation structure is one of the environmental factors that influence bird species richness. Prediction of bird species richness with 3D structure is expected to improve the accuracy since 3D vertical structure represented by functional traits is important for the prediction of bird species. These metrics are usually retrieved from airborne laser scanning (ALS) point cloud data. However, it is challenging to estimate vegetation structure over large areas due to the unavailability of ALS data. Global Ecological Dynamics Investigation(GEDI), a spaceborne Light Detection And Ranging (LiDAR) launched in 2019, could be a potential substitute for the in-situ ALS point cloud data, enabling the estimation of bird species richness at broad geographic extents.

Objective This study aims at testing a method to estimate the bird species richness using interpolated GEDI Level 2-A and Level 2-B products. To achieve this goal, we first assessed the accuracy of the relative heights of vegetation from GEDI Level 2-A and functional traits from GEDI Level 2-B. Then we created a wall-to-wall relative height raster and functional traits raster from GEDI. We used a random forest model to fit bird species richness from National Ecological Observatory Network (NEON) field plot data with vegetation structure variables derived from NEON ALS data and separately with GEDI relative height and functional traits variables.



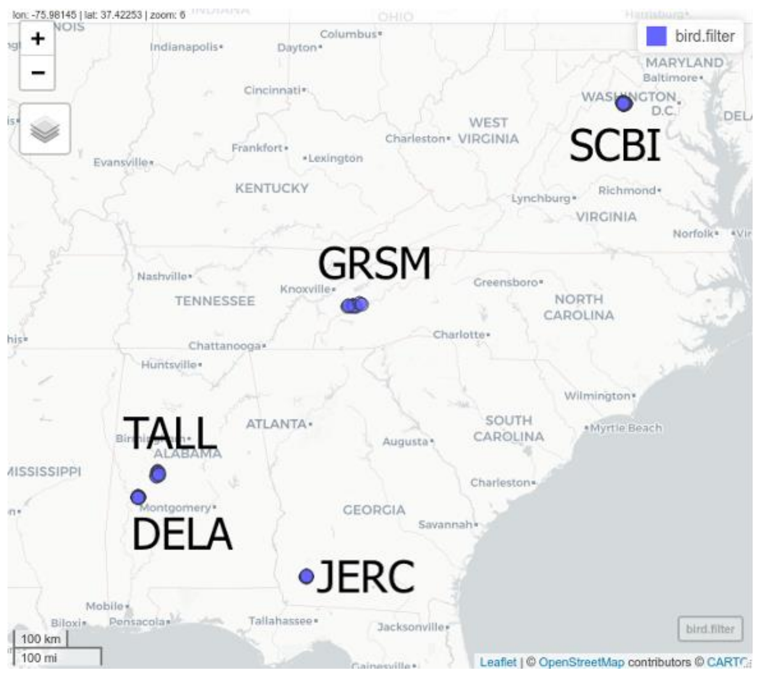
Research overview

Method and Data

Study area: Our study area covers five NEON domains in the eastern United States (Figure 1). Latitude ranges between 29.7°N and 38.9°N. Longitude ranges between -88.2°W and -76.5°W. Eastern US is represented by nine NEON domains; t we excluded 4 domains for which discrete return LiDAR data were not available. **Data:** We used the 2019-2021 NEON Breeding landbird point counts from nine observation sites within the five NEON domains to calculate bird species richness, the response variable for our random forest model. We calculated the Shannon diversity index (bird species richness) from the land bird point counts for each observation site. GEDI data was downloaded and processed. We selected data acquired in the 2018 leaf-on season and 2019 leaf-on season. For each bird observation site, we selected vegetation relative heights of every 10 percentile (rh0, rh10, rh20, rh30, rh40, rh50, rh60, rh70, rh80, rh90), as well as the 95 percentile (rh95), foliage height diversity (FHD), Plant Area Index (PAI), and gap percentage as possible predictors of bird species richness. We first clipped the NEON ALS point cloud into the selected footprint extent (250 m diameter). We calculated leaf area index (LAI) and FHD from NEON ALS data using a 3-D based method, in which point cloud is split into 1 m^3 voxel. LAI incorporates gap fraction (Pi), calculated for each horizontal layer by counting the number of points and dividing the number of the returns below a specified height threshold by the total number of the returns. In the case of the point cloud acquired from discrete return LiDAR, gap fractions were calculated by assessing the number of laser beams that reached the layer *z* and penetrated the layer *z*+*dz*:

$$P_i = \frac{N_{[0;z]}}{N_{total} - N_{[0;z+dz]}} \qquad LAI_i = -\frac{\ln(P_i)}{k}$$

$N_{[0;z]}$: the number of returns below z, N_{total}: the number of all the returns, and $N_{[0;z+dz]}$: number of returns below *z+dz*. LAI for each horizontal layer was calculated using the probabilistic approach with Beer-Lambert theory for light penetration into the canopy

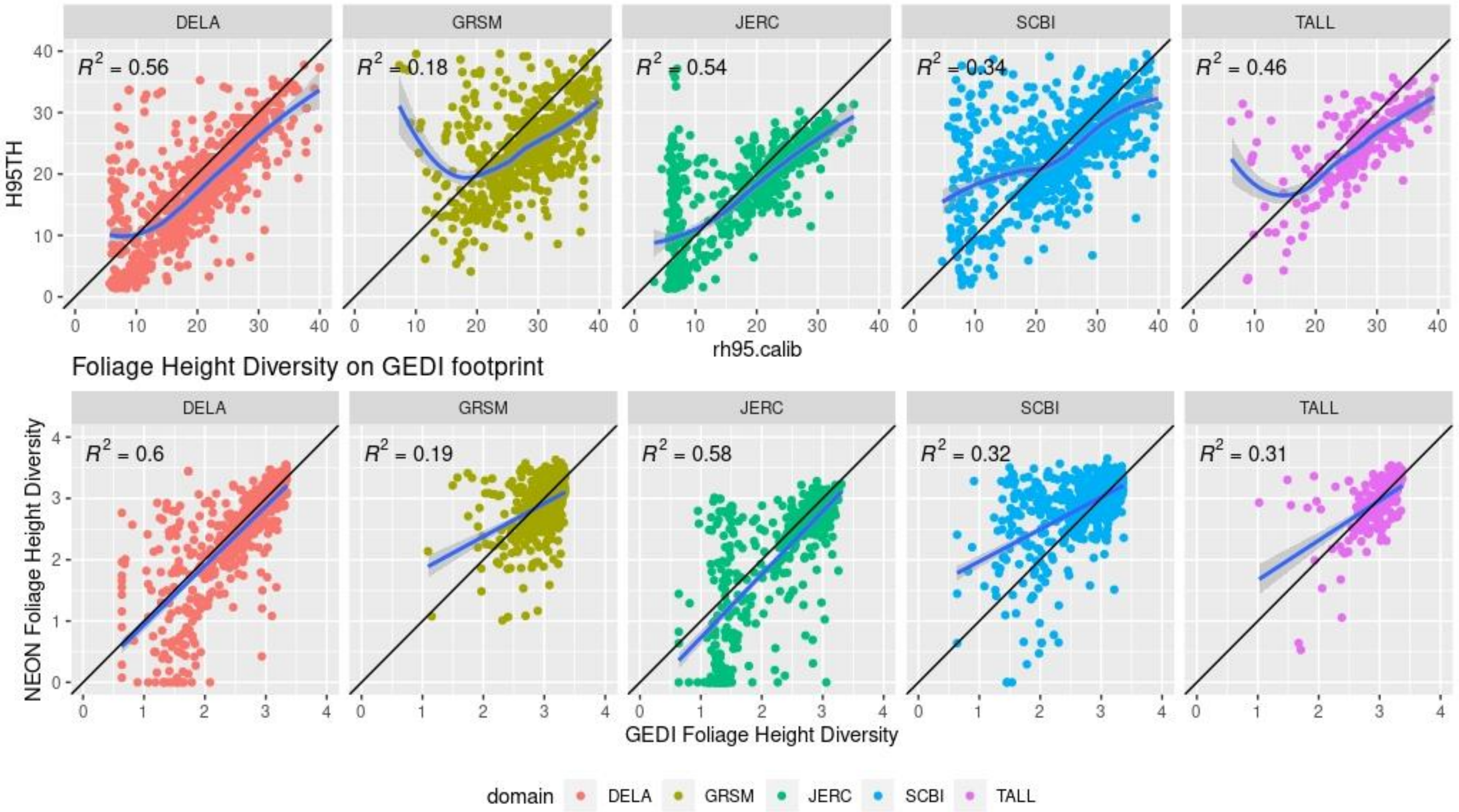


Five study sites in the USA

Results

The comparison of 95% relative height between NEON and GEDI showed an intermediate correlation. R^2 differed by study sites. In the region where forest density is low, the precision of the GEDI's estimation had a larger error. Another comparison of Foliage Height Diversity between NEON and GEDI resulted in lower R^2 values than 95% relative height.

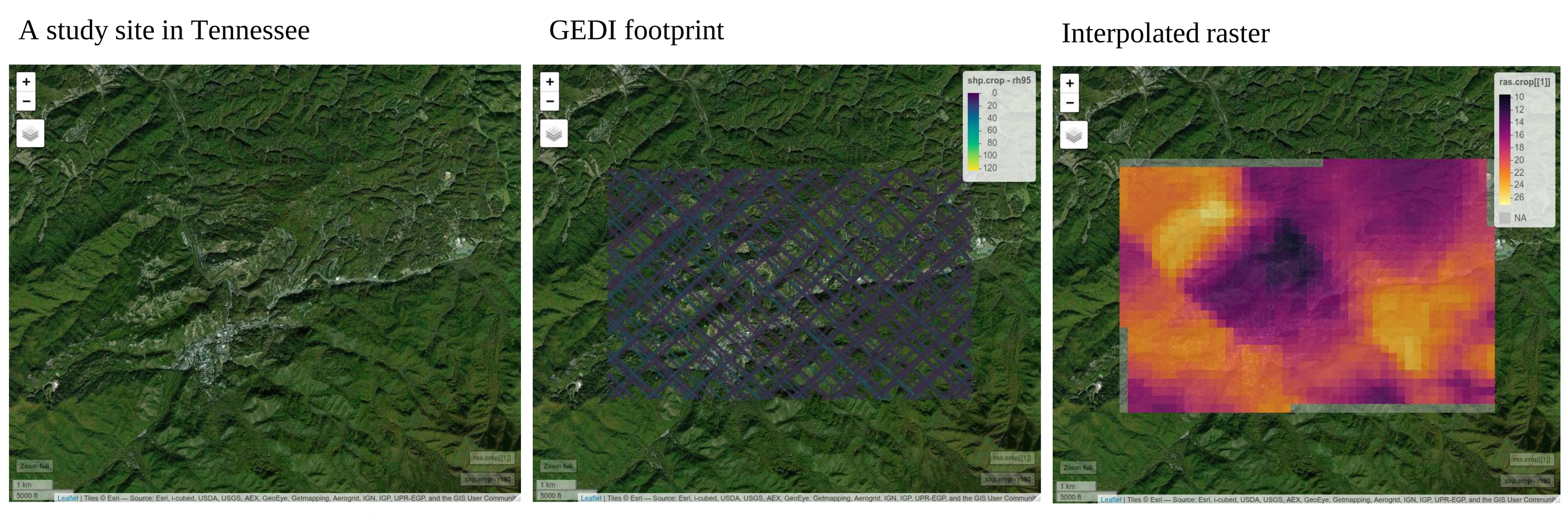
We built initial bird species richness prediction model from GEDI data and NEON Breeding landbird point counts. Accuracy assessment of the model showed 10 % point lower OOB R^2 (0.3) compared to the previous research that used airborne LiDAR data as the remote sensing data.



Discussion

The difference of R^2 in 95% relative height comparison over study sites could be attributed to the forest cover ratio (forest %) on the sampling point. As forest % increases, the correlation between GEDI and NEON is improved. Another factor that results in lower R^2 is the slope of the sampling point. GEDI's surface height estimation is known to be negatively affected by the slope. Thus, it is required to find a way to correct the negative effect of the slope.

To acquire better OOB R2 from the bird species richness prediction using GEDI, it is possible to (i) correct the slope effect, and (ii) introduce new features as feature variables of the random forest regression model we employed in this research.



Interpolation of GEDI's 95% height to produce raster data

Conclusion

Brid species richness is a key variable in assessing the environmental condition of a given area. Spatial estimation of the metrics had been conducted by using the small-scale LiDAR data, but large-scale analysis had been difficult. In this study, I attempt to estimate bird species richness at the regional level using the GEDI dataset, which provides 3-D information about forests from space. We used the NEON dataset as the source of field sampling data and LiDAR data. The tentative result of this research showed a moderate correlation between GEDI and LiDAR data. OOB R^2 of the bird species richness prediction model was 10 percent points lower than the previous model. Further improvement of the model required.