LiDAR-derived variables as a proxy for functional diversity of woody plant and response to forest management intensity

Project outside of the course scope, 7.5 ETCS

Yaquan Chang (lsc161)

Supervisor: Dr. Naia Morueta-Holme

August 30th, 2019

Table of Contents

Abstract

Land-use intensification can have a significant impact on biodiversity, but it is still unclear how functional richness is affected. Functional richness is an essential feature for biological assemblages, as it links taxonomic richness to the functioning of ecosystems. LiDAR has been detected as a proxy for the biodiversity of some organisms (e.g., fungi) since it could capture vegetation structure (e.g., canopy cover, height) in a fine resolution on the broad scale, but it is still unclear whether it can detect the functional diversity. In this project, I compared LiDAR-derived variables and field-derived variables as explanatory variables in models of functional diversity of woody plants. Then I investigated how forest management intensity impacts species richness and functional diversity of woody plants. My results showed that LiDAR-derived models have a slightly better performance than field-derived models when predicting functional diversity. But it was not able to predict species richness. The functional diversity of seed mass showed a slightly positive relationship along with forest management intensity. While functional diversity of specific leaf area showed no differences across the forest management intensity. I conclude that LiDAR-derived variables are able to predict several functional diversity based on the single trait level. The field inventory could supply some details (e.g., decay level, coarse woody debris) which increases the functional diversity predictions. Forest management intensity likely increases the divergence of seed mass due to high nutrient and diverse habitat but still with some overlaps in different management types. Moreover, it is still unclear how it affects functional diversity in the multidimensional level.

Keywords

Woody plants, functional diversity, LiDAR, vegetation structure

1. Introduction

The biodiversity loss one of the exceed planetary boundaries for earth due to climate change and human disturbance (Rockström et al., 2009). Human activities (e.g., hunting, grazing, and planting; Bengtsson et al., 2000) often cause homogeneous in landscape structure. The species richness may increase due to intermediate disturbance hypothesis (IDH), asserting that moderate levels of human intervention increase the species coexistence, such as introduced species and early successional native species (McKinney et al., 2008). However, human disturbance also pushes species to shift from their current habitat to more suitable habitat, by changing the microhabitat and forest structures. These shifts may interrupt community composition and therefore disrupt ecosystem functioning (Aguirre-Gutiérrez et al., 2017). Therefore, understanding how biodiversity changes ecosystem functioning raises ecologists concern (Cardinale et al., 2000; Naeem et al., 1994; Skidmore and Pettorelli, 2015). Functional diversity plays a pivotal role in linking ecosystem function to biodiversity (Jetz et al., 2016; Laureto et al., 2015) by focusing on the range of traits covered by species assemblages rather than the number of species (Cadotte et al., 2011). High functional diversity with the high dissimilarity of species shows more resilient to environmental changes but not always mean high species richness. For instance, introducing new species into the community will increase the biodiversity, but may decrease the average dissimilarity among species (Botta-Dukát, 2005); on the other hand, assemblages may change functional diversity without changing the species richness.

There are many different frameworks to measure functional diversity (e.g., functional richness, evenness, divergence, dispersion; Laliberté and Legendre, 2010), while the Rao's quadratic entropy (RaoQ; Botta-Dukát, 2005) and community weight metrics (CWM; Grime, 1998) have been widely used in measuring functional diversity (Ricotta and Moretti, 2011). RaoQ usually represents the dispersion pattern by calculating the functional dissimilarity

between two selected individuals in random with replacements (Botta-Dukát, 2005). When comparing to the random expectation, lower RaoQ suggests environmental filtering is central to the assembly of the community (beta niche), whereas a higher RaoQ supports the idea that the limiting similarity mechanism prevents the coexist of too similar species (alpha niche). While CWM is calculated by the mean trait value weighted by relative abundance, relating to the mass ratio hypothesis, assuming the traits of dominant species are vital to the ecosystem process (Grime, 1998; Lavorel et al., 2008).

Woody plants provide a lot of ecosystem services to other organisms, including supplying pollen to butterflies (Aguirre-Gutiérrez et al., 2017), and providing fruit and seeds to birds (Kissling et al., 2008). Thus, investigating the functional diversity of woody plants could give us an insight into other organisms. There are a variety of functional traits measured in woody plants. Seed mass (SM) and specific leaf area (leaf area per leaf dry mass; SLA) are two main plant traits which represent plant fitness in the forest. SM is always considered to be a crucial feature of the plant fitness for the following two main reasons: 1) seed mass has a negative correlation of the number of seeds which indicates the different dispersal strategy of plants (i.e., species with small seed generally produce more seeds at a time to have a ruderals strategy); 2) seed mass is positively correlated with survival of seeding represents the regeneration strategy of different species, since larger seeds usually develop larger seedings to tolerate resources limits (nutrients, lights, etc.). Therefore, due to the trade-off between the seeds number and seeding survival, the species with different seed mass have different life-strategies (Coomes and Grubb, 2003). SLA expresses the transformation of carbon assimilation capacity to photosynthetic surface area per unit mass, associating with leaf lifespan and resource capture (i.e., high SLA tends to be associated with high nitrogen, low lifespan, and high photosynthesis capacity; Cornwell et al., 2006).

European Beech (*Fagus sylvatica*) forest is one of the primary forest types in Europe which has been modified by human intervention for a long time (e.g., timber production, and agriculture, Bengtsson et al., 2000; Brunet et al., 2010). These activities have profound effects on forest structure and biodiversity (Brunet et al., 2010). After realizing this problem, people tried a variety of approaches to manage the forest, trying to counteract biodiversity loss. For instance, some people suggested “near nature management”, which keeps veteran trees and deadwood in the forest to mimic nature (Hahn et al., 2005); others created nature reserve or national park to create undisturbed habitat for wildlife (Esselink et al., 2000). However, it is still unclear how these management types impact on functional diversity although the impact of management intensity on biodiversity has been widely investigated (Burivalova et al., 2014; Putz et al., 2001).

Besides, functional diversity needs a large number of trait measurement, which requires time-consuming inventories and expert knowledge, which would be hard to apply on a broad scale. Some datasets record comprehensive functional traits which are convenient to researchers (Maitner et al., 2018). But only based on the field approach can only cover minor areas (Lelli et al., 2019). Light Detection and Ranging (LiDAR) is one of the remote sensing methods using the pulsed laser to measure the structure of forest (Lim et al., 2003), for instance, using canopy cover to capture forest density, the percentile of height to capture vertical structure, and using canopy standard deviation to capture horizontal heterogeneity of vegetation. These structure information have been used to detect biodiversity for many organisms like fungi (Thers et al., 2017), birds (Clawges et al., 2008), and woody plants (Turner et al., 2003), providing a supplement approach for the costly and arduous field inventories, especially in the broad scales (Zlinszky et al., 2015). Nevertheless, functional diversity has never been thoroughly investigated by LiDAR-derived variables. Thus, there is a need to investigate whether LiDAR can be a proxy for functional diversity, which can not

only provide a relatively cost-effective research method but also enhance our knowledge on how woody plants disperse between suitable habitats. This would be helpful to design reliable concepts for biodiversity management at a landscape scale.

In this study, I developed several LiDAR-derived variables, together with field inventory that I did this summer, trying to predict species richness and functional diversity. My objective of the study was going to investigate the following questions:

1) Whether the LiDAR-derived variables are reliable and can capture field-derived characteristics. Due to the complete information of LiDAR product, I hypothesize that would be possible.
2) Whether LiDAR-derived variables can be used to predict woody species richness and functional diversity, since vegetation structural and topographical characteristics showing the suitability of habitats, I hypothesize that would be possible.
3) How management intensity has an impact on species richness. I hypothesize that woody plant species richness would increase when forest management intensity at a moderate level due to the intermediate disturbance hypothesis.
4) How management intensity drives functional richness. I hypothesize that functional richness decreases with the increase of management intensity due to broader ecological niches and stronger dispersal abilities.

2. Methods

2.1.Study area

Gribskov is the fourth largest forests in Denmark located in northern Zealand covering approximately 5600 ha, which has been heavily influenced by people's activities (e.g., hunting, grazing, and timbering; (Lelli et al., 2019; Overballe-Petersen et al., 2013). The

1961-1990 average temperature of Gribskov is 7.7 ^{o}C, and mean annual precipitation is 697 mm with four to six months without frost (Frich et al.; Laursen et al., 1999). The Gribskov naturally would be the deciduous forest (account for approximately 54%), including European beech (*Fagus sylvatica* L.) and oak (Quercus sp.). However, nowadays 36% of the forest has replaced by exotic Norway spruce (*Picea abies* (L.) H. Karst). As of 2008, only 7% of the land was used for agriculture road construction (Overballe-Petersen et al., 2014). This project was conducted in the southern Gribskov (Fig. 1a), containing four different management types:

1) Traditional Managed forest (TM) – Traditional intensively managed beech forest for 70-100 years with simple structure, a large number of stumps, and the small proportion of deadwood and veteran trees
2) Near Natural Managed forest (NNM) – Intensively managed beech forest to mimic nature with heterogeneous structure and a large proportion of deadwood, veteran trees for about 30 years.
3) Recently Unmanaged forest (RU) – Unmanaged forest for about 30 years with trees older than 100 years
4) Long Unmanaged forest (LU) – Protecting forest from management for more than 50 years with veteran trees older than 200 years

In this study, I sampled twenty nested plots in four stands with different management types (15 m radius plot with 5 m radius circle inside Fig. 1a, b) based on the field inventory in 2015 (Lelli et al., 2019), trying to capture the gradient of species richness and functional richness among different management types.

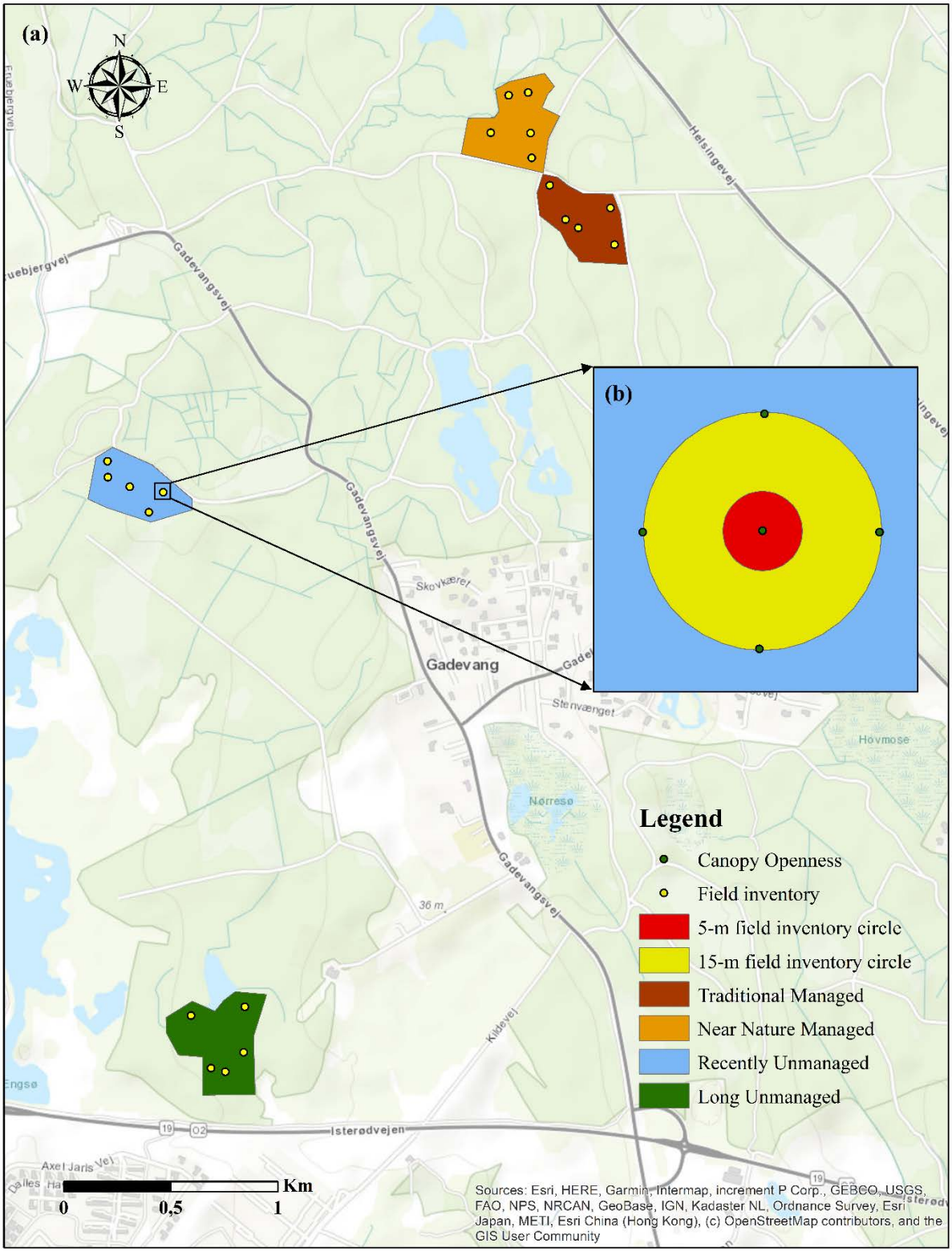

Fig. 1. Overview of field forest inventory in the different managed stands; (a) represents the different sampling plots in different stands; (b) represents the nested sample plots that I made in the field.

2.2.LiDAR-derived explanatory variables

Danish national airborne LiDAR scanning survey was taken in 2014-2015 (refers to DHM/Punktsky, https://kortforsyningen.dk/) with a projection of UTM32/ETRS89 (Nord-Larsen et al., 2017). I only used the point cloud of Gribskov, which was scanned in 2014 spring. The scan angle was 30^{o,} and flying height was 680 m above the ground with a speed of 240km/h. The collected point cloud is in full-waveform .las format which has an average density of 4.6 pts/m^2, containing various classifications based on American Society for Photogrammetry and Remote Sensing categories (ASPRS; e.g., ground, building, low vegetation, medium vegetation, high vegetation; Thers et al., 2017) except waterbodies. The accuracy of the point cloud is 0.15 m in horizontal and 0.05 m in vertical (Nord-Larsen et al., 2017). Final point cloud attributes include XY coordinate, Z altitude, intensity, return number, number of returns, ASPRS classification, and point source ID. Data is available online in 1 km^2 grid with formate of ".laz." The point cloud data in this study was processed into the following steps:

1) Point cloud data preparation
2) Metrics calculation (e.g., vegetation, intensity, and topography)

The point cloud data preparation included clipping the point cloud dataset from 20 15-meter radius plot in four stands of different management intensity (Fig. 1); normalizing point cloud by setting existing ground classification as 0 to remove topography effect; deleting point with Z values below 0 to remove the noise.

After the preparation of LiDAR data, I computed 14 metrics to retrieve vegetation structure, intensity, and topographic information entirely based on previous studies (Hudak et al., 2008; Thers et al., 2017). The vegetation, intensity, and topographic metrics were computed in 1 m^2 resolution and then calculated the mean value at the plot level (15 m radius circle; approximately 707 cells per plot). The vegetation metrics included vegetation height,

density, vertical, and horizontal distribution. The intensity metrics include skewness and ground proportion. Skewness metrics were usually used to separate objections (Yunfei et al., 2008) and here I chose from both single and all the returns, trying to test whether the intensity of the single return could be a better indicator of functional richness than all returns. The intensity of ground proportion was chosen to check whether it can replace with fallen deadwood and wetland information. Finally, the digital elevation model (DEM) and roughness were computed to represent topographic information (see Table 1 for the details of each matric).

Table 1. LiDAR-derived explanatory variables in this study.Point cloud data was retrieved from GeoDanmark, and all the metrics were calculated in 1 m^2 resolution at first and then computed average at the plot level. All the calculations were based on "lidR" package in R. Note that the bold metrics were selected to build the final model and "L" letter at the beginning represents LiDAR.

	Name of the explanatory variable	**Explanation**
Vegetation height	**L_Canopy_Ht**	The tree height derived from Canopy Height Model (CHM) using Individual Tree Detection (ITD) algorithm (Popescu and Wynne, 2004)
Vegetation density	**L_Canopy_Cov**	All returns over 2 m divided by all the returns
	L_Ground_Prop	Ground returns divided by all returns
Vertical distribution of vegetation	L_Ele_Skew	The skewness of elevation of all returns
	L_Perc_5	5%th percentile of elevation
	L_Perc_25	25%th percentile of elevation
	L_Perc_50	50%th percentile of elevation
Horizontal heterogeneity of vegetation	**L_CHM_SD**	The standard deviation of each cell and 8 adjacent cells from Canopy Height Model (CHM)
	L_STRATUM50_SD	Returns with altitude > 0 and ≤0.5 meter, divided by all vegetation returns (no ground returns), which may capture low scrub vegetation
Intensity values	L_IntSkew_All	The skewness of all return intensity values, which has a mixed effect of vegetation height, canopy heterogeneity, and reflectance properties of the ground and vegetation (Thers et al., 2017).
	L_IntSkew_single	The skewness of single return intensity values, with the same ecological interpretation of L_IntSkew_All
	L_IntGround_Prop	The proportion of ground intensity; which may capture the deadwood and presence of wetland information
Topographic metrics	**L_DEM**	Average elevation derived from Digtle Elevation Model (DEM)
	L_Rgd	The proportion of surface area to the projected area on the ground; measuring the heterogeneity of terrain, which may capture potential microhabitats

2.3.Field-derived explanatory variables

The field-derived variables were mainly chosen from Lelli's studies (Lelli et al., 2019) showing a good proxy for functional diversity. The field-derived variables were divided into vegetation metrics, coarse woody debris (CWD), and detailed information in the field. To get stable models, I sampled five nested replicates for each stand from July 18th to July 27th, 2019 with relative sunny weather, based on the previous field inventory in 2015 (Fig. 1a, 1b; Lelli

et al., 2019). To capture more variation in the forest and topography structure, I used LiDAR-derived metrics as a guide to select suitable plots prior to the field visits. I exported LiDAR-derived metrics in 27 m x 27 m resolution, which has the same area as the 15 m radius plots. Then I chose the plots with the highest and lowest values in each LiDAR-derived variable, counted the plot number, and selected the top five plots with extreme values in each stand. In the field, I used the GARMIN GPS map 64s to set plots location. The sample plots were 5 m radius circles nested with 15 m radius circles (Fig. 1b). In the 5-meter circle, diameter at breast height (DBH) was measured with a caliper, and height was measured with clinometer and laser rangefinder for all alive trees greater than 5 cm in DBH. Besides, height for alive bushes and treelets which heights over 10 cm were included. In the 15-meter circle, DBH and height were only recorded for alive trees with DBH over 15 cm. The identification of living woody plants was registered at the same time. Length, diameter, and decay level of both standing and fallen coarse woody debris (diameter > 10 cm) were measured. Accurately, the decay level was estimated in the following scale:

1. Recently dead, the wood still hard, with bark > 50%, and almost without moss;
2. The wood still hard but starting to become soft and with moss growing at the surface (often < 50% bark)
3. The original wood structure disappears, and wood becomes soft and covered by the moss

Finally, further measurement of canopy closure, number of stumps, and the presence of wetlands were registered as well. To calculate canopy closure, I firstly took photos vertically of 5 points in the 15-m circle (Fig. 1b). Back in the lab, I then calculated the mean of sky cluster in the RGB images, setting the 0.2 radius color range in blue sky cluster and 0.1 radius color range in cloud cluster (RGB = 1,1,1), computing the proportion of the whole sky (blue sky + cloud) in each photo (Fig. 2a – 2e). Subsequently, I checked manually for all the images recognition result and computed the average of all the images in each plot to see the

canopy openness. Eventually, I used 1 minus openness to get the canopy closure. Standardized basal area was scaled according to the size of different concentric circles (i.e., divided by 78.5 m^2 (5 cm < DBH < 15 cm), 706.9 m^2 (DBH > 15 cm)). Then the plot level estimation was computed by summing the scaled DBH. Biomass of CWD was calculated with the assumption of considering CWD as a cylinder. The biomass of fallen and standing CWD was computed separately, and then the proportion of 3rd decay level was calculated as well (see Table 2 for the details and rest variables).

Table 2. Field-derived variables in this study. Note that the bold metrics were selected to build the final model and "F" letter at the beginning represents the field.

Variable category	Name of the explanatory variable	Explanation
Vegetation height	F_Tree_Ht	Measuring tree height in the field with Clinometer
Vegetation density	F_Canopy_Closure	Measuring canopy closure by taking photos (Fig.1b) and computing the proportion of crown
	F_BA	Measuring Diameter of Breast Height (DBH) in the field and calculating standardized basal area
Coarse Woody Debris (CWD)	**F_CWDF**	Biomass of fallen CWD
	F_CWDS	Biomass of standing CWD
	F_CWD3_Prop	The proportion of CWD in 3rd decay level
Detailed information	**F_Wet**	Presence of water bodies
	F_Stumps	Number of cutting stumps
	F_Vete	Number of Veteran trees (with DBH > 80 cm)

2.4.Woody species richness and functional diversity

To investigate the changes of woody plants' species richness and functional diversity in respond to the management gradient, I choose seed mass (SM) and specific leaf area (leaf area per leaf dry mass; SLA) from BIEN data set (Maitner et al., 2018) as the functional traits of woody plants. These two traits are in quantity data types. Then I quantified functional diversity by computing RaoQ and CWM separately for each trait (Table 3). Since Rao's Q is affected by both the univariate trait distributions and the covariance between traits, I used single trait approach to avoid these effects.

Table 3. Response variables derived from fieldwork and BIEN Dataset. Note that function richness was calculated by the functional traits of specific leaf area (leaf area per dry mass; SLA) and seed mass (SM).

Response variables	**Explanation**
Species richness	No. of woody species in each plot
RaoQ_SLA	The dissimilarity of specific leaf area
CWM_SLA	The community weighted specific leaf area
RaoQ_SM	The dissimilarity of seed mass
CWM_SM	The community weighted seed mass

2.5.Data analyses

2.5.1. Pre-processing the model

To ensure the reliable of LiDAR-derived variables, I compared the field-derived tree height and canopy closure (Table 2) with LiDAR-derived canopy height and canopy cover (Table 1), respectively. Then I computed the Pearson correlation between all the variables (both LiDAR- and field- derived; Table 1, Table 2) with an independence threshold of 0.7 to avoid multi-collinearity of the model. Finally, I applied the principal component analysis (PCA) on the LiDAR and field derived variables of each stand in different management types before the model processing in order to examine the relationships of these response variables.

2.5.2. Modeling species richness and functional diversity

I used LiDAR-derived, and field-derived variables combined and separately (bold values in Table 1 and Table 2) to explain the species richness and functional diversity (Table 3) and evaluate whether LiDAR variables were sufficient to explain the response variables. Multivariate linear regression models (MLR) with normal distribution was used for modeling. To meet the model assumption, I log-transformed the response variables of RaoQ_SLA, RaoQ_SM, and CWM_SM based on natural logarithms before modeling. After the initial modeling, I selected explanatory variables with stepwise variable selections in both directions for the five response variables (Table 3). The decision of best models were supported by the

lowest Akaike information criteria (AIC; Wagenmakers and Farrell, 2004). Then I standardized regression coefficient to compare variable importance. Finally, I evaluated the homoscedasticity and QQ-normality plot visually to check the model assumptions were met.

All the analyses in this study were performed by R 3.6.0 (Team, 2018) with the packages of “lidR” (Roussel and Auty, 2018), “colordistance”(Weller and Westneat, 2019), “BIEN” (Maitner et al., 2018), “FD” (Laliberté and Legendre, 2010), “ecospat” (Di Cola et al., 2017), “stats” (Team, 2018), and “lm.beta” package (Behrendt., 2014). The “lidR” package was used to compute LiDAR-derived variables with the function “grid_canopy()” for grid canopy, “tree_detection()” for individual tree detections, and “grid_metrics()” for several user-defined functions (see Table 1 for the definition of different metrics). The “countcolors” was mainly used for image recognition to calculate the canopy closure based on “getKMeanColors()” function. Woody plant traits were obtained from “BIEN” package by using “BIEN_trait_traitbyspecies()” function. RaoQ and CWM were computed by the “dbFD()” function in “FD” package. The “ecospat” package was used to examine the Pearson correlation with the function “ecospat.cor.plot()”. Multiple linear regression was executed by “lm()” function in “stats” package and standardized regression coefficient was computed by “lm.beta()” function in “lm.beta” package.

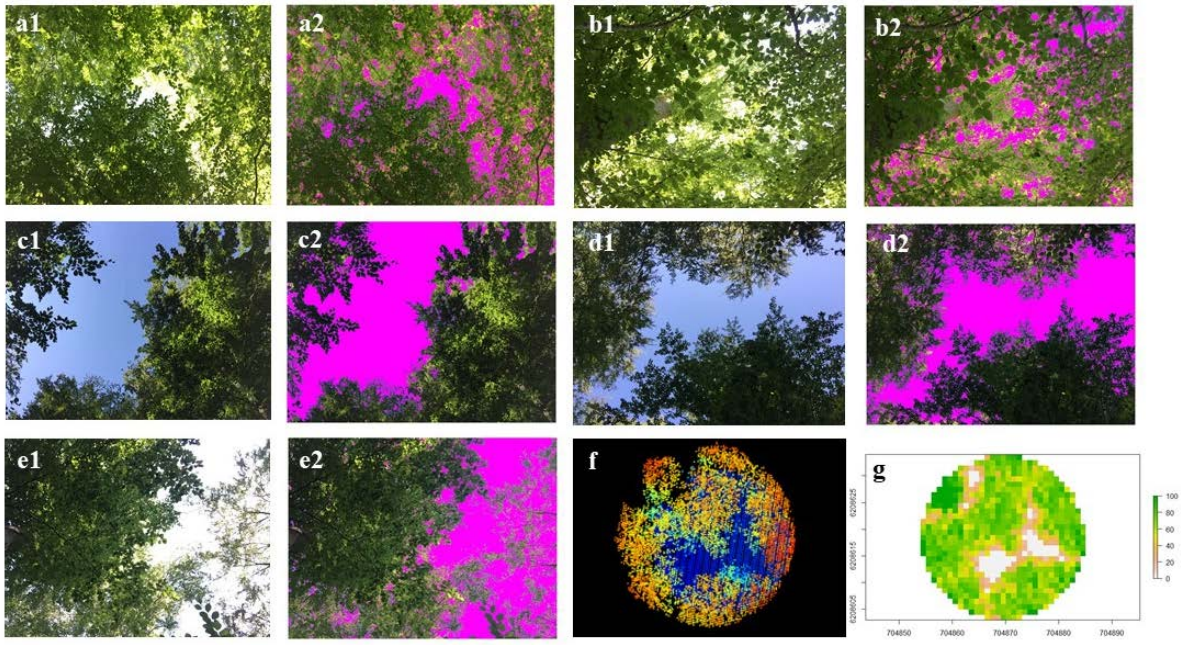

Fig. 2. Comparison between canopy openness based on RGB image recognition and LiDAR-derived canopy cover model. a-e represent image recognition in the near nature management plot 8_7, photos were taken at the middle point and 4 points around 15 m circle (Fig. 1b); 1 represents photo taken in the field; 2 represents recognition of sky clusters; f represents point cloud data of plot 8_7; g represents canopy cover raster in the plot 8_7.

3. Results

3.1. Field validation of LiDAR-derived metrics of forest structure

The canopy closure was assessed in the 5 points in each plot (Fig. 1b) by recognizing sky and cloud cluster, revealing relatively precise estimations by visual comparisons (Fig. 2a-e). LiDAR-derived canopy cover measured the whole canopy layer within the plot in 1 m resolution (Fig. 2f; 2g), which could represent the canopy closeness comprehensively. The canopy closure and tree height measured in the field revealed strong positive relationships with LiDAR-derived canopy cover and canopy height (r = 0.69, r = 0.72, respectively; Fig. 3), indicating the high quality of LiDAR-derived variables. Since the relationship in canopy closeness is under 1:1 line (Fig. 2), showing the LiDAR-derived canopy cover may underpredict canopy closeness. Besides, the skewness of single intensity return could not be able to predict the skewness intensity of all returns due to low Pearson correlation (r = 0.21; Fig. S1). Moreover, although the intensity of ground proportion (L_IntGround_Prop)

correlated negatively with ground features such as fallen coarse woody debris (F_CWDF), number of stumps (F_Stumps), and the presence of wetlands (F_Wet), it could not be able to predict due to low Pearson correlations (r = 0.45, 0.21, 0.3, respectively; Fig. S1).

3.2. PCA analysis

Principal component analysis (PCA) was used to determine the clustering of all the explanatory variables (both LiDAR and field) among the four management types (Fig. 4). Management types differed most along the first axis (PC1), which accounts for 23.2% of the variance in the range of explanatory variables. The long unmanaged stand was most distinct along PC1 showing the presence of wetland (F_Wet) and a large amount of fallen deadwood biomass (F_CWDF) especially in 3rd decay level (L_CWD3_Prop) with low axis scores. While a large number of stumps (F_Stumps), high elevation (L_DEM), and ruggedness (L_Rgd) of the terrain characterized the managed stands (traditional and near nature managed) with high axis scores in the opposite direction. The PC2 explained 19.1variations, showing the gradient of canopy height (L_Canopy_Ht), canopy cover (L_Canopy_Cov), and 5th percentile of height with positive scores in axis 2, while intensity skewness (L_IntSkew_Single) and stand coarse woody debris (F_CWDS) associated with negative axis 2 scores. But these variables could not distinguish the different management types (p > 0.1; ANOVA test) since recently unmanaged stand exhibited a much wider variety in PC 2 than other management types (Fig. 4) with high stand deadwood (F_CWDS).

3.3. Linear Model result

Among the stepwise models showed below, the combined model performed best with lowest AIC except for RaoQ_SLA. Field-derived models simulated better than LiDAR-derived models in predicting species richness and RaoQ_SLA. LiDAR-derived models were generally better than field-derived models at functional diversity predictions (Table 4).

Species richness revealed gradient which positively correlated with coarse woody debris (i.e., F_CWDS, F_CWDS), and the number of stumps (F_Stumps) in the combined models (Table 4) with the highest explanation (r^2 = 0.61). The LiDAR-derived model only remains intercept, which was not able to predict species richness in this study, indicating species richness was mainly predicted by the field derived variables.

Field-derived model performed best among the three different models in predicting RaoQ_SLA (Table 4), which correlated negatively with the basal area (F_BA) and positively correlated with fallen deadwood (F_CWDF) and presence of wetlands (F_Wet). However, the combined models further improved r^2 (from 0.42 to 0.71) by adding information positively on the LiDAR-derived canopy cover (L_Canopy_Cov), canopy height standard deviation (L_CHM_SD), and DEM; negatively on the canopy height (L_Canopy_Ht), 5^{th} percentile of height (L_Perc_5), single intensity skewness (L_IntSkew_Single), and ruggedness (L_Rgd). Although RaoQ_SLA correlated with the PC1 elements (F_BA, F_CWDF, and F_Wet, etc.), it shows no difference in the management gradient. CWM_SLA has a robust negative correlation with the number of stumps (F_Stumps) and canopy cover (L_Canopy_Cov) in the combined and LiDAR-derived model, respectively. This indicates specific leaf area was mainly determined by the forest management intensity and light conditions, while CWM_SLA also revealed no difference along the management gradient (Fig. 5).

The RaoQ_SM revealed the gradient in management types correlated positively with biomass of fallen deadwood (F_CWDF) and presence of wetland (F_Wet) in the combined and field models (Table 4), indicating long unmanaged stand with a large amount of fallen deadwood and presence of wetland tends to have dissimilarity in seed mass (Fig. 5). The Rao_SM in the LiDAR-derived model correlated negatively with L_DEM and positively with L_Perc_5, revealing topography and vertical vegetation structure would also capture seed

mass divergence. When predicting CWM_SM, F_BA, F_CWD3_Prop, and L_DEM detect the CWM together in the combined model (Table 4). While in the LiDAR-derived model, vegetation height (L_Canopy_Ht), vegetation structure (L_CHM_SD), and topography factors (L_DEM and L_Rgd) also contributed to the CWM_SM. In the field-derived model, basal area detected most on the CWM_SM. These explanatory variables captured the management gradient (Fig. 4), indicating communities tend to have larger seed mass in the unmanaged forest compare to the traditional managed forest (Fig. 5).

4. Discussion

I evaluated the capacity of LiDAR as a proxy of woody plant species richness and functional diversity, comparing LiDAR and field-derived predictors across forest management gradient. In line with previous studies which used LiDAR as a complementary source of habitat information (Zellweger et al., 2014), my LiDAR and field-derived variables supplemented each other well in the combined model. While in the separate model, LiDAR-derived model performed slightly better in 3 cases (i.e., CWM_SM, RaoQ_SM, and CWM_SM), and field-derived model showed better performance in the rest two cases (species richness and RaoQ_SLA). This finding substantiates my hypothesis that LiDAR survey is a valuable alternative method to predict functional diversity. However, species richness cannot be predicted with only LiDAR-derived variables unlike other organisms, including fungi (Peura et al., 2016; Thers et al., 2017), butterflies (Aguirre-Gutiérrez et al., 2017), and animals (Simonson et al., 2014), which may be caused by the model selections (e.g., generalized linear model, random forest model) and study scales (Lopatin et al., 2016; Madonsela et al., 2018). Moreover, contrary to my hypothesis based on IDH, the species richness has no difference along with the management types (p = 0.47; ANOVA test).

Although some studies found vascular plants increase richness with increase of management intensity (Paillet et al., 2010), there are still some studies showing no difference among the forest management gradient (Both et al., 2019), which may due to small study scales could not be able to represent the whole gradient. In the functional diversity analysis, seed mass showed dissimilarity along with the forest management gradient, although there are still many overlaps with each other while the specific leaf area showed no difference among the forest management gradient.

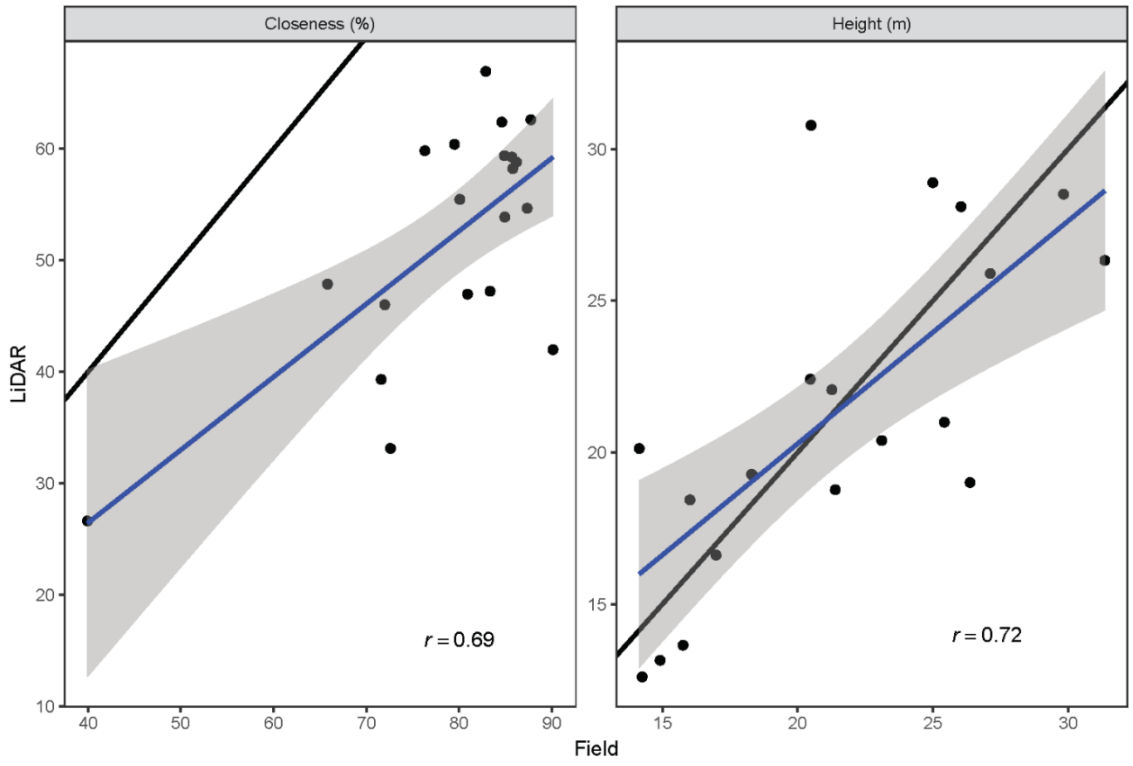

Fig. 3. Comparison between field-derived variables and LiDAR-derived variables; canopy closure (%) was calculated by the image recognition (Fig. 2) in the field to represent closeness while LiDAR-derived canopy cover was computed by the proportion of returns over 2 m (Table 1); height (m) in the field was collected by clinometer while in LiDAR was collected from individual tree detection. The blue lines in the figure are the regression lines and black lines in the figure are the 1:1 line. r indicates the Pearson correlation between field and LiDAR-derived variables.

4.1.Field validation

In the closeness comparison, the LiDAR-derived canopy closeness was defined by canopy cover indicating proportion forest area occupied by the vertical projection of canopy, which often used to estimate overstorey cover (Korhonen et al., 2006). While field-derived

canopy closeness was defined by canopy closure, showing the percentage of the sky obscured by the vegetation viewed from a single point (Jennings et al., 1999). Although these two concepts often used as analogies, there are still some slight difference between them (Paletto and Tosi, 2009) which may cause some uncertainties of the field validation. Besides, another uncertainty for the field validation is the time difference of collecting data. LiDAR data was collected in the spring of 2014 while the field data were collected from 2019 summer. That why the LiDAR-derived canopy cover seems to underpredict the canopy closeness when comparing to field-derived canopy closure. Moreover, canopy closure was chosen from 5 points each plot (Fig. 1b), four of them were set on the boundary, which may not reflect the complete closure. The Pearson correlation of LiDAR-derived canopy cover and field-derived canopy closure is still 0.69, showing the LiDAR-derived variables are in high quality. The LiDAR-derived ground proportion of intensity (L_IntGround_Prop) showed no significant correlation with field-derived fallen coarse woody debris, the number of stumps, and presence of wetlands, it showed these features by visual comparison (Fig. S2). In this study, I only used the mean ground proportion of intensity, which may not be a good indicator of ground features. Range and standard deviation of the LiDAR-derived variables may show a better predictor than only mean of the metrics (Thers et al., 2017). The skewness of single return intensity (L_IntSkew_Single) have no correlation of all return (L_IntSkew_All, fig. S1), since the most of single return is from hard surface (e.g., building, ground, and top dense clusters of vegetation) which can be used to predict forest structure (e.g., height; Lim et al., 2003).

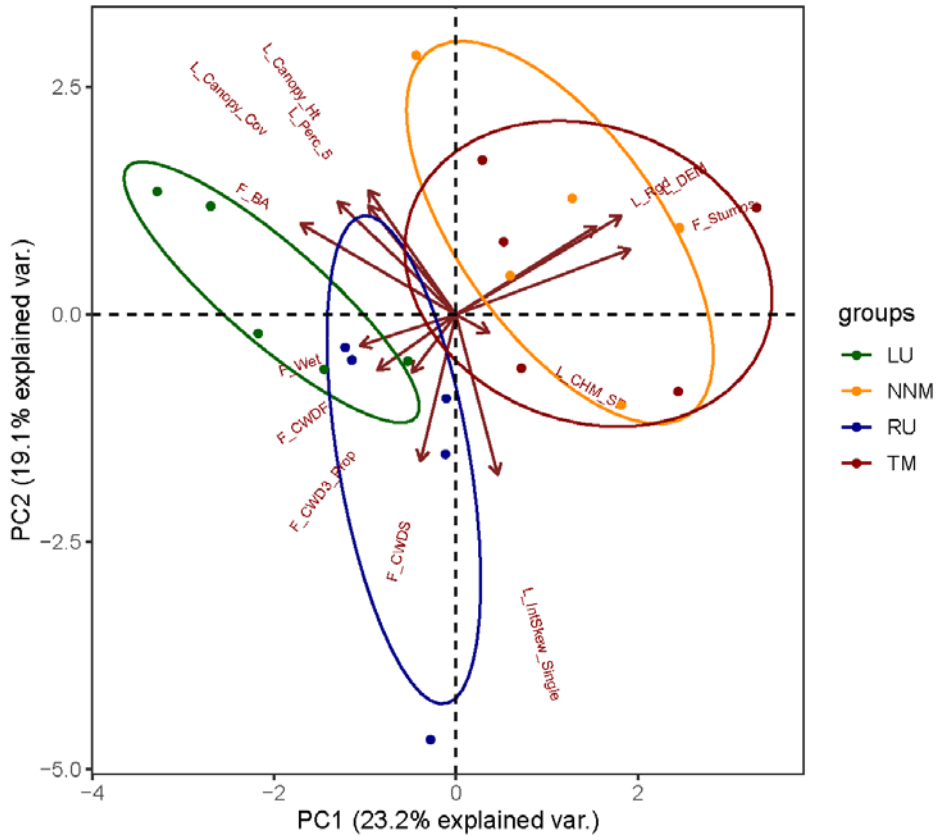

Fig. 4. PCA ordination shows the clusters of the explanatory variables (bold in Table 2 and Table 3) with ellipse group data of different management types. LU represents long unmanaged stand; NNM means near nature managed stand; RU indicates recently unmanaged stand, and TM represents the traditionally managed stand.

4.2. Species richness

Fallen coarse woody debris and the number of stumps have a significant positive correlation with species richness. Coarse woody debris provides nutrients and canopy gaps in the forest, allowing more bushes and treelets to grow, accelerating forest succession which may increases species richness. The number of stumps represents the forest management intensity, together with the canopy gaps in the forest, showing the disturbance in the forest could let light get through the canopy layer. Therefore, the understory could be able to grow, However, fallen coarse woody debris and number of stumps have opposite directions in the PCA analysis (Fig. 4), which may cause the species richness has weakly sensitive change across the forest management gradient.

4.3.Functional diversity predictions

In the functional diversity predictions, the LiDAR-derived models generally predict better than the field-derived models, while the combined models performed generally best among these models. However, Field-derived model performed best among three models when predicting RaoQ_SLA (Table 4). Fallen coarse woody debris and presence of wetlands correlated positively with the RaoQ_SLA, indicating that specific leaf area is sensitive to edaphic factors which is consistent with other studies (Pakeman, 2013), the high nutrient may increase the divergence of SLA. However, all the predictors did not show the significance in the prediction, indicating these factors were not able to capture RaoQ_SLA, which is consistent with other studies (Both et al., 2019), RaoQ_SLA showed no difference across the management gradient (Fig. 5), which may be due to study scale and SLA is more sensitive to climate gradient (Dubuis et al., 2013). Besides, there are some outliers in the leverage plot (beyond Cook's distance, e.g., 2_7, 4_3; Cook et al., 1980) may push the model into this direction. The 2_7 plot has a lot of deadwood since it has overlapped with other experiment plots (Table S2, Fig. S6, Fig. S8). This would put more weight in the fallen deadwood factor. CWM_SLA was mainly related negatively to L_Canopy_Cov since closed-canopy made woody plants with lower height hard to obtain light and only species with small specific leaf area, small photosynthetic capacity and high lifespan could survive in this situation (Wright et al., 2004). 4_3 is another plot in outliers (Fig. S10, S11), with a low canopy cover (26.63%) and some bushes with relative high SLA like *Crataegus laevigata* and *Rubus idaeus* may lead cause CWM_SLA models into this direction. Generally, the Q-Q normality plots and leverage plots show the relative reliability of the models (Fig. S3 – S17).

The dissimilarity of seed mass has a negative correlation with L_DEM but positive correlation with L_Perc_5, F_Wet, F_CWDF, which indicates a gradient of forest management intensity (Fig.4, 5). Seed mass showed a divergence pattern with a decrease of

management intensity (Fig. 5), but still have overlaps across the gradient, which may be caused by only four stands include in this study. In the long unmanaged forest, some species like *Quercus robur* and *Corylus avellana* with large seed mass showed up (Table S2) due to the high nutrient and more wetlands with low management intensity. Considering the simple species composition in Gribskov, the RaoQ_SM may be mainly affected by these relatively rare species. The CWM_SM related negatively with L_DEM and positively with F_BA, showing the gradient of forest management intensity as well. Long unmanaged forest with more veteran trees would increase biodiversity (Bergmeier et al., 2010), together with wetlands and high nutrient, providing suitable habitat for the species with larger seeds (e.g., *Corylus avellana*).

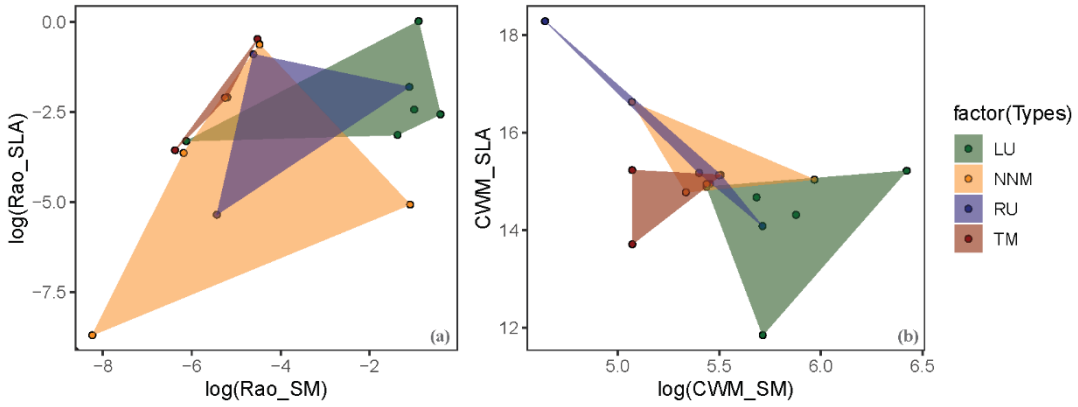

Fig. 5. Functional spaces occupied by stands in different management types. Scatterplots in (a) showing the degree of divergence (RaoQ) with log-transformed seed mass (SM) and specific leaf area (SLA) traits. It only includes 16 plots since some plots only have beech and the RaoQ is 0, which is infinite to transform to log. While in (b), the scatterplots position represents community weighted means (CWM) of seed mass (SM) and specific leaf area (SLA). The color of different shades includes all the plots within the stands. LU represents long unmanaged stand; NNM means near nature managed stand; RU indicates recently unmanaged stand, and TM represents the traditionally managed stand.

Table 4. selected models for combined and separate LiDAR/field derived variables. The LiDAR-derived variables were selected from the bold variable in Table 1. Field derived variables were collected from bold variables in Table 2. The five response variables are species richness, RaoQ_SLA, CWM_SLA, RaoQ_SM, and CWM_SM. The models were stepwise by AIC values. The standard linear regression coefficients were set in the brackets. Note that RaoQ_SLA, RaoQ_SM, CWM_SM have been log-transformed.

Response variables	**LiDAR + Field**		**LiDAR**		**Field**	
	AIC	Variables	AIC	Variables	AIC	Variables
Species richness	60.79	F_CWDF (0.69**) F_CWDS (0.58) F_CWD3_Prop (0.37) F_Wet (0.37) F_Stumps (0.98*) L_IntSkew_Single (-0.48) L_Rgd (-0.40)	65.38		60.91	F_CWDF (0.55*) F_CWDS (0.33) F_Wet (0.36) F_Stumps (0.45)
RaoQ_SLA	73.81	F_CWDF (1.06) F_Wet (0.34) F_Stumps (1.24) L_Canopy_Ht (-0.65) L_Canopy_Cov (0.56) L_Perc_5 (-0.52) L_CHM_SD (1.80) L_IntSkew_Single (-1.25) L_DEM (0.50) L_Rgd (-1.62)	73.51	L_Canopy_Ht (-0.43) L_CHM_SD (0.41)	70.78	F_BA (-0.44) F_CWDF (0.37) F_Wet (0.48)
CWM_SLA	44.01	F_CWDF (-0.15) F_CWD3_Prop (-0.34) F_Wet (-0.18) F_Stumps (-0.89***) L_Canopy_Ht (-0.50*) L_Canopy_Cov (-0.66**) L_Perc_5 (0.41*) L_CHM_SD (0.23) L_DEM (0.31)	51.93	L_Canopy_Cov (-0.86***) L_IntSkew_Single (-0.30) L_Rgd (-0.33)	65.41	F_BA (-0.42) F_Stumps (-0.44)
RaoQ_SM	59.19	F_CWDF (0.70**) F_CWD3_Prop (0.33) F_Wet (0.65**) F_Stumps (0.83*) L_Perc_5 (0.41*) L_CHM_SD (0.85) L_ IntSkew_Single (-0.47) L_DEM (-0.36) L_Rgd (-0.40)	72.71	L_Canopy_Ht (-0.36) L_Canopy_Cov (0.30) L_Perc_5 (0.56*) L_DEM (-0.62*) L_Rgd (0.31)	73.13	F_CWDF (0.51*) F_Wet (0.48*)
CWM_SM	-1.93	F_BA (0.40*) F_CWDS (-0.32) F_CWD3_Prop (-0.29*) L_Canopy_Cov (0.19) L_Perc_5 (0.29) L_DEM (-0.35*) L_Rgd (0.20)	3.46	L_Canopy_Ht (0.61**) L_Perc_5 (0.31) L_CHM_SD (-0.70**) L_ DEM (-0.53**) L_Rgd (0.57**)	7.10	F_BA (0.65***) F_CWDS (-0.30) F_CWD3_Prop (-0.28)

Note: p-value significance: <0.05 *; <0.01 **; <0.001 ***

4.4.Limitations and future work

4.4.1. Explanatory variables

Since there is no noise removal function “lidR” package, I could only get rid of the points which below ground. Thus, some noise over the ground cannot be removed. In this study, I used a large number of LiDAR metrics to simulate the species richness and functional diversity. These raw data have less straightforward interpretation (Thers et al.,

2017). Some people choose to convent raw LiDAR metrics into some ecological parameters, which can be interpreted directly (e.g., topography wetness index, TWI ; Peura et al., 2016). The ruggedness was computed based on the proportion of surface area to the projected area on the ground, which is sensitive to the spatial resolution. The L_CHM_SD is based on the surface model, which includes the forest canopy and bare ground surface. Although the image recognition of field-derived canopy closure was not affected by the weather (Fig. 3e) due to considering sky and cloud cluster, it was strongly affected by time, since there are more variations of sky color in the night but it sets the radius in advance. This needs to check manually and maybe reset the suitable radius in some images. The biomass of coarse woody debris was estimated into the cylinder. However, many of them were decayed, especially in 3rd decay level, showing an overestimation of the deadwood biomass. Last but not least, due to the time limitation, I was only able to do 20 plots within four stands, which may not be enough to build a stable model, and four stands may not be able to capture the distinct forest management gradient.

4.4.2. Response variables

In this study, I set species richness as the first response variable. However, in the conservation perspective, species richness may not be a suitable indicator of conservation value when compared with conservation-relevant species (e.g., red list species; Lelli et al., 2019). Besides, I only use specific leaf area and seed mass to investigate functional diversity, and many traits were not included due to lacking data (e.g., wood density, CN ratio). I only use single based trait methods to quantify functional diversity, which may not be able to see the big picture of how functional diversity distributed in the whole community. There are some multidenominational frameworks to choose, such as functional richness, evenness, divergence, and dispersion frameworks (Mason et al., 2005). Functional richness represents how much niche has been filled by species in the community which has been measured by the

minimum convex hull volume with all the species included (Cornwell et al., 2006), but it cannot integrate relative abundance information and highly sensitive to the outliers. Therefore rare species with extreme trait values will inflate functional richness (Cornwell et al., 2006; Laliberté and Legendre, 2010). By taking the relative abundance into consideration, functional evenness and divergence measure the regularity and divergence of traits distribution in niche space, respectively. These three metrics in the framework are independent with each other and can be applied to both single and multidenominational traits (Mason et al., 2005; Villéger et al., 2008). However, it can be only used for quantitative traits in the communities, which has more than 2 species and does not consider the dispersion shift (Laliberté and Legendre, 2010). Functional dispersion, as a new multidenominational framework, indicating the mean distance of individual species in the multidimensional trait space to the center of all species weighted by relative abundance, which is highly correlated with RaoQ and moderately related to functional richness and evenness (Laliberté and Legendre, 2010). So, lower functional dispersion indicates environmental filtering, and higher functional dispersion supports the idea of limiting similarity mechanisms (Ricotta and Moretti, 2011). It has no limitations traits and species number. One potential advantage of functional dispersion over RaoQ is that in the unweighted case (i.e., with presence-absence data), since RaoQ takes relative abundance into account. Although RaoQ can be applied to multidenominational traits, when coupled with CWM as a functional diversity framework (Ricotta and Moretti, 2011), RaoQ usually used in the single-based framework to see the environmental filtering and limiting similarity in single trait scale (Dubuis et al., 2013; García-Palacios et al., 2017; Grevé et al., 2018) since CWM can only be used in the single trait level. Single trait level analysis could provide direct insight into how function diversity distributed in each niche axis, whereas multidimensional trait analysis could give a big picture of how functional diversity distributed in the whole community.

4.4.3. Modeling processing

The model includes 15 explanatory variables and 5 response variables, which may have too many variables for the linear model. Reducing the dimensions with PCA analysis or NMDS distance might be able to simplify the model. Last but not least, this linear model includes count data (F_Stumps), proportion data (F_CWD3_Prop), and binary data (F_Wet). The continuous proportion follows the Beta distribution (Douma and Weedon., 2019), count data follows the Poisson distribution; binary data follows the binomial distribution, which may not be suitable for the linear regression model. The generalized linear model (GLM) would be more suitable for this case.

5. Conclusion

I confirmed my hypotheses that the LiDAR-derived variables could retrieve information from vegetation structure and topography information, which reveals a better prediction than field-derived variables in the functional diversity predictions. This is useful, especially in the large scale which forest structures cannot sufficiently be quantified. While Field-derived variables can provide some details which also could be crucial to the predictions, e.g., coarse woody debris, the proportion of 3rd decay level. Therefore, the combined model performed best by combining the advantages of the two methods. Functional diversity of seed mass showed the gradient along with forest management intensity while there are still a lot of overlaps across management stands, which may due to lacking enough replicates. The species richness and functional diversity of specific leaf area were not sensitive with the management intensity. In my future thesis, I hope I could use some distance-based framework to measure multidimensional trait space across a large gradient in a broad scale, to investigate whether the functional diversity could be affected by the land-use intensity in a larger scale.

Acknowledgment

I would like to express my sincere gratitude to my supervisor, Dr. Naia Morueta-Holme, for her patient guidance, enthusiastic encouragement, and useful critique of my project. I am sincerely thankful to Dr. Jacob Heilmann-Clausen and Dr. Niels Strange who provided me with the previous field inventory information and field equipment. I am also profoundly grateful to Dr. Michael Krabbe Borregaard, Dr. Hans Henrik Bruun, Ph.D. students Yujing Yan, and Xiangyan Su, for their statistical and concept instructions.

Reference

Aguirre-Gutiérrez, J. et al., 2017. Butterflies show different functional and species diversity in relationship to vegetation structure and land use. Global ecology and biogeography, 26(10): 1126-1137.

Behrendt, S., 2014. lm. beta: Add Standardized Regression Coefficients to lm-Objects. R package version 1.5-1.

Bengtsson, J., Nilsson, S.G., Franc, A. and Menozzi, P., 2000. Biodiversity, disturbances, ecosystem function and management of European forests. Forest ecology and management, 132(1): 39-50.

Bergmeier, E., Petermann, J. and Schröder, E., 2010. Geobotanical survey of wood-pasture habitats in Europe: diversity, threats and conservation. Biodiversity and Conservation, 19(11): 2995-3014.

Both, S. et al., 2019. Logging and soil nutrients independently explain plant trait expression in tropical forests. New Phytologist, 221(4): 1853-1865.

Botta-Dukát, Z., 2005. Rao's quadratic entropy as a measure of functional diversity based on multiple traits. Journal of vegetation science, 16(5): 533-540.

Brunet, J., Fritz, Ö. and Richnau, G., 2010. Biodiversity in European beech forests-a review with recommendations for sustainable forest management. Ecological Bulletins: 77-94.

Burivalova, Z., Şekercioğlu, Ç.H. and Koh, L.P., 2014. Thresholds of logging intensity to maintain tropical forest biodiversity. Current Biology, 24(16): 1893-1898.

Cadotte, M.W., Carscadden, K. and Mirotchnick, N., 2011. Beyond species: functional diversity and the maintenance of ecological processes and services. Journal of applied ecology, 48(5): 1079-1087.

Cardinale, B.J., Nelson, K. and Palmer, M.A., 2000. Linking species diversity to the functioning of ecosystems: on the importance of environmental context. Oikos, 91(1): 175-183.

Clawges, R., Vierling, K., Vierling, L. and Rowell, E., 2008. The use of airborne lidar to assess avian species diversity, density, and occurrence in a pine/aspen forest. Remote Sensing of Environment, 112(5): 2064-2073.

Cook, R.D. and Weisberg, S., 1980. Characterizations of an empirical influence function for detecting influential cases in regression. *Technometrics*, *22*(4), pp.495-508.

Coomes, D.A. and Grubb, P.J., 2003. Colonization, tolerance, competition and seed-size variation within functional groups. Trends in Ecology & Evolution, 18(6): 283-291.

Cornwell, W.K., Schwilk, D.W. and Ackerly, D.D., 2006. A trait-based test for habitat filtering: convex hull volume. Ecology, 87(6): 1465-1471.

Di Cola, V. et al., 2017. ecospat: an R package to support spatial analyses and modeling of species niches and distributions. Ecography, 40(6): 774-787.

Douma, J.C. and Weedon, J.T., 2019. Analyzing continuous proportions in ecology and evolution: a practical introduction to beta and Dirichlet regression. Methods in Ecology and Evolution.

Dubuis, A. et al., 2013. Predicting current and future spatial community patterns of plant functional traits. Ecography, 36(11): 1158-1168.

Esselink, P., Zijlstra, W., Dijkema, K.S. and van Diggelen, R., 2000. The effects of decreased management on plant-species distribution patterns in a salt marsh nature reserve in the Wadden Sea. Biological Conservation, 93(1): 61-76.

Frich, P., Rosenørn, S., Madsen, H. and Jensen, J.J., COPENHAGEN 1997.

García-Palacios, P., Shaw, E.A., Wall, D.H. and Hättenschwiler, S., 2017. Contrasting mass-ratio vs. niche complementarity effects on litter C and N loss during decomposition along a regional climatic gradient. Journal of Ecology, 105(4): 968-978.

Grevé, M.E. et al., 2018. Effect of forest management on temperate ant communities. Ecosphere, 9(6): e02303.

Grime, J., 1998. Benefits of plant diversity to ecosystems: immediate, filter and founder effects. Journal of Ecology, 86(6): 902-910.

Hahn, K., Emborg, J., Larsen, J.B. and Madsen, P., 2005. Forest rehabilitation in Denmark using nature-based forestry. Restoration of boreal and temperate forests, eds: 299-317.

Hudak, A.T., Crookston, N.L., Evans, J.S., Hall, D.E. and Falkowski, M.J., 2008. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. Remote Sensing of Environment, 112(5): 2232-2245.

Jennings, S., Brown, N. and Sheil, D., 1999. Assessing forest canopies and understorey illumination: canopy closure, canopy cover and other measures. Forestry: An International Journal of Forest Research, 72(1): 59-74.

Jetz, W. et al., 2016. Monitoring plant functional diversity from space. Nature plants, 2(3): 16024.

Kissling, W.D., Field, R. and Böhning-Gaese, K., 2008. Spatial patterns of woody plant and bird diversity: functional relationships or environmental effects? Global Ecology and biogeography, 17(3): 327-339.

Korhonen, L., Korhonen, K.T., Rautiainen, M. and Stenberg, P., 2006. Estimation of forest canopy cover: a comparison of field measurement techniques.

Laliberté, E. and Legendre, P., 2010. A distance-based framework for measuring functional diversity from multiple traits. Ecology, 91(1): 299-305.

Laliberte, E. et al., 2010. Land-use intensification reduces functional redundancy and response diversity in plant communities. Ecology letters, 13(1): 76-86.

Laureto, L.M.O., Cianciaruso, M.V. and Samia, D.S.M., 2015. Functional diversity: an overview of its history and applicability. Natureza & Conservação, 13(2): 112-116.

Laursen, E.V., Thomsen, R.S. and Cappelen, J., 1999. Observed Air Temperature, Humidity, Pressure, Cloud Cover and Weather in Denmark: With Climatological Standard Normals 1961-90. Danish Meteorological Institute Copenhagen.

Lavorel, S. et al., 2008. Assessing functional diversity in the field–methodology matters! Functional Ecology, 22(1): 134-147.

Lelli, C. et al., 2019. Biodiversity response to forest structure and management: Comparing species richness, conservation relevant species and functional diversity as metrics in forest conservation. Forest ecology and management, 432: 707-717.

Lim, K., Treitz, P., Wulder, M., St-Onge, B. and Flood, M., 2003. LiDAR remote sensing of forest structure. Progress in physical geography, 27(1): 88-106.

Lopatin, J., Dolos, K., Hernández, H., Galleguillos, M. and Fassnacht, F., 2016. Comparing generalized linear models and random forest to model vascular plant species richness using LiDAR data in a natural forest in central Chile. Remote sensing of environment, 173: 200-210.

Madonsela, S., Cho, M.A., Ramoelo, A., Mutanga, O. and Naidoo, L., 2018. Estimating tree species diversity in the savannah using NDVI and woody canopy cover. International journal of applied earth observation and geoinformation, 66: 106-115.

Maitner, B.S. et al., 2018. The bien r package: A tool to access the Botanical Information and Ecology Network (BIEN) database. Methods in Ecology and Evolution, 9(2): 373-379.

Mason, N.W., Mouillot, D., Lee, W.G. and Wilson, J.B., 2005. Functional richness, functional evenness and functional divergence: the primary components of functional diversity. Oikos, 111(1): 112-118.

Naeem, S., Thompson, L.J., Lawler, S.P., Lawton, J.H. and Woodfin, R.M., 1994. Declining biodiversity can alter the performance of ecosystems. Nature, 368(6473): 734.

McKinney, M.L., 2008. Effects of urbanization on species richness: a review of plants and animals. *Urban ecosystems*, *11*(2), pp.161-176.

Nord-Larsen, T., Riis-Nielsen, T. and Ottosen, M., 2017. Forest resource map of Denmark. Mapping of Danish forest resources using ALS from 2014, 15.

Overballe-Petersen, M.V., Nielsen, A.B., Hannon, G.E., Halsall, K. and Bradshaw, R.H., 2013. Long-term forest dynamics at Gribskov, eastern Denmark with early-Holocene evidence for thermophilous broadleaved tree species. The Holocene, 23(2): 243-254.

Overballe-Petersen, M.V., Raulund-Rasmussen, K., Buttenschøn, R.M. and Bradshaw, R.H., 2014. The forest Gribskov, Denmark: lessons from the past qualify contemporary conservation, restoration and forest management. Biodiversity and conservation, 23(1): 23-37.

Paillet, Y. et al., 2010. Biodiversity differences between managed and unmanaged forests: meta-analysis of species richness in Europe. Conservation biology, 24(1): 101-112.

Pakeman, R.J., 2013. Intra-specific leaf trait variation: management and fertility matter more than the climate at continental scales. Folia Geobotanica, 48(3): 355-371.

Paletto, A. and Tosi, V., 2009. Forest canopy cover and canopy closure: comparison of assessment techniques. European Journal of Forest Research, 128(3): 265-272.

Peura, M. et al., 2016. Mapping a 'cryptic kingdom': Performance of lidar derived environmental variables in modelling the occurrence of forest fungi. Remote Sensing of Environment, 186: 428-438.

Popescu, S.C. and Wynne, R.H., 2004. Seeing the trees in the forest. Photogrammetric Engineering & Remote Sensing, 70(5): 589-604.

Putz, F.E., Blate, G.M., Redford, K.H., Fimbel, R. and Robinson, J., 2001. Tropical forest management and conservation of biodiversity: an overview. Conservation Biology, 15(1): 7-20.

Ricotta, C. and Moretti, M., 2011. CWM and Rao's quadratic diversity: a unified framework for functional ecology. Oecologia, 167(1): 181-188.

Rockström, J. et al., 2009. A safe operating space for humanity. nature, 461(7263): 472.

Roussel, J. and Auty, D., 2018. lidR: Airborne LiDAR Data Manipulation and Visualization for Forestry Applications. R package version 1.0. 0.

Simons, N.K., Weisser, W.W. and Gossner, M.M., 2016. Multi-taxa approach shows consistent shifts in arthropod functional traits along grassland land-use intensity gradient. Ecology.

Simonson, W.D., Allen, H.D. and Coomes, D.A., 2014. Applications of airborne lidar for the assessment of animal species diversity. Methods in Ecology and Evolution, 5(8): 719-729.

Skidmore, A.K. and Pettorelli, N., 2015. Agree on biodiversity metrics to track from space: Ecologists and space agencies must forge a global monitoring strategy. Nature, 523(7561): 403-406.

Team, R.C., 2018. R: A language and environment for statistical computing; 2015.

Thers, H. et al., 2017. Lidar-derived variables as a proxy for fungal species richness and composition in temperate Northern Europe. Remote Sensing of Environment, 200: 102-113.

Turner, W. et al., 2003. Remote sensing for biodiversity science and conservation. Trends in ecology & evolution, 18(6): 306-314.

Villéger, S., Mason, N.W. and Mouillot, D., 2008. New multidimensional functional diversity indices for a multifaceted framework in functional ecology. Ecology, 89(8): 2290-2301.

Wagenmakers, E.-J. and Farrell, S., 2004. AIC model selection using Akaike weights. Psychonomic bulletin & review, 11(1): 192-196.

Weller, H.I. and Westneat, M.W., 2019. Quantitative color profiling of digital images with earth mover's distance using the R package colordistance. PeerJ, 7: e6398.

Wright, I.J. et al., 2004. The worldwide leaf economics spectrum. Nature, 428(6985): 821.

Yunfei, B. et al., 2008. Classification of LIDAR point cloud and generation of DTM from LIDAR height and intensity data in forested area. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 37(7): 313-318.

Zellweger, F., Morsdorf, F., Purves, R.S., Braunisch, V. and Bollmann, K., 2014. Improved methods for measuring forest landscape structure: LiDAR complements field-based habitat assessment. Biodiversity and conservation, 23(2): 289-307.

Zlinszky, A., Heilmeier, H., Balzter, H., Czúcz, B. and Pfeifer, N., 2015. Remote sensing and GIS for habitat quality monitoring: New approaches and future research. Multidisciplinary Digital Publishing Institute.

Appendix:

Table S1. The number of woody plants showed in different plots. Plot number begin with 2 indicates long unmanaged forest; plot number begin with 4 indicates recently unmanaged forest; plot number begin with 7 represents traditional managed forest; plot number begin with 8 represents near nature managed forest.

Plot No.	Acer pseudoplatanus	Betula pendula	Betula pubescens	Corylus avellana	Crataegus laevigata	Euonymus europaeus	Fagus sylvatica	Frangula alnus	Fraxinus excelsior	Picea abies	Prunus cerasifera	Quercus robus	Quercus Rubra	Rubus idaeus	Ulmus glabra
2_10	0	0	0	1	0	0	15	0	0	0	0	2	0	0	0
2_2	0	0	0	0	0	0	9	0	3	6	0	0	1	0	0
2_4	0	0	0	0	0	0	12	0	5	0	0	1	0	0	0
2_7	7	0	0	0	0	0	16	0	3	0	0	1	0	0	0
2_9	0	0	0	0	0	0	10	0	1	0	0	0	0	0	0
4_10	0	0	0	0	0	0	25	0	0	0	0	0	0	0	0
4_2	0	0	0	0	0	7	12	0	0	0	0	1	0	0	1
4_3	0	0	0	0	1	0	3	2	0	0	0	0	0	2	0
4_6	0	0	0	0	0	0	21	0	0	0	0	0	0	0	0
4_8	0	4	0	0	0	0	44	0	0	0	0	0	0	0	1
7_2	19	0	0	0	0	0	18	0	0	0	0	0	0	2	0
7_4	0	4	2	0	0	0	18	0	0	4	0	0	0	0	0
7_7	0	0	0	0	0	0	27	0	2	0	0	0	0	0	0
7_8	0	0	0	0	0	0	34	0	0	0	0	0	0	0	0
7_9	0	0	0	0	0	0	31	0	0	0	0	0	0	0	0
8_1	2	0	0	0	0	8	47	0	0	0	0	0	0	1	0
8_4	6	0	0	0	0	0	19	0	0	0	1	0	0	10	0
8_6	1	0	0	0	0	0	69	0	0	0	0	0	0	0	0
8_7	1	0	0	0	0	0	17	0	0	0	0	1	0	0	0
8_9	2	0	0	0	0	0	57	0	4	0	0	0	0	0	0

Table. S2. LiDAR and field-derived explanatory variables for each plot used in the PCA ordination and linear model. Field-derived variables are basal area (F_BA); fallen coarse woody debris (F_CWDF); stand coarse woody debris (F_CWDS); proportion of deadwood in 3rd decay level (F_CWD3_Prop); presence of wetlands (F_Wet); number of stumps (F_Stumps). LiDAR-derived variables are canopy height (L_Canopy_Ht); canopy cover (L_Canopy_Cov); 5th percentile of height (L_Perc_5); standard deviation of canopy height (L_CHM_SD); intensity skewness from single return (L_IntSkew_Single); dagitl elevation model (L_DEM); and ruggudness (L_Rgd). Plot number begin with 2 indicates long unmanaged forest; plot number begin with 4 indicates recently unmanaged forest; plot number begin with 7 represents traditional managed forest; plot number begin with 8 represents near nature managed forest.

Plot No.	F_BA	F_CWDF	F_CWDS	F_CWD3_Prop	F_Wet	F_Stumps	L_Canopy_Ht	L_Canopy_Cov	L_Perc_5	L_CHM_SD	L_IntSkew_Single	L_DEM	L_Rgd
2_10	0.639117	6.045456	3.307385	6.555482	1	2	25.89931	62.6098	0.63837	1.852795	-0.19649	22.03905	1.003209
2_2	0.273834	8.22307	2.499168	18.10182	1	11	20.3982	66.94146	0.267834	2.062855	0.00034	21.75488	1.00875
2_4	0.616241	6.167451	0.228084	45.14664	1	0	28.51355	59.37985	0.242285	2.070341	-0.18413	23.96393	1.007871
2_7	0.522442	37.5868	0.00053	76.00737	0	3	19.25435	53.87906	0.215319	1.771163	-0.05712	28.0158	1.011342
2_9	0.47512	10.80414	2.356194	48.6231	0	0	26.33268	58.22867	0.319342	1.896224	-0.06903	23.2219	1.001751
4_10	0.318105	1.834406	0	100	0	3	16.61847	41.99256	0.192798	1.580781	-0.09473	23.96393	1.007871
4_2	0.256961	15.95903	0	62.54715	0	6	21.00006	62.39682	0.42766	2.092511	-0.0368	16.86572	1.004164
4_3	0.081831	5.247525	12.6743	14.20385	1	1	19.01009	26.62859	0.14685	2.799999	0.035814	15.82986	1.001666
4_6	0.196157	6.747267	0	100	0	4	12.6064	54.66961	0.19069	1.840172	-0.04329	21.24786	1.005542
4_8	0.567512	1.204979	0	92.50927	0	4	18.43929	55.4638	0.246146	1.887937	-0.03915	23.28898	1.004101
7_2	0.205145	1.377259	0.017106	69.25993	1	5	18.7755	58.81564	0.227596	2.02538	-0.17055	53.8045	1.012539
7_4	0.197436	3.838314	5.405861	69.61385	0	9	22.41489	60.40496	0.28894	1.962441	0.00089	52.76528	1.004576
7_7	0.351146	5.05496	0.497186	37.26367	0	16	19.28423	33.13486	0.155047	2.164324	0.072715	47.41691	1.009044
7_8	0.329504	2.595147	0	0	0	5	30.78298	47.87553	0.198325	2.516	-0.20886	46.50461	1.007366
7_9	0.170087	0.518356	0	0	0	12	20.13131	46.02233	0.160351	2.481456	-0.09805	49.27347	1.027421
8_1	0.257457	1.084081	0	0	0	15	13.64983	47.23275	0.086123	1.327011	-0.18072	49.60062	1.012342
8_4	0.164935	12.30091	0	18.99461	0	8	13.15298	39.31927	0.113971	1.878857	-0.06239	42.98431	1.009201
8_6	0.281838	0.656623	0	0	0	9	22.06946	59.26748	0.14387	1.479002	-0.03771	40.94287	1.004484
8_7	0.359317	2.330252	0	0	0	12	28.10239	59.83277	1.174856	2.011642	-0.11633	38.86506	1.008202
8_9	0.344574	0.190421	0.35225	96.16144	0	10	28.89593	46.97081	0.291637	3.125977	-0.1208	40.09151	1.020435

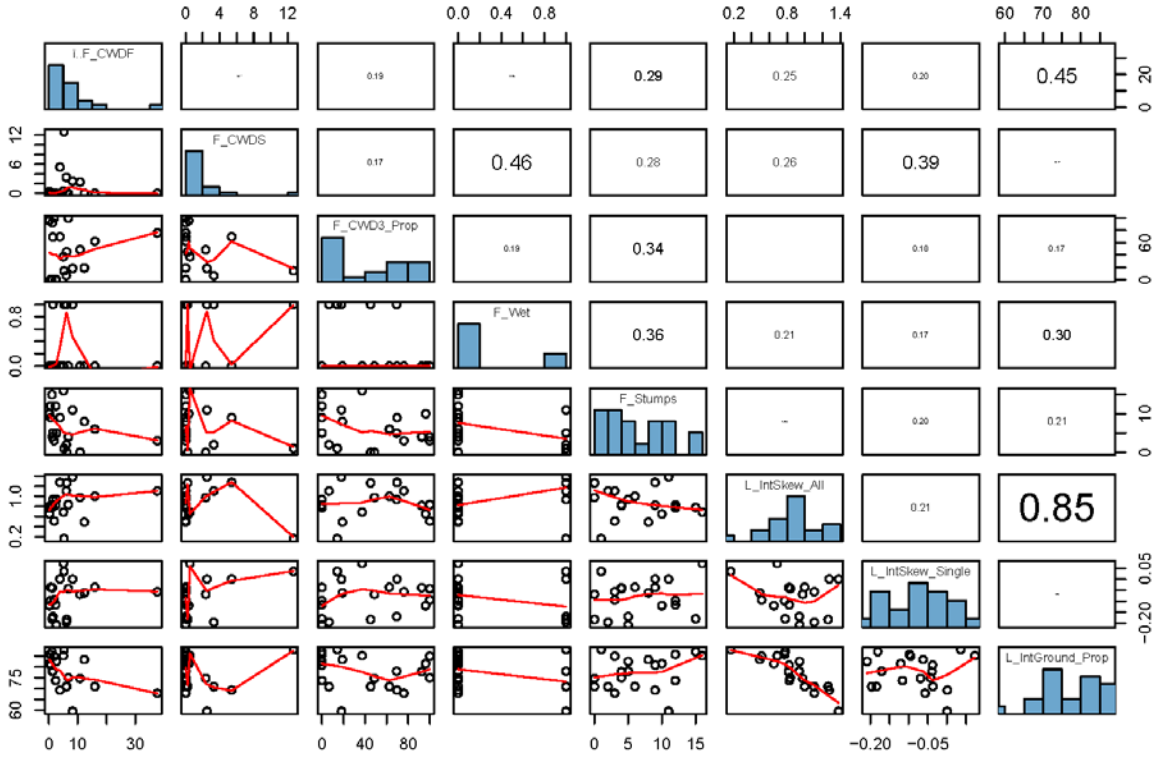

Fig. S1. The Pearson correlation between LiDAR-derived and field-derived variables. LiDAR-derived variables include the skewness of all return intensity values (L_IntSkew_All), the skewness of single return intensity values (L_IntSkew_Single), and intensity of ground proportion. Field-derived microhabitats consist fallen coarse woody debris (F_CWDF), stand coarse woody debris (F_CWDS), the proportion of coarse woody debris in 3rd decay level (F_CWD3_Prop), presence of wetlands (F_Wet), and number of stumps (F_Stumps).

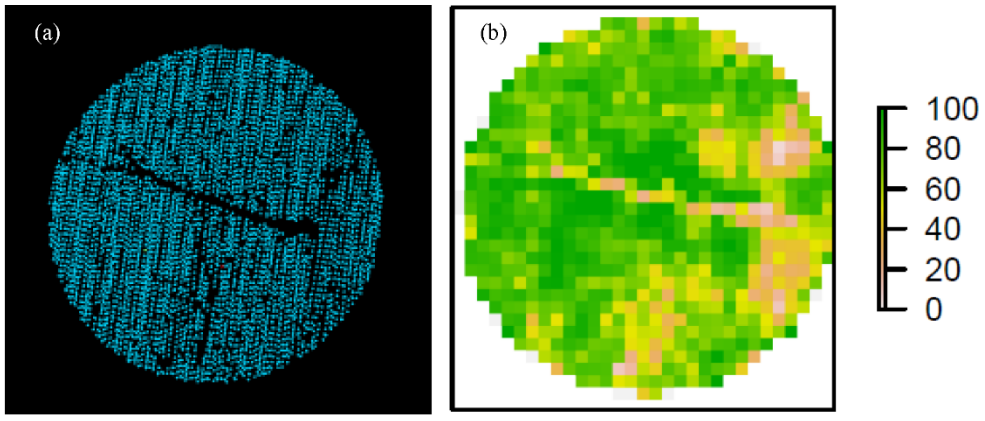

Fig. S2. The comparison between ground returns and intensity of ground proportion in example 2_10. (a) is the LiDAR-derived ground return, (b) is the intensity of ground proportion in 1m^2 resolution.

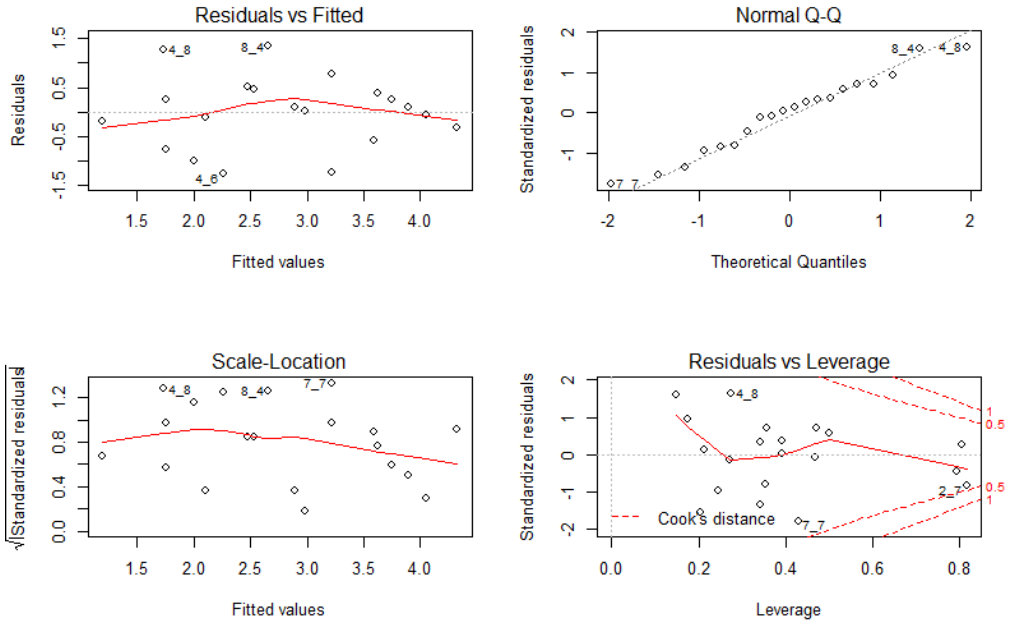

Fig. S3. LiDAR and field-derived species richness prediction

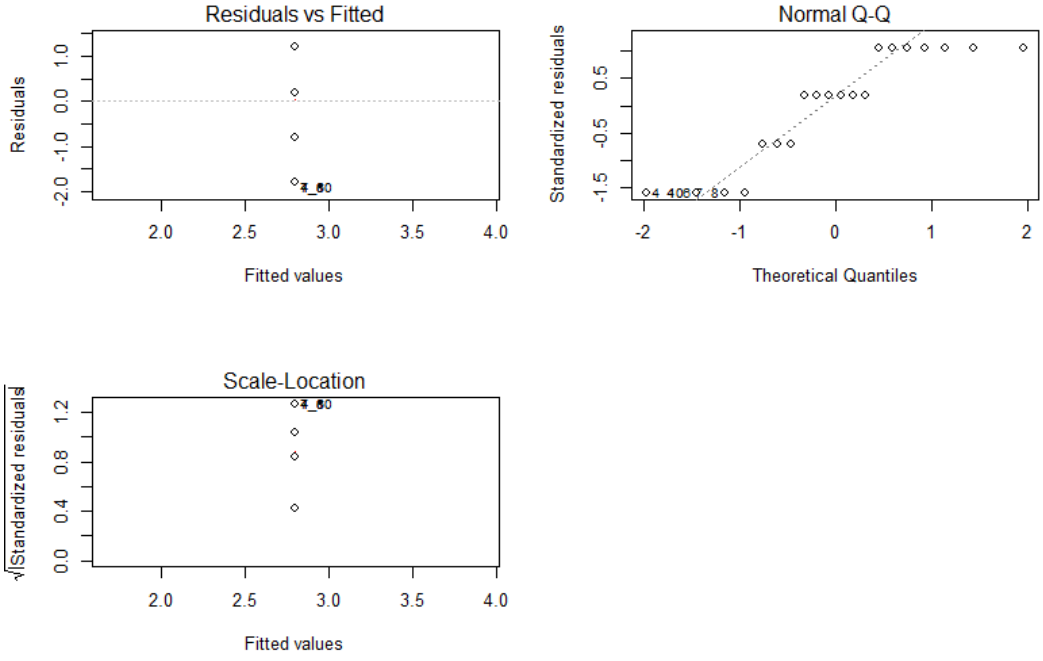

Fig. S4. LiDAR-derived species richness prediction

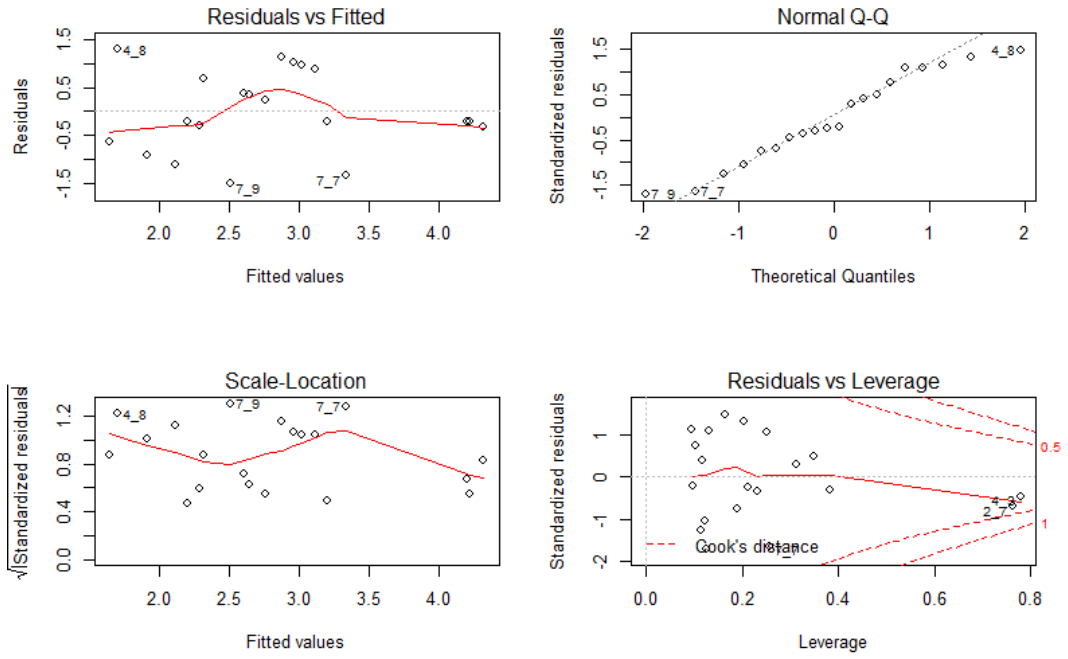

Fig. S5. Field-derived species richness prediction

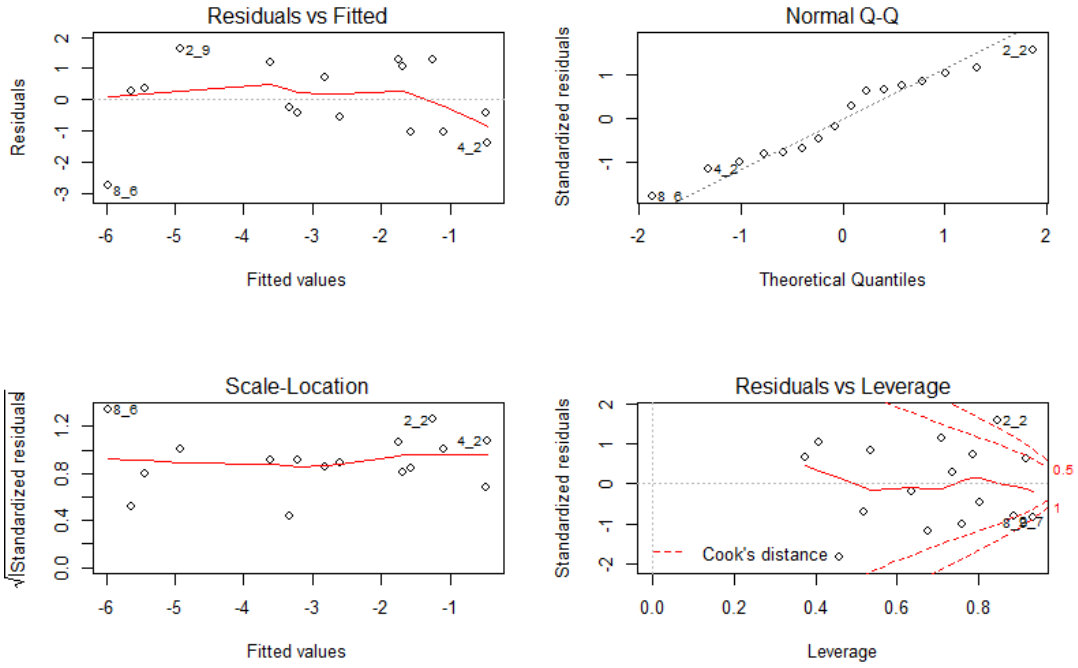

Fig. S6. LiDAR and field-derived RaoQ_SLA prediction

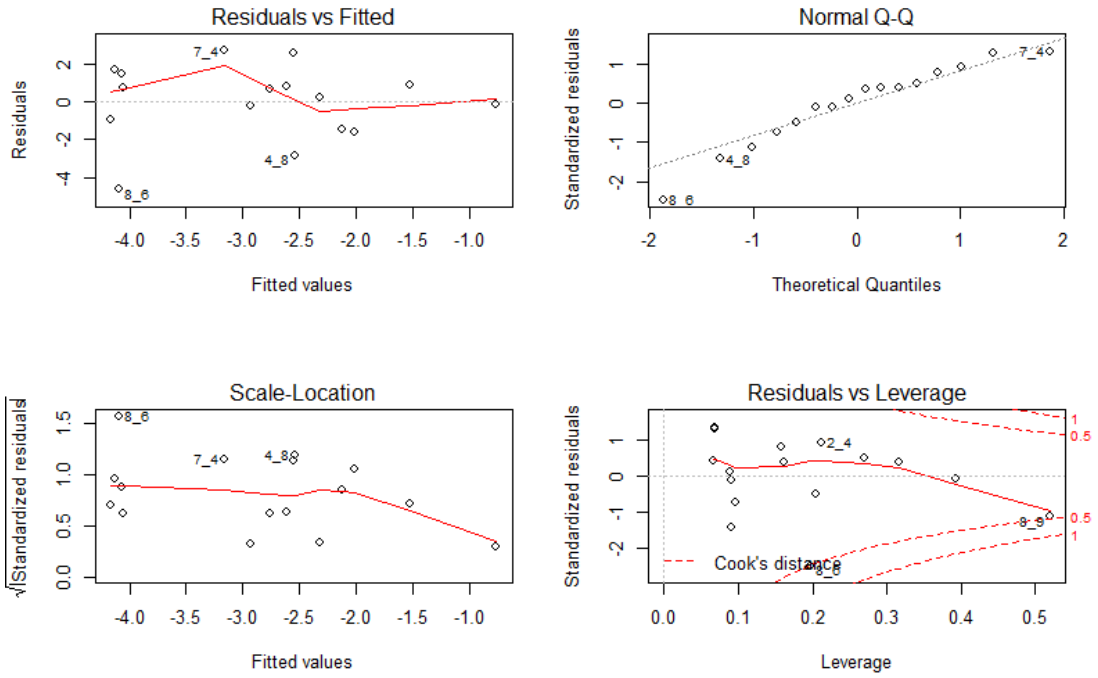

Fig. S7. LiDAR-derived RaoQ_SLA prediction

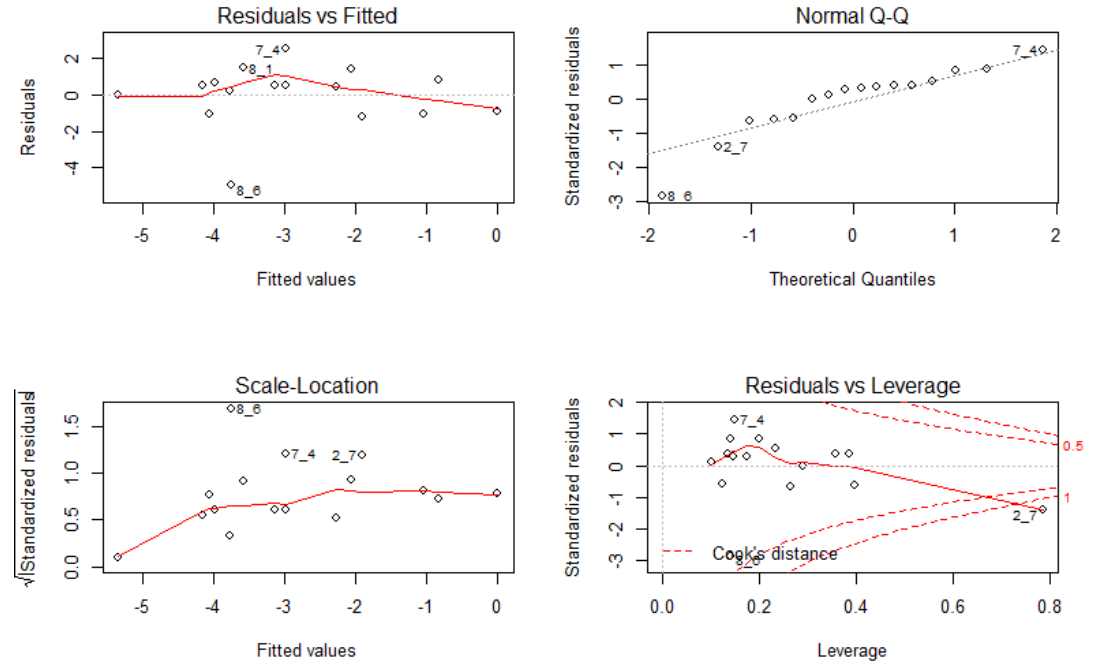

Fig. S8. Field-derived RaoQ_SLA prediction

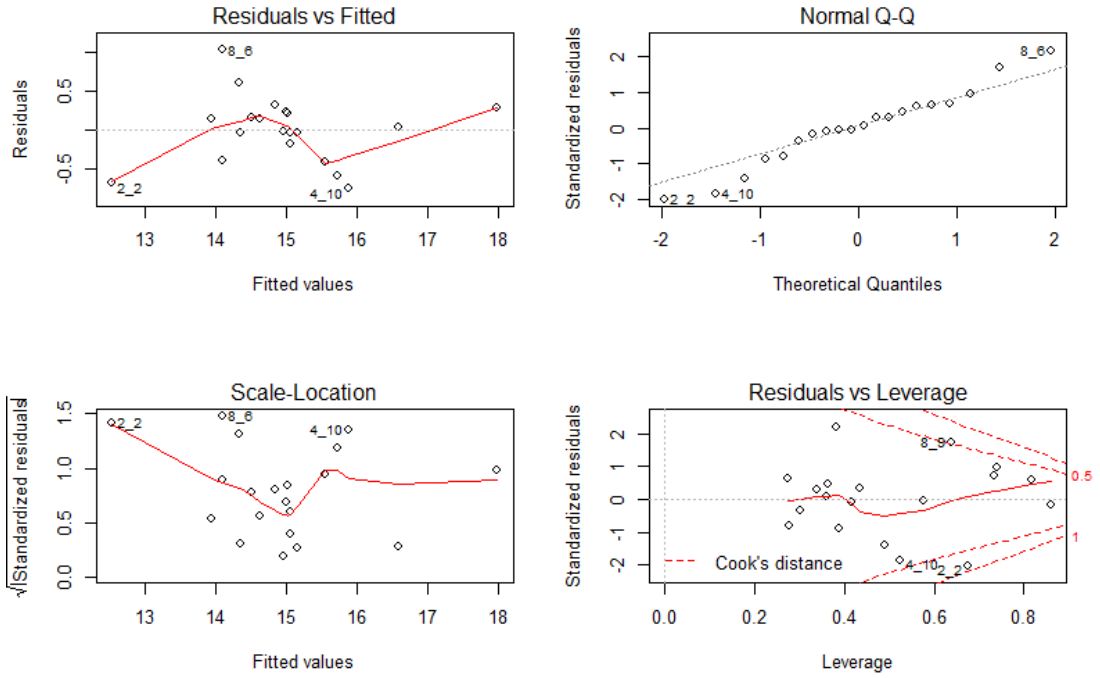

Fig. S9. LiDAR and field-derived CWM_SLA prediction

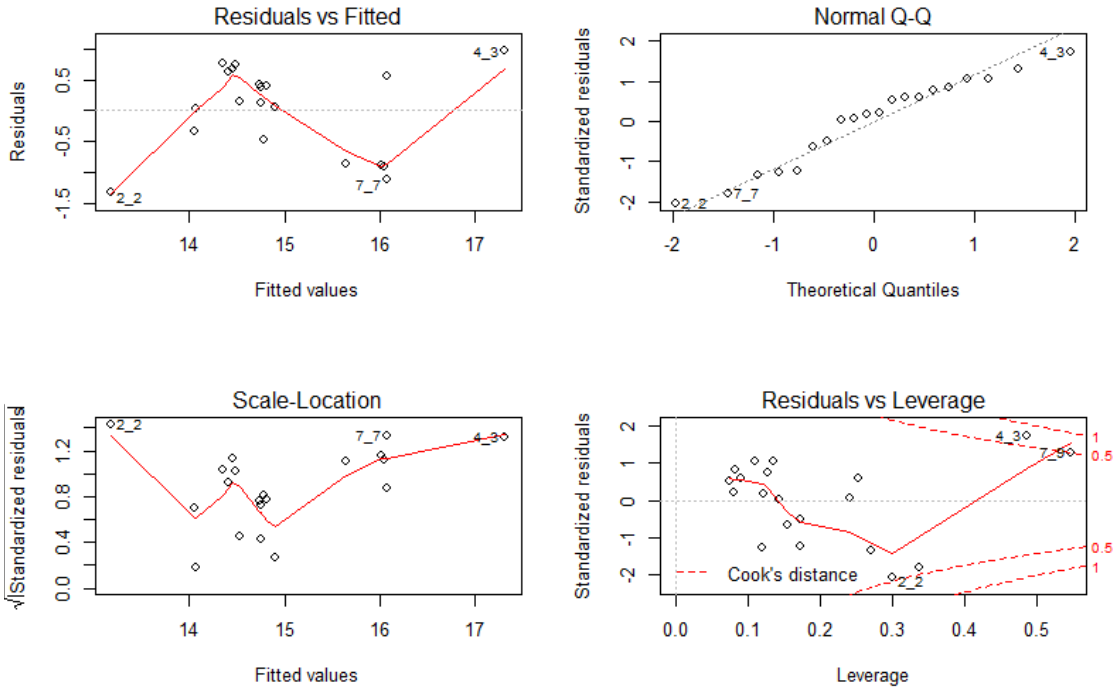

Fig. S10. LiDAR-derived CWM_SLA prediction

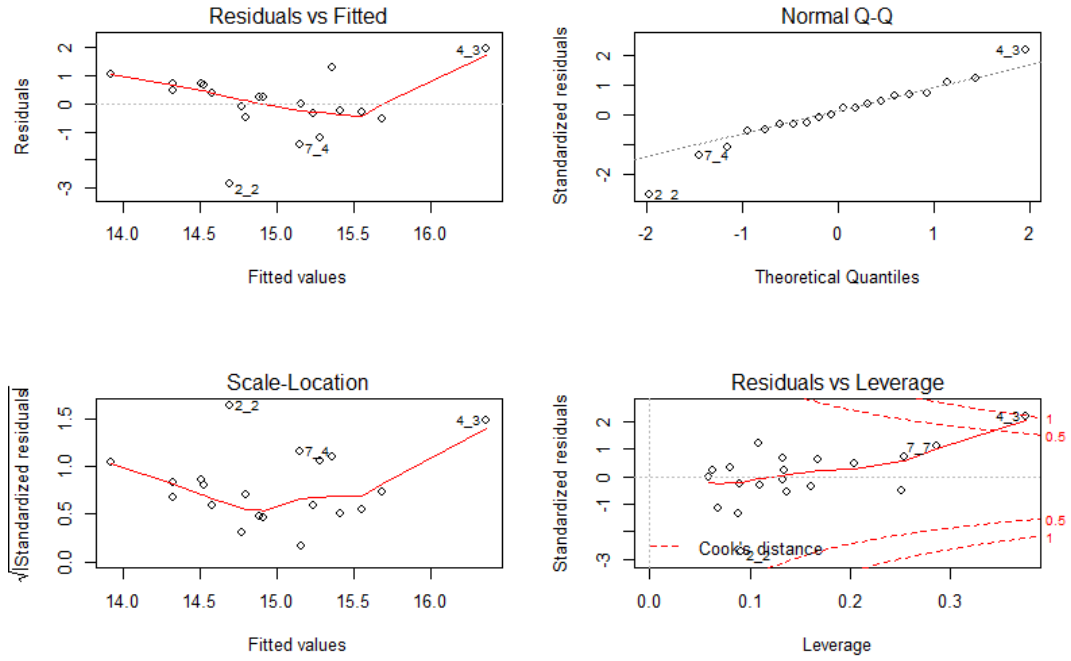

Fig. S11. Field-derived CWM_SLA prediction

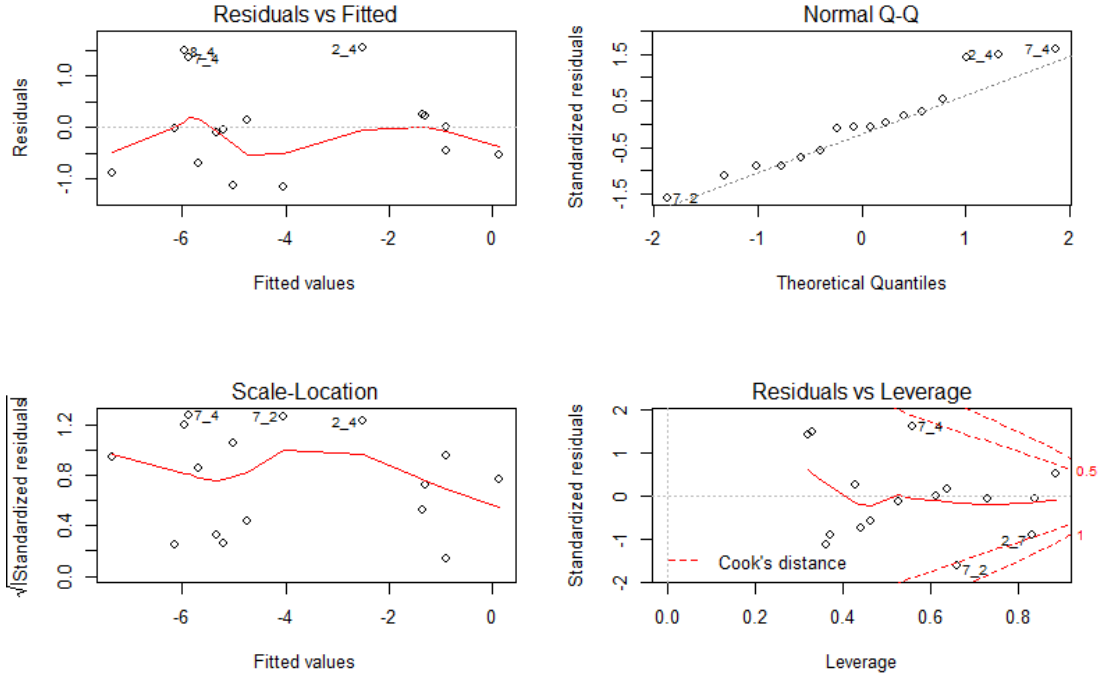

Fig. S12. LiDAR and field-derived RaoQ_SM prediction

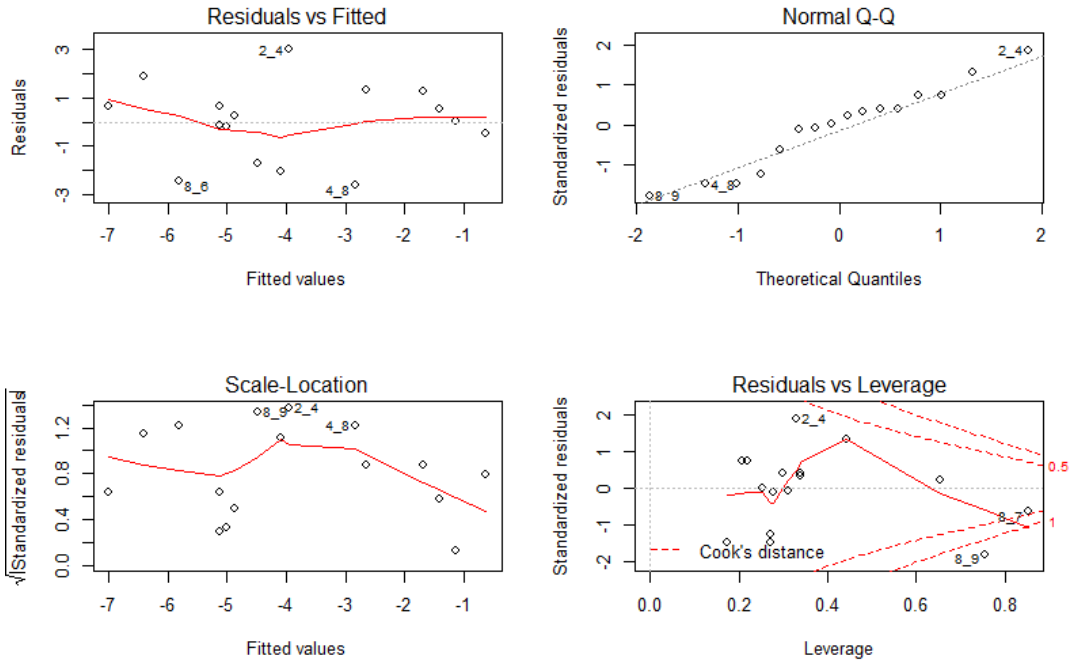

Fig. S13. LiDAR-derived RaoQ_SM prediction

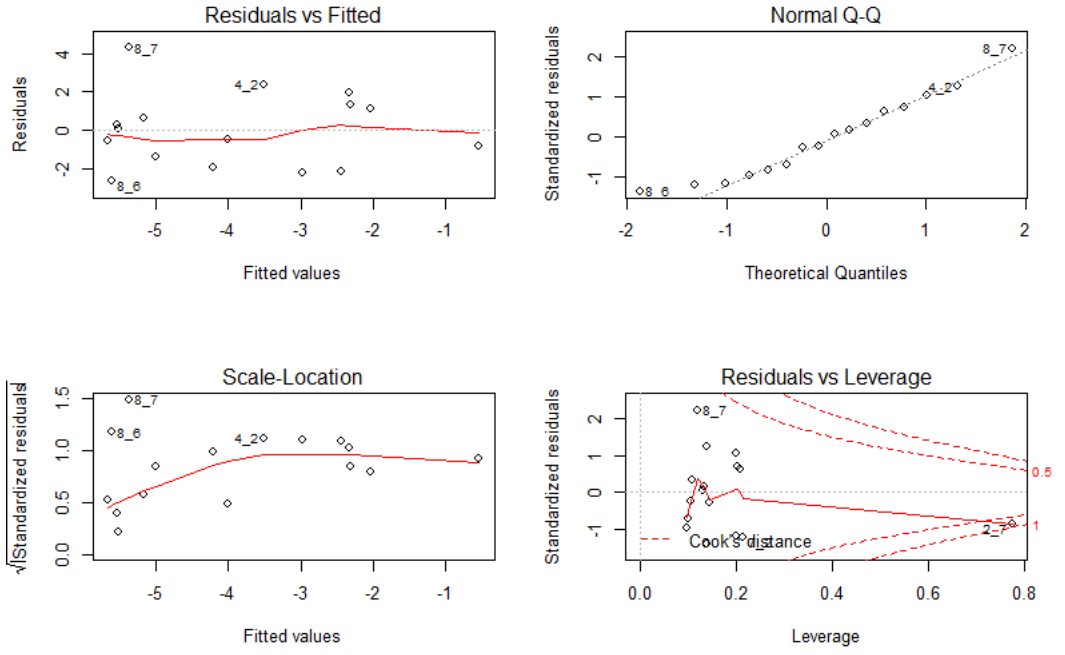

Fig. S14. Field-derived RaoQ_SM prediction

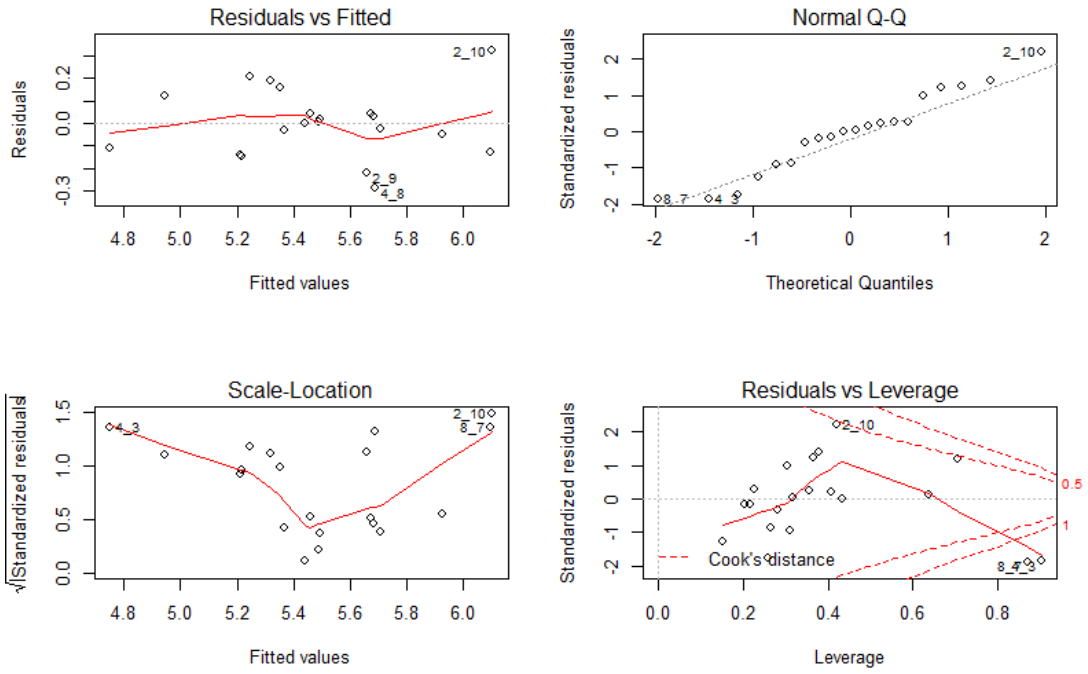

Fig. S15. LiDAR and field-derived CWM_SM prediction

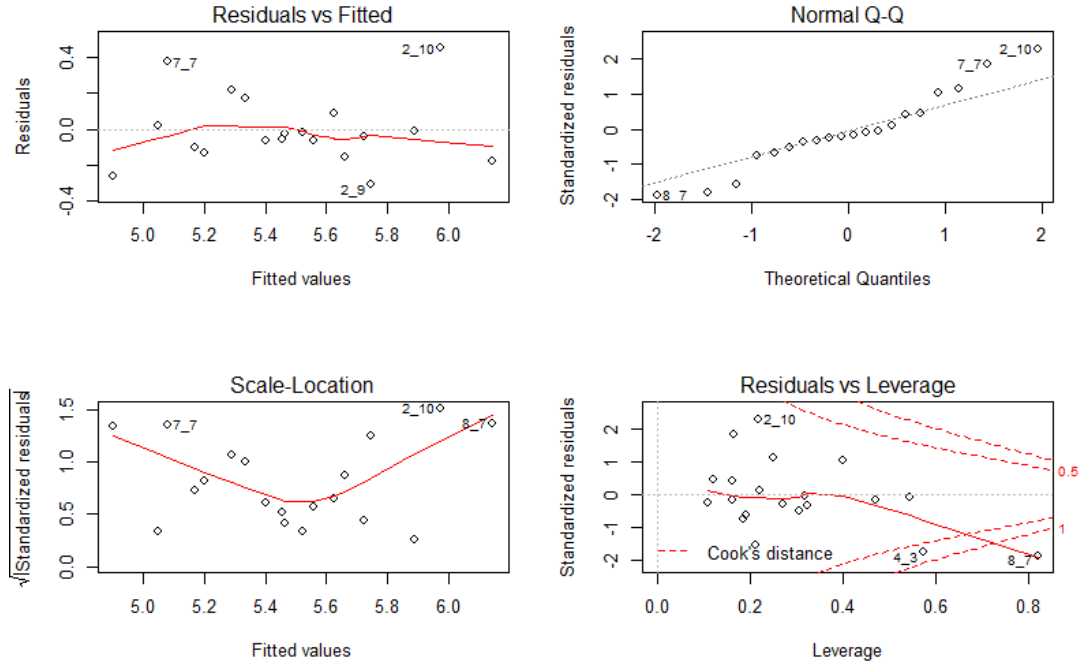

Fig. S16. LiDAR-derived CWM_SM prediction

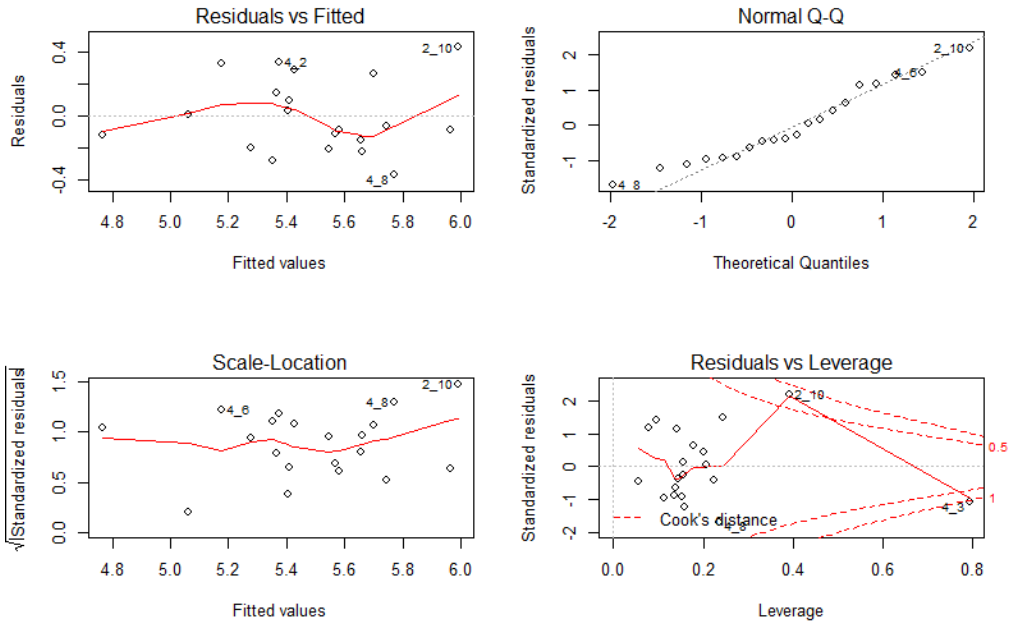

Fig. S17. Field-derived CWM_SM prediction