Article

Potential of Airborne LiDAR Derived Vegetation Structure for the Prediction of Animal Species Richness at Mount Kilimanjaro

Alice Ziegler [1,*], Hanna Meyer [2], Insa Otte [3], Marcell K. Peters [4], Tim Appelhans [1], Christina Behler [5], Katrin Böhning-Gaese [6,7], Alice Classen [4], Florian Detsch [1], Jürgen Deckert [8], Connal D. Eardley [9], Stefan W. Ferger [6], Markus Fischer [10], Friederike Gebert [11], Michael Haas [12], Maria Helbig-Bonitz [5], Andreas Hemp [13], Claudia Hemp [6], Victor Kakengi [14], Antonia V. Mayr [4,5], Christine Ngereza [15,16], Christoph Reudenbach [1], Juliane Röder [15], Gemma Rutten [10], David Schellenberger Costa [17], Matthias Schleuning [6], Axel Ssymank [18], Ingolf Steffan-Dewenter [4], Joseph Tardanico [4], Marco Tschapka [5,19], Maximilian G. R. Vollstädt [20], Stephan Wöllauer [1], Jie Zhang [4], Roland Brandl [15] and Thomas Nauss [1]

1 Environmental Informatics, Faculty of Geography, University of Marburg, Deutschhausstraße 12, 35032 Marburg, Germany; tim.appelhans@gmail.com (T.A.); fdetsch@web.de (F.D.); reudenbach@uni-marburg.de (C.R.); woellaus@staff.uni-marburg.de (S.W.); nauss@geo.uni-marburg.de (T.N.)

2 Institute Landscape Ecology, University of Muenster, Heisenbergstraße 2, 48149 Münster, Germany; hanna.meyer@uni-muenster.de

3 Department for Remote Sensing, Institute of Geography and Geology, University of Würzburg, Oswald-Külpe-Weg 86, 97074 Würzburg, Germany; insa.otte@uni-wuerzburg.de

4 Department of Animal Ecology and Tropical Biology, Biocenter, University of Würzburg, Am Hubland, 97074 Würzburg, Germany; marcell.peters@uni-wuerzburg.de (M.K.P.); alice.classen@uni-wuerzburg.de (A.C.); antonia.mayr@uni-wuerzburg.de (A.V.M.); ingolf.steffan-dewenter@uni-wuerzburg.de (I.S.-D.); joseph.tardanico@uni-wuerzburg.de (J.T.); jie.zhang@uni-wuerzburg.de (J.Z.)

5 Institute for Evolutionary Ecology and Conservation Genomics, University of Ulm, Albert-Einstein-Allee 11, 89069 Ulm, Germany; ChristinaBehler@gmx.de (C.B.); m.helbig-bonitz@gmx.de (M.H.-B.); marco.tschapka@uni-ulm.de (M.T.)

6 Senckenberg Biodiversity and Climate Research Centre (SBiK-F), Senckenberganlage 25, 60325 Frankfurt am Main, Germany; katrin.boehning-gaese@senckenberg.de (K.B.-G.); stefan.ferger@yahoo.de (S.W.F.); claudia.hemp@uni-bayreuth.de (C.H.); matthias.schleuning@senckenberg.de (M.S.)

7 Institute for Ecology, Evolution and Diversity, Goethe University Frankfurt, Biologicum, Max-von-Laue-Straße 13, 60439 Frankfurt am Main, Germany

8 Museum für Naturkunde, Leibniz Institute for Evolution and Biodiversity Science, Invalidenstraße 43, 10115 Berlin, Germany; juergen.deckert@mfn.berlin

9 School of Life Sciences, University of KwaZulu-Natal, Private Bag X01, Scottsville, Pietermaritzburg 3209, South Africa; connal.eardley@gmail.com

10 Institute of Plant Sciences, University of Bern, Altenbergrain 21, 3013 Bern, Switzerland; markus.fischer@ips.unibe.ch (M.F.); ggpmrutten@gmail.com (G.R.)

11 Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland; friederike.gebert@web.de

12 State Museum of Natural History, Department of Entomology, 70191 Stuttgart, Germany; haasmichael990@gmail.com

13 Department of Plant Systematics, University of Bayreuth, Universitätsstraße 30, 95440 Bayreuth, Germany; andreas.hemp@uni-bayreuth.de

14 Tanzania Wildlife Research Institute, Arusha P.O. Box 661, Tanzania; kakengi1@yahoo.com

15 Department of Ecology, Animal Ecology, University of Marburg, Karl-von-Frisch-Straße 8, 35032 Marburg, Germany; cngereza@yahoo.com (C.N.); juliane.roeder@biologie.uni-marburg.de (J.R.); brandlr@mailer.uni-marburg.de (R.B.)

16 National Museum of Tanzania, 06 Shaaban Robert Street, Dar es Salaam P.O. Box 511, Tanzania

17 Institute of Ecology and Evolution, Friedrich Schiller University Jena, Dornburger Strasse 159, 07743 Jena, Germany; david.schellenberger.costa@uni-jena.de

18 Federal Agency for Nature Conservation, Konstantinstr. 110, 53179 Bonn, Germany; ssymanka@t-online.de

19 Smithsonian Tropical Research Institute, Balboa Ancón, PAnama City P.O. Box 0843-03092, Panama

20 Center for Macroecology, Evolution and Climate, University of Copenhagen, Universitetsparken 15, 2100 København, Denmark; maximilian.vollstaedt@sund.ku.dk

* Correspondence indicated by asterisk

* Correspondence: alice.ziegler@geo.uni-marburg.de

Abstract: The monitoring of species and functional diversity is of increasing relevance for the development of strategies for the conservation and management of biodiversity. Therefore, reliable estimates of the performance of monitoring techniques across taxa become important. Using a unique dataset, this study investigates the potential of airborne LiDAR-derived variables characterizing vegetation structure as predictors for animal species richness at the southern slopes of Mount Kilimanjaro. To disentangle the structural LiDAR information from co-factors related to elevational vegetation zones, LiDAR-based models were compared to the predictive power of elevation models. 17 taxa and 4 feeding guilds were modeled and the standardized study design allowed for a comparison across the assemblages. Results show that most taxa (14) and feeding guilds (3) can be predicted best by elevation with normalized RMSE values but only for three of those taxa and two of those feeding guilds the difference to other models is significant. Generally, modeling performances between different models vary only slightly for each assemblage. For the remaining, structural information at most showed little additional contribution to the performance. In summary, LiDAR observations can be used for animal species prediction. However, the effort and cost of aerial surveys are not always in proportion with the prediction quality, especially when the species distribution follows zonal patterns, and elevation information yields similar results.

Keywords: biodiversity; species richness; LiDAR; elevation; partial least square regression; arthropods; birds; bats; predictive modeling

1. Introduction

We are facing a decrease of global biodiversity [1,2] and the rate of this loss is accelerating with ongoing climate change [3] as well as the rapid transformation of natural habitats by human landuse [4]. To mitigate the effects of this biodiversity loss on the functionality of ecosystems [5–7], monitoring of species and functional diversity is an important prerequisite for focused management strategies [8–10]. To facilitate a unified monitoring system, a set of essential biodiversity variables (EBVs) were developed during the last years, e.g., [11–13]. However, gathering those variables during field campaigns is only possible in a number of limited situations, as area-wide coverage is unfeasible due to high costs as well as the lack of experts. This is particularly true for surveys across large areas with a steep elevation gradient, as complex terrain hinders accessibility [14].

Facing this challenge of comprehensive mappings of EBVs [15,16], Pettorelli et al. [17] proposed a subset of the EBVs that could potentially be surveyed using satellite remote sensing. Some of these remote sensing based EBVs already meet reasonable quality requirements (e.g., land cover, leaf area index or phenology), others like species occurrences or taxonomic diversity require further research. The retrieval of the latter is often based on correlations between land-cover properties and information on taxonomic richness from field studies that is used to build remote sensing models [18–22]. Multispectral or hyperspectral retrievals primarily rely on compositional information (e.g., [23–26], reviewed in [27]). LiDAR retrievals utilize information on the vertical structure of the ecosystem, e.g., [28,29] and seem to be particularly successful for flying organisms [30–39], although the biological background of this observation is poorly understood.

To advance the application of species richness related EBVs from remote sensing, performance must be compared across many taxa. While meta-analyses across different case studies allow some conclusions, the individual study design, different computations of error estimates and the uniqueness of the study regions make it difficult to actually compare the results regarding the model performance for different taxa. This becomes even more challenging if the study has been conducted in mountainous terrain, which is common for ecological space-for-time approaches as global biodiversity hotspots often tend to be in mountainous areas [40]. Here, elevational change of the land cover in combination

with a fixed amount of work force generally limits the number of ground samples that can be collected within one land cover zone. These individual limitations in training and testing datasets lead to a variety of testing approaches with varying degrees of reliability and comparability of error estimates.

This study analyzes the predictive performance of airborne LiDAR-derived variables for mapping the species richness of 17 taxonomic groups from four feeding guilds in a comprehensive manner. The study area is located at the southern slopes of Kilimanjaro (Figure 1) and field observations stretch from an elevation of 800 to 4400 m. Since the taxonomic assemblages cannot be directly observed using LiDAR, vegetation structure is used as a surrogate for species richness. It is assumed that the taxa or the aggregated feeding guilds can be predicted differently well by LiDAR data. For example, the "plant diversity hypothesis", links consumer richness (especially herbivores) to plant diversity [41–43]. Therefore, it is expected that the performances of structural models decline from herbivores to predators [14,44], as the distance of their position in the food chain to plants increases. Furthermore, as structural properties, at scales accessible for remote sensing, tend to be more relevant for animals with larger body-sizes [45–48], it is expected that the performance of species richness models increases with increasing body size. Similar consideration of correlation between structure and species traits apply for flying taxa that perceive the landscape with a grain not as detailed as walking taxa (environmental grain hypothesis, e.g., Kaspari and Weiser [49], Sarty et al. [50]). The unique dataset also allows for critically evaluating whether LiDAR-derived information brings any gain at all compared to models that rely solely on the known decrease of species richness with elevation or elevation correlated environmental properties [43]. This study investigates the potential of LiDAR-derived variables characterizing vegetation structure as predictors for animal species richness at the southern slopes of Mount Kilimanjaro. To disentangle the structural LiDAR information from co-factors related to elevational vegetation zones, LiDAR-based models were compared to the predictive power of elevation models.
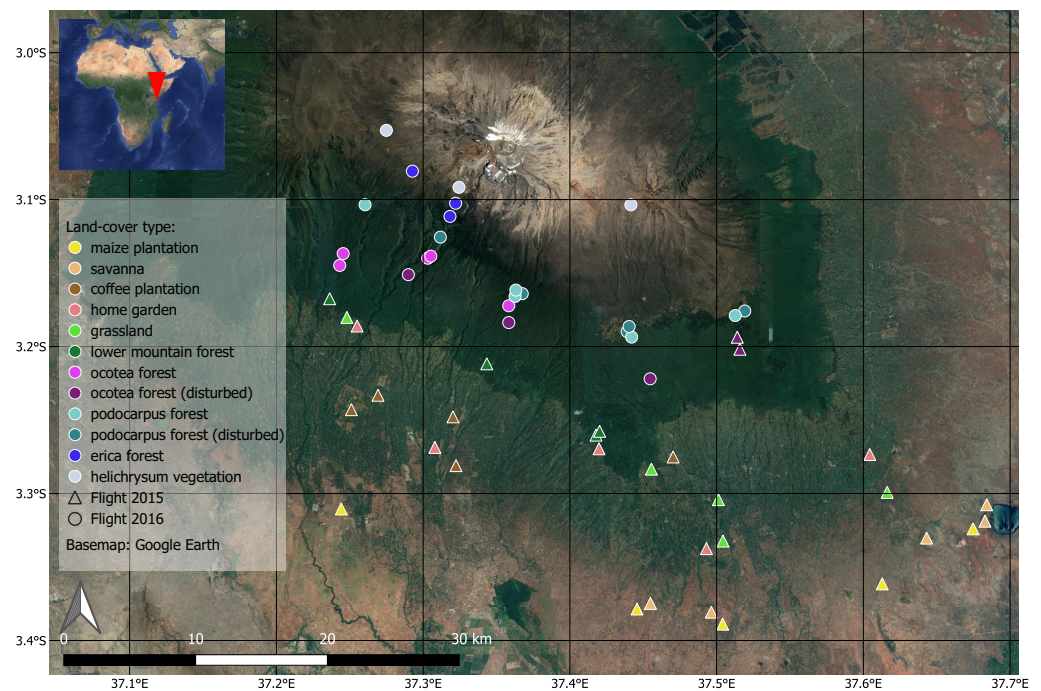

Figure 1. Study area with sampling plots. Colors of symbols show different land covers, shapes show the different flight missions from 2015 and 2016. The background image indicates the large-scale vegetation zones along the elevational gradient (background: Google Maps [51]).

2. Materials and Methods

2.1. Study Area and Sampling Design

The study area is the southern slopes of Mount Kilimanjaro. In the framework of a comprehensive research unit, study plots of 50 m × 50 m were established across 12 land cover zones with five replica plots per zone. The 60 plots describe two ecological gradients: an elevational gradient from 800 m a.s.l. to 4400 m a.s.l. and a disturbance gradient from (near) natural to anthropogenic land-cover types within elevation zones (Figure 1 and Table A1).

2.2. Data and Preprocessing

2.2.1. Diversity Data

On 59 of the 60 study plots, data for estimating the diversity of 17 taxa were available. Sampling followed standardized approaches as described in detail in Peters et al. [43] and Peters et al. [52]. For an overview of the taxa and sampling methods, see Table A2. The species richness data were aggregated (sum of species) to four feeding guilds. See Tables A3 and A4 for the proportional allocation of species per taxon to feeding guild. As true bugs, spiders and springtails were not identified to the species level, the entire group of spiders was assumed to be predators, springtails count as decomposers and true bugs were ignored when estimating species richness of feeding guilds.

2.2.2. LiDAR Data

The LiDAR data set was acquired during two missions with a Riegl LMS-Q780 sensor carried by an Airbus Helicopter at an altitude between 850 m and 1750 m above ground level. The northern (high land) plots were sampled in March 2015 and the southern (low land) plots in November 2016 (Figure 1). The temporal distance was assumed to be negligible since both acquisition dates fall into the early rainy season and plots of the same land-cover type are covered within the same flight campaign except for disturbed ocotea forests. The LiDAR pulses contain between one and seven returns with a vertical accuracy of 0.15 m and a horizontal accuracy of 0.20 m (95% confidence interval). The mean point density is 34 points per square meter but varies due to terrain and flight conditions. Outliers were removed and points were classified into ground and non-ground following the standard procedure using the LAStools preprocessing software [53]. Rasterized LiDAR layers (e.g., digital terrain model (DTM), digital surface model (DSM), and canopy height model (CHM)) were generated by the open source remote sensing data base (RSDB) at a resolution of 1 m [54]. The resulting DTM preserves fine details in regions with high ground point densities and plausible elevation estimates in regions with low ground point densities (e.g., dense forest).

To derive a set of potential predictor variables from the LiDAR observations, several indices, which characterize structural properties of the 50 m × 50 m areas, were computed for each of the plots using the RSDB (see Table 1) [54]. The compiled 97 LiDAR metrics used included e.g., canopy height metrics (maximum, standard deviation, median, quartiles, etc.), return number metrics (maximum, standard deviation of different layers, etc.) and ecological estimates (leaf area index, above-ground biomass and gap fraction, etc.; for the complete list of variables, see Table 1).

Table 1. Overview of structural variables characterizing the vegetation and a description on their calculation. Most indices were calculated on the LiDAR (Light Detection And Ranging) point cloud of each plot (50 m × 50 m); only a few were calculated on 1 m × 1 m cells of the canopy height model (marked: based on CHM).

Name	Explanation
canopy	
maximum canopy height (CH)	Maximum canopy height
mean CH	Mean canopy height
median CH	Median canopy height
percentile of CH	10% Percentile of canopy heights
. . . (in 10% steps)	x% Percentile of canopy heights
standard deviation CH	Standard Deviation of canopy height
skewness CH	Skewness of canopy height Distribution
variance CH	Variance of canopy height
curtosis CH	Excess Kurtosis of canopy height Distribution
coefficient of variation CH	Coefficient of Variation of canopy height
area ratio	area ratio of raster pixels (based on CHM) [55]
vegetation structure	
Return Density (RD) of different layers	Return density of 1 m layer
. . . (1 m steps up to 8 m/29 m)	Return density of x meter layer
RD canopy (>5 m)	Return density of canopy vegetation layer
RD regeneration (2–5 m)	Return density of regeneration vegetation layer
RD understory (<2 m)	Return density of understory vegetation layer
RD ground	Return density of ground layer
penetration rate (PR) of layers	Penetration rate of 1 m layer
. . . (1 m steps up to 8 m/29 m)	Penetration rate of x meter layer
PR canopy (>5 m)	Penetration rate of canopy vegetation layer
PR regeneration (2–5 m)	Penetration rate of regeneration vegetation layer
PR understory (<2 m)	Penetration rate of understory vegetation layer
maximum returns	highest number of return points per LiDAR laser pulse
mean returns	mean of return points per LiDAR laser pulse
standard deviation returns	standard deviation of return points per LiDAR laser pulse
median returns	median of return points per LiDAR laser pulse
standard deviation first return	standard deviation of first return points per plot

Table 1. *Cont.*

Name	Explanation
vegetation	
above ground biomass	Aboveground Biomass ($6.85 * TCH^{0.952}$) (top-of-canopy height (TCH) based on CHM) [56]
foliage height diversity	Foliage height diversity (Shannon Index grouped by layers of understory, regeneration and canopy) [19]
leaf area index	Leaf-area index, with k = 0.3, h.bin = 1, GR.threshold = 5 (based on CHM) [56]
vegetation coverage (VC) of different layers	vegetation coverage in 1 m height (based on CHM)
. . . (for 2, 5, and 10 m)	vegetation coverage in x meter height (based on CHM)
gap fraction	fraction of clear area above 10 m compared to whole area (to be detected as gap: minimum of 9 cells, based on CHM)

The land covers in the study area can be grouped into forest and non-forest (see Figure 1 and Table A1 for details). Due to their complex multi-layered structure, forested plots appear considerably different to non-forested plots hence the sets of LiDAR variables for these two types differ slightly. For the current study, this means that, on non-forested plots, variables describing vegetation layers reached a maximum of 8 m height, and on forest plots, vegetation height reached a maximum of 29 m (indicated in Table 1). The two thresholds correspond to variables where at least 50% of the plots had vegetation in this height. In the following, all modeling approaches were always carried out for forested and non-forested plots separately, to account for these fundamental differences.

2.3. Predictive Modeling of Diversity

The computations and analyses in this study were performed using the R environment 3.5 in conjunction with the caret package [57,58]. Partial least squares regression (PLSR) is useful for models in data settings with a smaller number of observations relative to the number of predictor variables. It can also handle multicolinearity, a situation that is unavoidable, when using LiDAR-derived variables [59–62]. To reduce the impact of overfitting caused by correlated variables, a forward feature selection (FFS) implemented in the CAST package [63] was used, which ensures a more stable variable selection than recursive feature elimination approaches [64].

To distinguish between effects on species richness predictability based on a pure elevation gradient versus habitat structure, three different model groups (elevation and its square only, structural metrics only and structural metrics to predict residuals of elevation based model) were established. Then, within each of these three groups of models, an individual prediction model was separately built for each taxon and feeding guild for forested and non-forested areas. The same combinations of plots for training and testing were used across all models (Table 2 for overview).

Table 2. Overview of the different models calculated in this study.

Model Name	Independent Variables (Predictors)	Target Variables (Response)
elevation	elevation, elevation ^{2}	species richness
structure	structural LiDAR variables	species richness
residuals	structural LiDAR variables	residuals from elevation model
combination	sum of elevation model and structural model	species richness

In the first group of models, only the elevation and its square were used to predict species richness ("elevation model"). In the second group of models, only the structural metrics derived from LiDAR (no elevation) were considered ("structure model"). In the third group of models, the same structural metrics were used to predict the residuals of the elevation model ("residual model"). Hence, the predictions of the residual model do not represent the complete species richness, but only that part which cannot be explained by the elevation model. Therefore, to be able to compare the results of this model with the elevation and structure model, the results of the residual model were added to the elevation model ("combination model"). Even though it is not a separate model in a strict sense, but only the sum of the elevation model and the residual model, this mixed approach will be called "combination model" in the following. The pure residual model, on the other hand, can be used to compare the plain structure dependence of taxa and feeding guilds without effects of elevation. Prediction results from forested and non-forested areas were assembled to one error estimation per response variable, to compare the general model performance of taxa and feeding guilds for the whole study area.

To test if the model performance depends on species traits, the correlation of model performance of each taxa and feeding guild to the respective body size and the mode of movement were tested. For body size, the Spearman rank correlation coefficient was calculated. Groups were sorted by body size from large (large mammals with up to 1.7 m length) to small groups (parasitoid wasps with only a few millimeters). For the test between

model performance and flying/non-flying groups of organisms, the Mann–Whitney U-Test was performed.

Validation Strategy and Model Tuning

Due to the limited number of observation samples per taxonomic group, choosing an appropriate tuning and testing strategy of the various models was of major importance. As illustrated in Figure 2, model training and testing consisted of two separate cross validation cycles. The outer 20-fold-cross validation withholds one random plot of each land-cover type in every resample. Those samples were held back from model training to qualify them for estimating the model performance for new locations in the study region. This repeated approach allows for more stable validation results given the limited number of plots. The inner cross validation was embedded within the PLSR machine learning approach. It uses the same method of leaving one plot per land-cover type out in each resample. The inner cross validation was used for model tuning and variable selection only. Tuning affected the number of principal components used in the PLSR and varied between one and two. Feature selection was implemented according to Meyer et al. [64].

For quantifying the predictive performance, the root mean squared error (RMSE) was computed for each fold of the outer cross validation. Previously, the results from forested and non-forested areas were combined, to be able to make a general statement per taxon and feeding guild. Since species richness varied considerably across the taxonomic groups, the RMSE of each group was normalized with the standard deviation of the species richness per group of the plots used in each model.
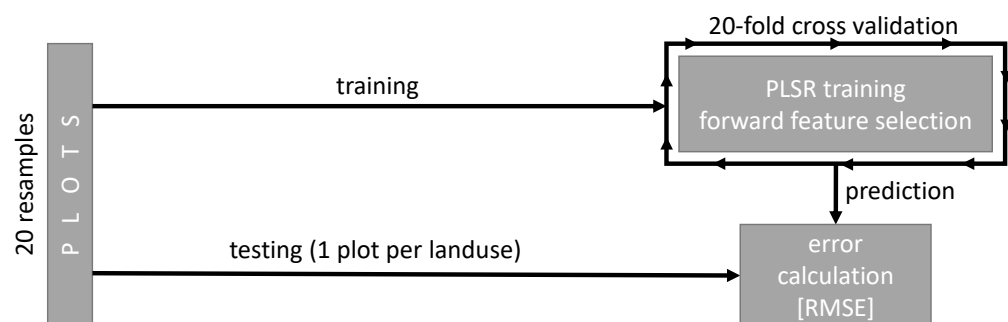

Figure 2. The model training (upper right loop) uses a partial least squares regression (PLSR) and a forward feature selection with a 20-fold cross validation. Validation is carried out by predicting the values of the testing plots. The division of testing and training plots (outer loop) follows a repeated stratified sampling approach, with randomly chosen resamples of one plot per land cover for the testing, leaving the rest of the plots for training. Validation is based on the median root mean square error (RMSE) of the individual resamples, normalized by the standard deviation of these RMSE values.

3. Results

The elevation model performs best for 14 out of 17 taxa (Figure 3a and Table 3). On average, the RMSE/sd values for these 14 taxa are 0.21 lower than in the structural and 0.23 lower than in the combined model. The structural model performs better only for parasitoid wasps and the combined model is the best for aculeate wasps and insectivorous bats.

For all three model types, the interquartile range (IQR) of large mammals, springtails, bees, parasitoid wasps and insectivorous bats is rather small, while syrphid flies, moths, dung beetles and grasshoppers show large variations of RMSE/sd values. Only large mammals, millipedes and springtails show a significant superior model performance for the best model (here elevation) compared to both other models (Tukey test). For an individual taxon, a median performance of the RMSE/sd of half the standard deviation or better is only reached for ants, grasshoppers, springtails, bees, parasitoid wasps, other aculeate wasps and insectivorous bats. Bees reach the best model results across all taxa and model types with an RMSE/sd of 0.34 (elevation model).
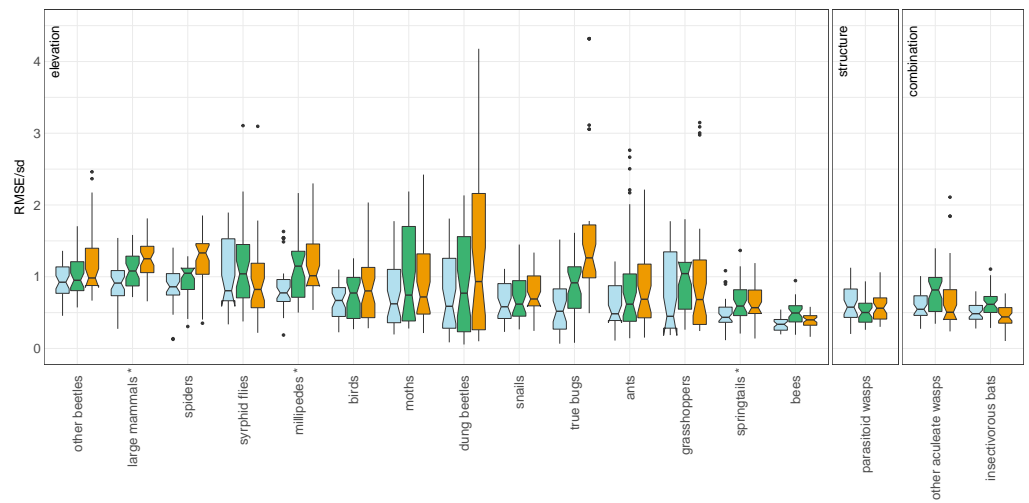
(**a**)
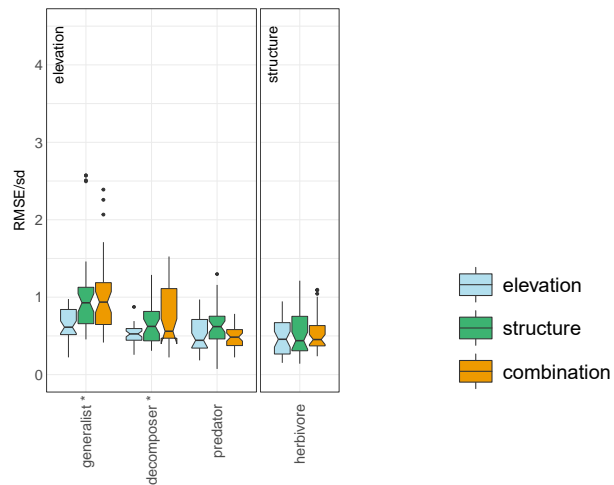
(**b**)

Figure 3. Modeling performances for each taxon (**a**) and feeding guild (**b**) in terms of the root mean square error normalized by standard deviation (RMSE/sd). Smaller values show a better model performance. Colors represent the different model types. Taxa are grouped into "elevation", "structure" and "combination" depending on which of the three models shows the best median RMSE/sd. Stars indicate if the best model is significantly better than both of the other models. Within the groups, taxa and feeding guilds are sorted by descending RMSE/sd. The boxes include the median and the inter quartile range (IQR) with notches indicating roughly the 95% confidence interval. Whiskers are extending to ±1.5 times the IQR and points indicate single error values outside of this range.

Table 3. Mean ± standard deviation across all plots of observed species richness per plot and median of the root mean square error normalized by standard deviation of field samples per taxon and feeding guild ± the standard deviation of model runs. Lower RMSE/sd values mean a better model performance.

Taxon/Feeding Guild	Species Richness per Plot	Elevation Model	Structure Model	Combined Model	Residual Model
	Mean ± Standard Deviation	Median RMSE/sd ± Standard Deviation			
ants	2.7 ± 3.4	0.48 ± 0.29	0.62 ± 0.72	0.69 ± 0.57	1.5 ± 1.3
bees	6.7 ± 5.9	0.34 ± 0.1	0.49 ± 0.17	0.4 ± 0.11	1.5 ± 0.4
birds	16 ± 8.6	0.67 ± 0.25	0.77 ± 0.31	0.8 ± 0.45	1.9 ± 1
dung beetles	5.2 ± 7.7	0.59 ± 0.5	0.77 ± 0.73	0.93 ± 1.3	1.6 ± 2.2
grasshoppers (locusts, crickets)	10 ± 14	0.45 ± 0.54	1 ± 0.41	0.68 ± 0.82	1.1 ± 1.3
insectivorous bats	5.5 ± 3.3	0.48 ± 0.13	0.62 ± 0.18	0.44 ± 0.14	1.4 ± 0.46
large mammals	2.3 ± 1.8	0.91 ± 0.28	1.1 ± 0.24	1.3 ± 0.26	2.2 ± 0.45
millipedes	1.2 ± 1.7	0.77 ± 0.34	1.1 ± 0.48	1 ± 0.43	1.5 ± 0.65
moths	9.6 ± 11	0.62 ± 0.48	0.74 ± 0.69	0.72 ± 0.61	1.2 ± 1
other aculeate wasps	3.2 ± 4	0.55 ± 0.21	0.82 ± 0.26	0.5 ± 0.45	1.2 ± 1.1
other beetles	8.6 ± 4.9	0.92 ± 0.24	0.95 ± 0.27	0.98 ± 0.46	1.7 ± 0.81
parasitoid wasps	16 ± 14	0.58 ± 0.26	0.5 ± 0.18	0.56 ± 0.19	1.1 ± 0.38
snails (slugs)	6.8 ± 5.6	0.58 ± 0.26	0.62 ± 0.31	0.69 ± 0.28	1.5 ± 0.61
spiders	5 ± 2.2	0.86 ± 0.26	1 ± 0.23	1.3 ± 0.39	2.5 ± 0.72
springtails	4.6 ± 2.4	0.43 ± 0.2	0.59 ± 0.27	0.57 ± 0.25	1.4 ± 0.63
syrphid flies	3.5 ± 3.3	0.8 ± 0.47	1 ± 0.55	0.82 ± 0.5	1.2 ± 0.74
true bugs	2.1 ± 2.3	0.57 ± 0.39	0.91 ± 0.37	1.3 ± 1.1	2.4 ± 2.2
decomposer	12 ± 7.8	0.53 ± 0.13	0.62 ± 0.25	0.56 ± 0.38	1.8 ± 1.2
generalist	6.8 ± 3.8	0.61 ± 0.21	0.93 ± 0.57	0.94 ± 0.48	2.2 ± 1.2
herbivore	40 ± 30	0.46 ± 0.23	0.44 ± 0.3	0.45 ± 0.26	1.3 ± 0.77
predator	41 ± 22	0.44 ± 0.22	0.62 ± 0.23	0.48 ± 0.15	1.2 ± 0.37

Regarding feeding guilds, the elevation model performs best for generalists, decomposers and predators with an RMSE/sd value that is 0.20 and 0.13 lower on average than the structural and the combined model, respectively (Figure 3b and Table 3). Only for generalists and decomposers was the best model (elevation) performing significantly better than the other two. The structural model performs best for herbivores but only with slight differences in the RMSE/sd to the combined (0.01) and the elevation model (0.02).

To explore the potential of modeling species richness outside the gradient of Mount Kilimanjaro, Figure 4 shows a comparison of the plain residual models (RMSE/sd). These results are independent of the elevation and do not model species richness, but the residuals of the elevation model. Therefore, it is possible to compare the ranking of species performance as it would be suspected if it was only dependent on structure without a superimposing elevational gradient. Taxa and feeding guilds are sorted by their median error estimates which range between 1.1 and 2.5 (Table 3). Smaller values here mean a closer relationship to structural metrics. Value ranges of model performances within each group lie within the same magnitude, except for dung beetles which show a high variation. The RMSE/sd of the residual model shows a ranking of the feeding guilds from predators (1.2), over herbivores (1.3), to decomposers (1.8) and generalists (2.2). The analysis of the best subsets of prediction variables could not identify regular patterns (Figures A1–A4).

There is no statistical relationship between model performance and body size of the assemblages (Table 4). However, there is a difference for the combined and residual model with a better performance for the flying than for the non-flying taxonomic groups (see Table 4).
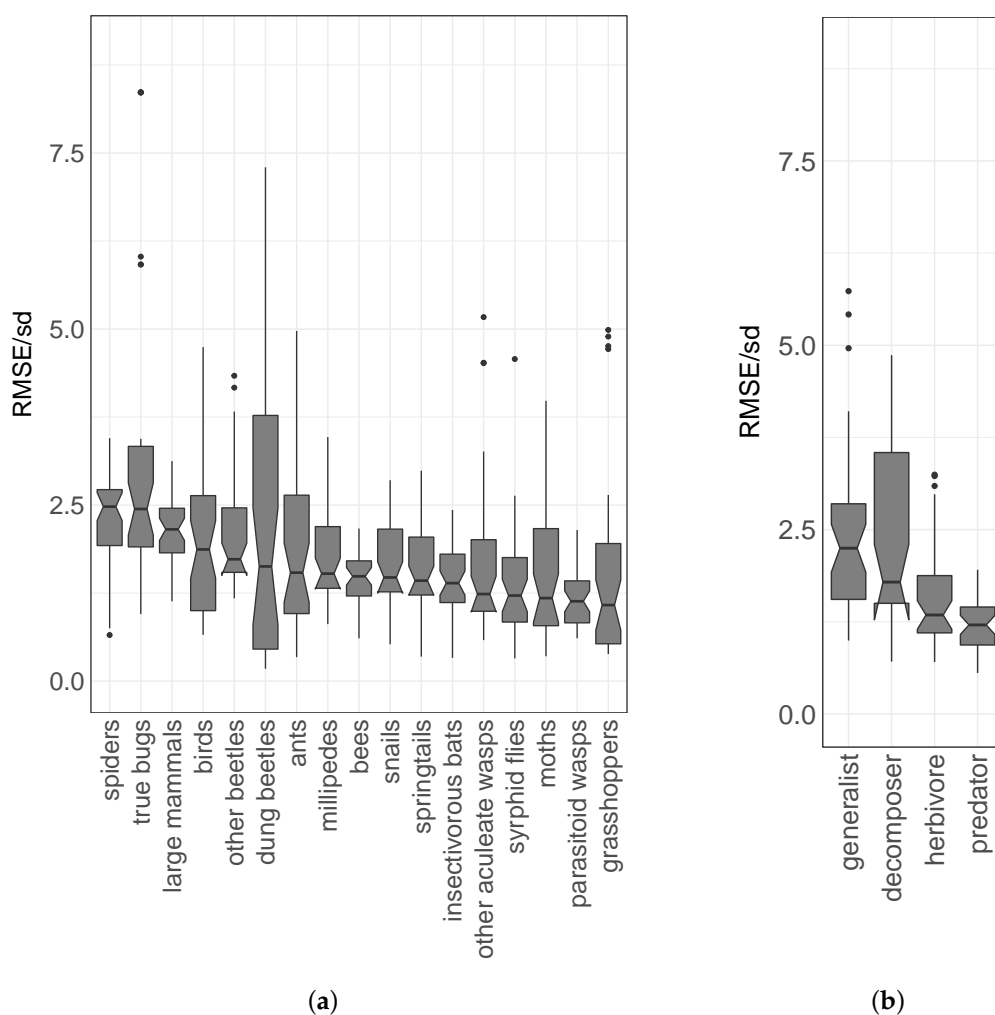
(**a**)   (**b**)

Figure 4. Modeling performance for the residuals of the elevation model for each taxon (**a**) or feeding guild (**b**) as root mean square error normalized by standard deviation [RMSE/sd]. Taxa (**a**) and feeding guilds (**b**) are sorted by increasing median modeling performance and therefore increasing influence of vegetation structure on the target variable (which means decreasing median RMSE/sd). For the description of plot elements, see Figure 3.

Table 4. Results of rank tests comparing the performances of models measured by RMSE/sd (as shown in Figures 3 and 4) with respect to selected traits (body size, mode of movement) of the taxa. For the tests between the performance of the models and body size, the Spearman rank correlation coefficient (r) was used. Body size was sorted from large to small groups. For the test between flying and non-flying groups of organisms the Mann–Whitney U-Test was used. Significant results, in terms of the *p*-value, are marked bold.

	Body Size		**Mode of Movement**	
	r	*p*	*p*	
RMSE/sd elevation	−0.14	0.58	0.28	
RMSE/sd structure	−0.13	0.62	0.28	
RMSE/sd combination	−0.10	0.70	**0.011**	**(flying <non-flying)**
RMSE/sd residuals	−0.22	0.39	**0.006**	**(flying <non-flying)**

4. Discussion

The study evaluated the potential of LiDAR data to predict species richness at Mount Kilimanjaro. The influence of the respective effects of elevation and vegetation structure on

species richness were investigated by comparing the model performances of models that used elevation as the only predictor, models that used LiDAR variables only and models that used LiDAR variables to predict the residuals of the elevation models.

Generally, performances of the different models varied only slightly within each taxon, with no significant difference of the best performing model to both other models, except for three taxa (large mammals, millipedes, springtails). However, there is a trend, indicating that the elevation model performs best for 14 out of 17 taxa. All taxa which do perform significantly better with one specific model, belong to that group. In the cases where the structure or combined model performs best, the performances differ only marginally and differences are not significant.

As expected, results of the model performances of feeding guilds indicate that herbivores are influenced more by structure than generalists and decomposers (Figure 4b). However, considering the feeding guilds, generalists and decomposers are the only groups for which the best of the three models (elevation model) is significantly better than the other two. It is ecologically reasonable that generalists, which obviously use a wide variety of food, are, at least for feeding reasons, not specifically connected to the vegetation structure. Opposed to that, for herbivores, structure was suspected to be the most relevant predictor as they rely solely on vegetation and therefore structure should influence feeding patterns and the occurrence and diversity of species Even though performances improve slightly with the structure model, differences were not significant. Decomposers rely on the existence of organic material. Still, as long as the supply of organic material is given, it seems reasonable that other environmental factors, which are linked to elevation, would have a greater impact. In conclusion, prediction results for the feeding guilds show a tendency of the hypothesized correlation between a lower feeding guild and a higher dependency on the structure. However, these differences are very small and therefore not convincing. The model performances of the feeding guilds are generally comparable with the ones of the individual taxa. Nevertheless, as more field samples are included in the feeding guilds, sampling uncertainties are partly leveled out by the higher number of sampled individuals.

This study further aimed at comparing the general potential of LiDAR-derived variables for the prediction of the structurally dependent proportion of species richness for different taxa and feeding guilds. For this comparison, the elevation corrected residual model provides the relevant information (Figure 4). In line with the discussion about the best model type, generalists and decomposers seem to be the group not tightly connected to habitat structure, whereas herbivores seem to depend more on vegetation structure (Figure 3a). Along with the other models, there is no notable difference in the overall performances between taxa and feeding guilds for the residual models.

A comparison between the model types allows for drawing conclusions about the influence of elevation and structure as relevant predictors for biodiversity. In their study at Mount Kilimanjaro about diversity gradients at different levels of aggregation, Peters et al. [43] already showed that mean annual temperature is the most important variable to predict animal species richness in the region. In the present study, some taxa are significantly more influenced by elevation than by properties of the structure itself, but generally, median performances between models differ only slightly. The residual model attempts to illuminate patterns within the remaining structural properties that are not attributable to the strong gradient in the study area. However, only samples from four to five replica plots per land-cover type have been available for model training which limits the performance. This might promote over-fitted models for structural properties leading to larger prediction errors when applied to unknown locations, whereas the elevation model is able to find a general pattern within all plots as they are well distributed along the elevation gradient (Figure 1). Still, even a slightly worse structural or combined model compared to the elevation model validates the general usability of LiDAR data for predicting species richness, even though the effort of LiDAR missions then seems questionable. In the variable selection of the LiDAR metrics, no patterns emerge (Figures A1–A4). Neither individual variables nor variable groups (Table 1) appear in clear patterns across models. The LiDAR

variables were calculated for individual plots (50 m × 50 m). For some taxa, it might be beneficial to account for the structure of a larger spatial environment. Therefore, in future studies, it could be tested whether variable cell sizes of the LiDAR metrics can improve prediction models.

The hypothesized positive correlation between body size and the modeling performance is not supported by the data. However, the mode of movement significantly correlates with the prediction performance in the combined and the residual models. Especially in the residual model, flying taxa outperform the others. The six taxa with the smallest median error are species with the ability to fly. Only the flying taxa bees and birds (Rank 9 and 14 out of 17) lie within the worse performing half of taxa, showing the generally poorer performance of non-flying taxa. The comparably poor performance of predicting birds with structural metrics alone is rather surprising, as birds are the most studied taxonomic group in species–habitat structure relationships [65] and there are many studies that demonstrate promising correlations of bird diversity and different structural features (e.g., Müller et al. [35], Smart et al. [66], Vogeler et al. [67] or see the detailed review of Davies and Asner [68]).

The results of this study are based on 59 plots. Even though the total number of study sites seem to be sufficient for modeling purposes, compared to similar studies, the different land-cover types that follow the elevation gradient, in addition to the necessary division into forest and non-forest areas for modeling, limit the number of repetitions. Hence, model building has to be carefully adjusted to the limited number of plots. The possibility of also using land-cover type as a categorical predictor variable was discarded due to the low number of replicates. However, land cover is indirectly included in the model by the natural orientation of land cover along the elevational gradient at Mt. Kilimanjaro.

In general, it is not easy to evaluate the results in the context of other studies, since a comparison of the results can only provide indications of the success of the modeling. This is because the studies were conducted in different landscapes, for different taxa, but most importantly, with different measures of biodiversity. Species richness, beta diversity, and other metrics are related but not identical. In previous studies, the role of elevation is handled in different ways. The studies of Müller and Brandl [14] and Vierling et al. [69], for example, analyze the influence of LiDAR-derived variables compared to other abiotic and biotic variables for the prediction of spider species distribution and forest beetle assemblages. Results show a comparable or even much better performance of LiDAR variables to ground based measures [69]. As the variable elevation is a by-product of LiDAR point clouds, these studies included elevation in the group of LiDAR-derived variables, with elevation being a rather important variable. However, elevation changes within the study area are limited to about 800 meters with a rather homogeneous forest cover.

The studies of Zellweger et al. [39] and Rechsteiner et al. [37] are situated in a more mountainous terrain (>1200 m elevation difference); however, the LiDAR derived variables are limited to structural ones and elevation is not used as a predictor, although elevation is intrinsically included in structural variables at least along elevational vegetation zones. With similar complex terrain (around 4000 m elevation difference), Zellweger et al. [70] used structural as well as topographic and climate variables. Even though elevation was not used directly, climatic variables (including temperature) showed the highest importance for modeling beta-diversity of birds and butterflies, even exceeding results when vegetation structure was included in the models. This seems consistent with the results of the present study, given that elevation is a main proxy for temperature [71]. Overall, all these studies show clearly that the elevation gradient might be able to explain a major part of structural variables. An observation in a similar sense is made by Acebes et al. [65] in their review of 173 papers. They find, especially for forested areas, that, while canopy height is most commonly used as a LiDAR metric to model species-habitat structure, canopy cover and terrain topography performed better overall when they where used.

The study of Müller et al. [72] covers an elevation gradient of around 800 m and does only take vertical profile metrics derived from LiDAR data into account. They could show

that canopy arthropod diversity is driven by different structural features in the vegetation. Using a similar number of study plots as in our study, but exclusively in spruce forest, the overall vegetation structure is much more homogeneous than at Mt. Kilimanjaro. Thus, finer differences are likely to be masked by the large variability in vegetation structure in the models. The same is true for the study of Schooler and Zald [73], who analyzed the predictability of small mammals diversity in temperate mixed forest and found that it could be predicted by LiDAR derived structural metrics.

Therefore, when using LiDAR data, non-structural properties (e.g., elevation, temperature, or other abiotic variables—depending on the study area) should be investigated separately to avoid false conclusions concerning the effect of LiDAR-derived vegetation structure. Those abiotic conditions are relevant for modeling and therefore models from one study area are not necessarily representative in other areas [38]. Using a separate residual model shows great potential to avoid spurious correlation that leads to erroneous predictions when the model is applied to new locations.

At Mount Kilimanjaro, with its substantial elevation gradient, the utilization of LiDAR data does not significantly improve modeling results. A larger sampling size per land cover is required to further improve the robustness of conclusions drawn for the selection of models. To approach the long-term goal of comprehensive mapping of EBVs like species occurrence or taxonomic diversity with the use of remotely sensed data, areas with a less complex land-cover gradient in homogenous landscapes need to be addressed in future studies to understand the influence of structure better.

To provide comparable results, further studies need to be conducted on multi-taxon approaches with field surveys and data sets of similar granularity. Study areas with different terrain complexities should be considered. In doing so, a solid base for valuable model-building strategies can be generated and can assist the research community in quantifying EBVs in the future.

Appendix A

Table A1. Overview of land-cover types with elevation range of their occurrence, the number of plots in each land cover and the basic ecosystem category (forest or non-forest).

Land Cover	Elevation [m a.s.l]	Number of Plots	Ecosystem
maize plantation	866–1009	5	non-forest
savanna	871–1153	5	non-forest
coffe plantation	1124–1648	5	non-forest
homegarden	1169–1788	5	non-forest
grassland	1303–1748	5	non-forest
lower mountain forest	1560–2040	5	forest
ocotea forest	2120–2750	5	forest
ocotea forest (disturbed)	2220–2560	5	forest
podocarpus forest	2720–2970	5	forest
podocarpus forest (disturbed)	2770–3060	5	forest
erica forest	3500–3880	5	forest
helichrysum vegetation	3880–4390	4	non-forest

Table A2. General sampling methods as well as some details and the calculation of the species richness of the biodiversity data. Further details on the sampling approaches can be found in Peters et al. [43] and Peters et al. [52]. If not indicated otherwise, the species richness is calculated as the total (cumulative) number of species per study site (equal sampling effort for all sites).

Taxon	General Sampling Method	Sampling Specifics	Species Richness Calculation
ants	plastic tubes with diverse set of resource baits on the ground	2 h at times of peak ant activity	
bees	pan traps in different vegetation heights	48 h with 3 sampling rounds	
birds	audiovisual point counts	15 min before sunrise, completed before 9 am	
dung beetles	baited pitfall trap	72 h sampling time	
grass hoppers	sightings and two rounds of sweep net sampling	non-forested ares: repeatedly 1.5 h walking on parallel tracks, forested areas: 1.5 h shaking understory vegetation	sampling effort per site was adjusted to measure asymptotic species richness
insectivorous bats	acoustic monitoring (point stop method)	every corner of plot visited for 5 min	averaging species richness values of single surveys
large mammals	camera trap, analysis of dung remains	5 camera trapsper site, 70 trap-days per site	number of all non-domestic mammal species recorded on study site
millipedes	pitfall traps and sightings	sightings: 2 h	
moths	light traps	20 min with 4 sampling rounds on obstacle-free branch in 1.5–2 m height	averaging species richness values of single surveys
other aculeate wasps	pan traps in different vegetation heights	48 h with 3 sampling rounds	
other beetles	pitfall traps	7 days sampling time with a total of 5 sampling rounds	
parasitoid wasps	pan traps in different vegetation heights	48 h with 3 sampling rounds	
snails	sightings (large taxa) and collection of leaf litter (small taxa)	sightings: four rounds of fixed time surveys of 30 min, collection: 1 litre leaf litter	
spiders	pitfall traps	7 days sampling time	
springtails	pitfall traps	7 days sampling time	
syrphid flies	pan traps in different vegetation heights	48 h with 3 sampling rounds	total (cumulative) number of species richness with varying number of samples among study sites
true bugs	sweep net sampling	100 sweeps along two 50 m transects	

Table A3. Fractional breakdown of feeding guilds. The numbers for each taxon represent the relative contribution to the species richness of a given feeding guild. Therefore, each column adds up to a total of 1. Taxa are listed in alphabetic order. True bugs were not identified to the species level and are therefore not included.

	Decomposer	Generalist	Herbivore	Predator
ants	0	0.24	0.02	0.02
bees	0	0	0.14	0
birds	0	0.24	0.08	0.1
dung beetles	0.41	0	0	0
grasshoppers	0	0.03	0.17	0.01
insectivorous bats	0	0	0	0.02
large mammals	0	0.11	0.01	0
millipedes	0.16	0	0	0
moths	0	0	0.47	0
other aculeate wasps	0	0	0	0.12
other beetles	0.24	0.37	0.07	0.17
parasitoid wasps	0	0	0	0.51
snails	0.08	0.02	0.03	0.02
syrphid flies	0.11	0	0	0.02

Table A4. Fractional breakdown of taxa. The numbers for each feeding guild represent the relative contribution of species richness to a given taxon. Therefore, each row adds up to a total of 1. Taxa are listed in alphabetic order. True bugs were not identified to the species level and are therefore not included.

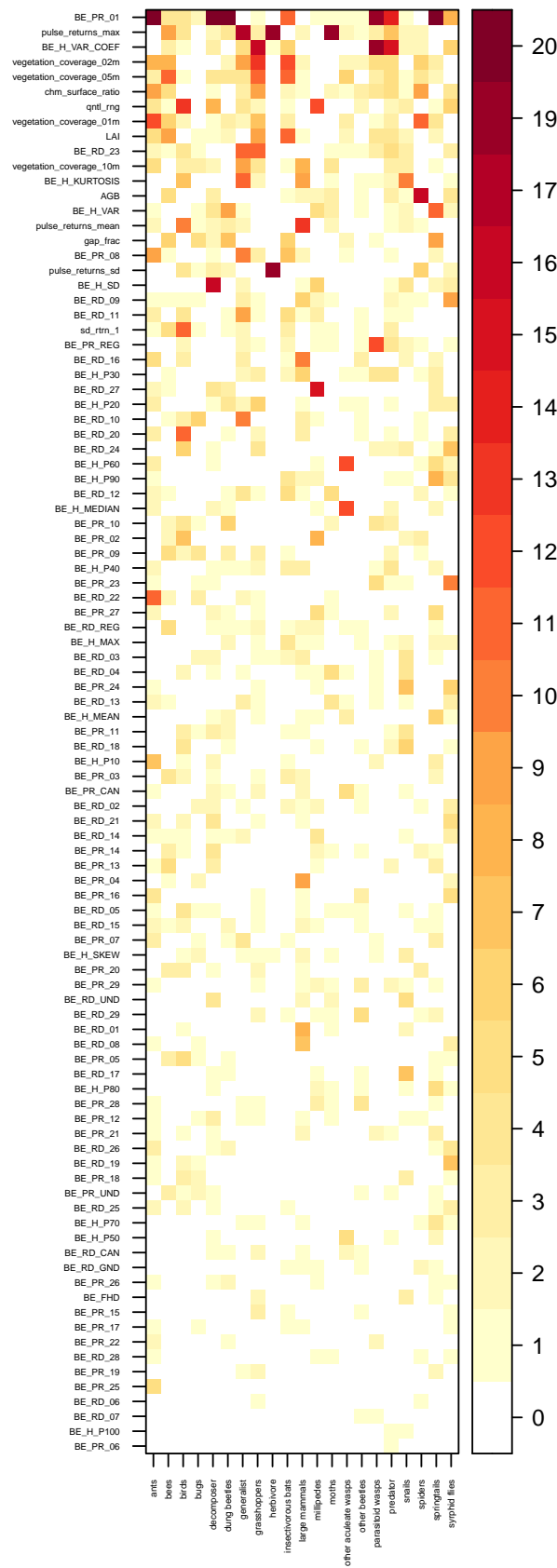|	Decomposer	Generalist	Herbivore	Predator
ants	0	0.47	0.24	0.29
bees	0	0	1	0
birds	0	0.18	0.31	0.51
dung beetles	1	0	0	0
grasshoppers	0	0.03	0.9	0.08
insectivorous bats	0	0	0	1
large mammals	0	0.52	0.33	0.15
millipedes	1	0	0	0
moths	0	0	1	0
other aculeate wasps	0	0	0	1
other beetles	0.13	0.17	0.17	0.53
parasitoid wasps	0	0	0	1
snails	0.24	0.05	0.41	0.31
syrphid flies	0.44	0	0.09	0.47

Figure A1. Variable selection for the structure model of each taxon in forest. Colors show how often variables were included during 20-fold cross-validation. Structural variables are sorted by the total number of times they where selected. See Woellauer et al. [54] for variable details.

Figure A2. Variable selection for structural variables for the residual model and combined model of each taxa in forest. Colors show how often variables were included during 20-fold cross-validation. Structural variables are sorted by the total number of times they where selected. See Woellauer et al. [54] for variable details.
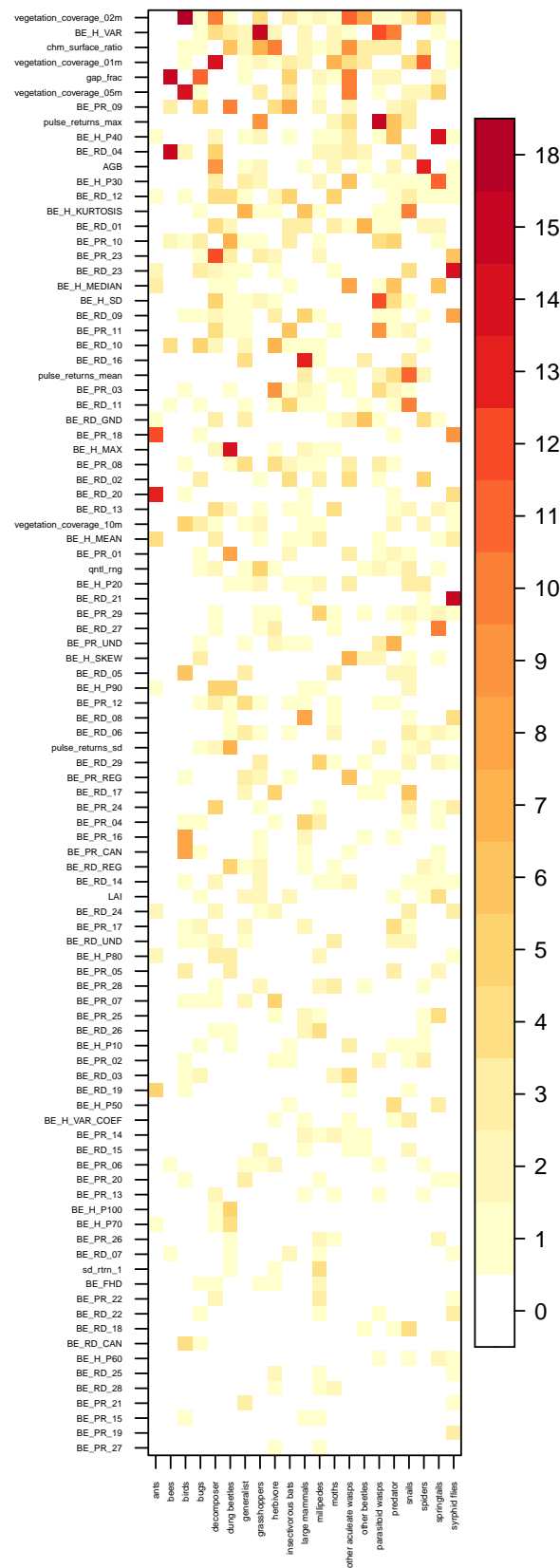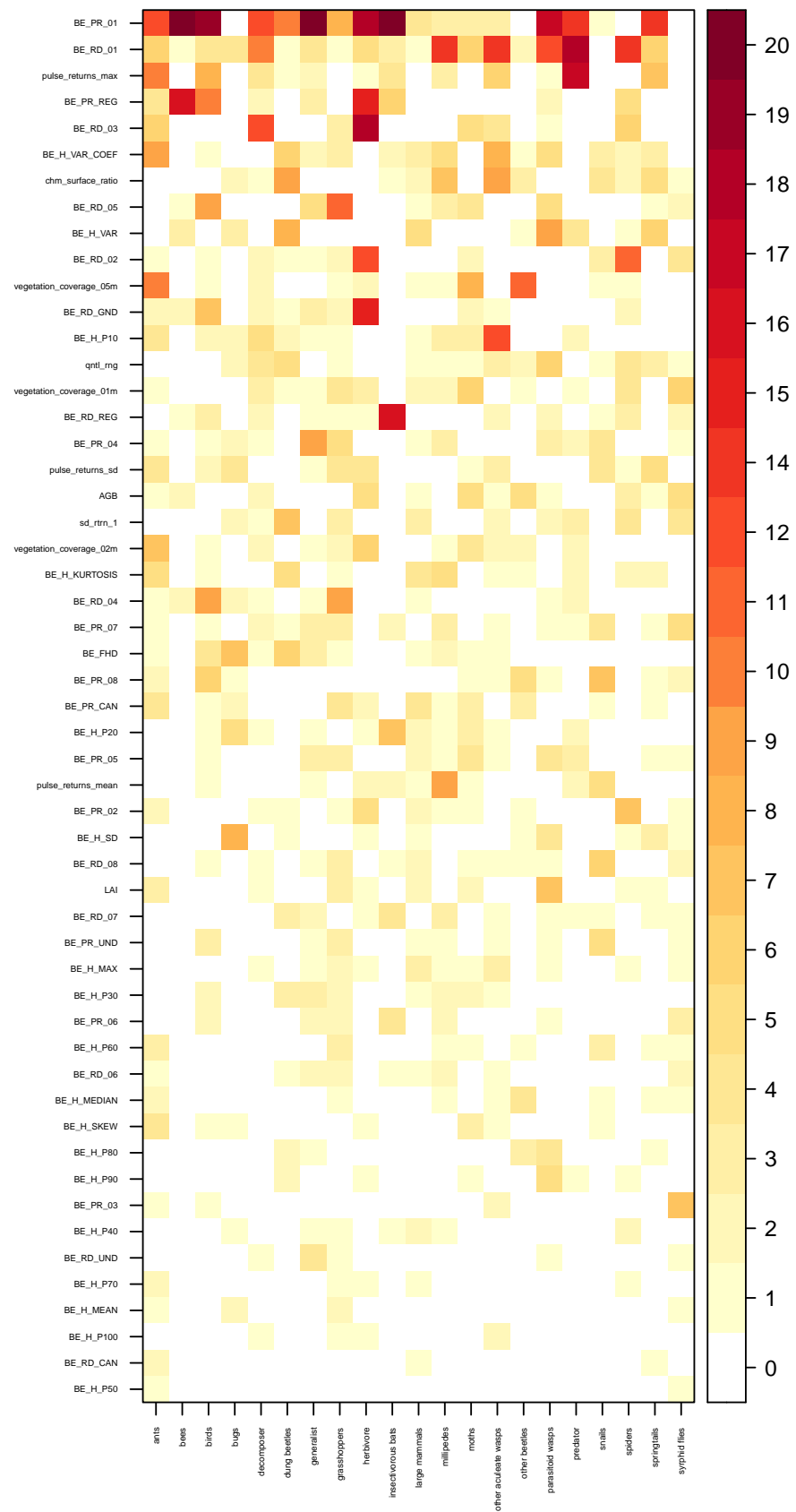
Figure A3. Variable selection for the structure model of each taxon in non-forest. Colors show how often variables were included during 20-fold cross-validation. Structural variables are sorted by the total number of times they where selected. See Woellauer et al. [54] for variable details.

Figure A4. Variable selection for structural variables for the residual model and combined model of each taxa in non-forest. Colors show how often variables were included during 20-fold cross-validation. Structural variables are sorted by the total number of times they where selected. See Woellauer et al. [54] for variable details.
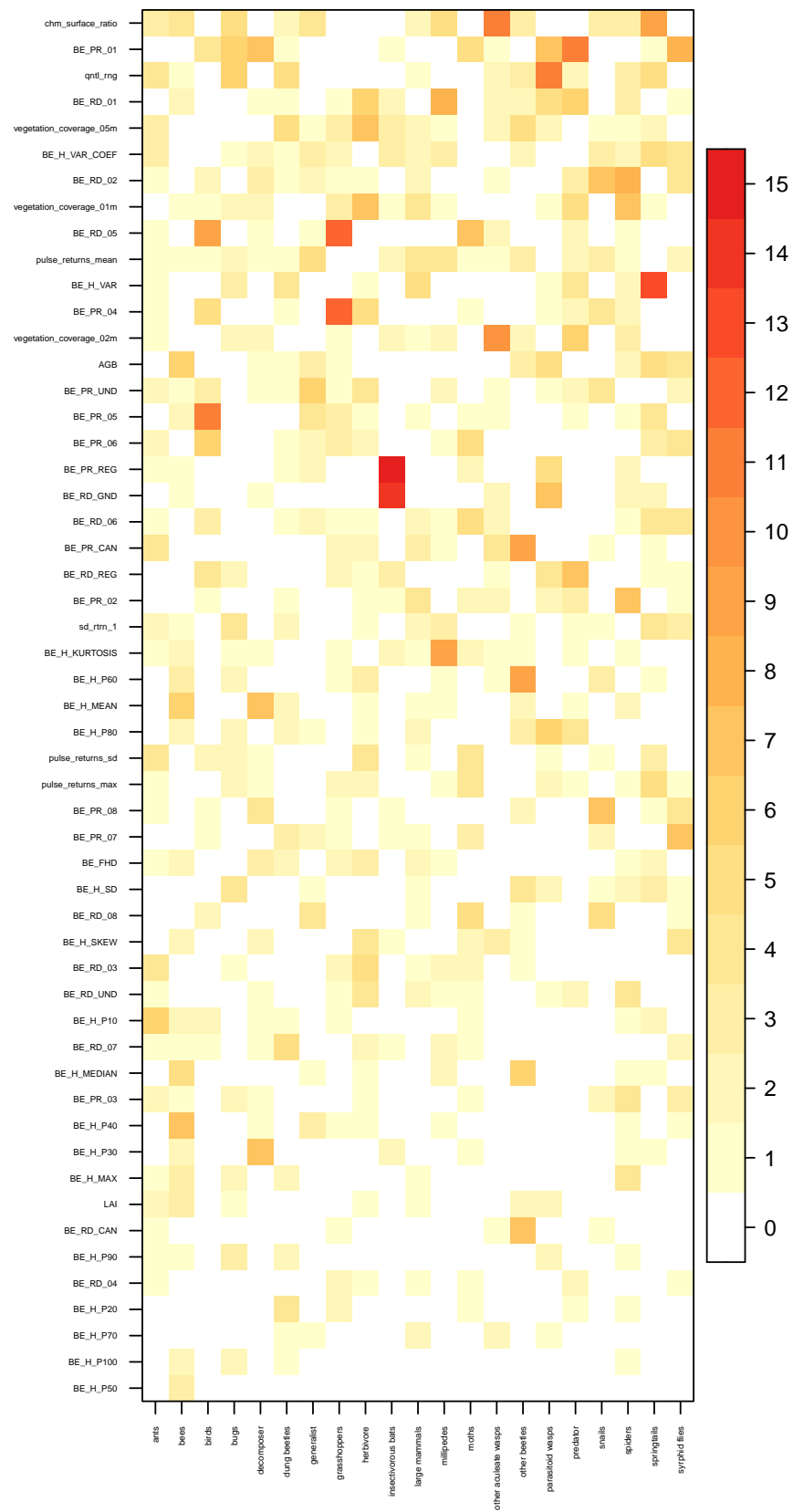
References

1. Ceballos, G.; Ehrlich, P.R.; Dirzo, R. Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E6089–E6096. [CrossRef]
2. Pereira, H.M.; Navarro, L.M.; Martins, I.S. Global Biodiversity Change: The Bad, the Good, and the Unknown. *Annu. Rev. Environ. Resour.* **2012**, *37*, 25–50. [CrossRef]
3. Warren, R.; Price, J.; Graham, E.; Forstenhaeusler, N.; VanDerWal, J. The projected effect on insects, vertebrates, and plants of limiting global warming to 1.5 °C rather than 2 °C. *Science* **2018**, *360*, 791–795. [CrossRef] [PubMed]
4. Newbold, T.; Hudson, L.; Hill, S.; Contu, S.; Lysenko, I.; Senior, R.; Börger, L.; Bennett, D.; Choimes, A.; Collen, B.; et al. Global effects of land use on local terrestrial biodiversity. *Nature* **2015**, *520*, 45–50. [CrossRef] [PubMed]
5. Hooper, D.; Chapin, F.; Ewel, J.; Hector, A.; Inchausti, P.; Lavorel, S.; Lawton, J.; Lodge, D.; Loreau, M.; Naeem, S.; et al. Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. *Ecol. Monogr.* **2005**, *75*, 3–35. [CrossRef]
6. Loreau, M.; Naeem, S.; Inchausti, P.; Bengtsson, J.; Grime, J.; Hector, A.; Hooper, D.; Huston, M.; Raffaelli, D.; Schmid, B.; et al. Ecology-Biodiversity and ecosystem functioning: Current knowledge and future challenges. *Science* **2001**, *294*, 804–808. [CrossRef]
7. Soliveres, S.; van der Plas, F.; Manning, P.; Prati, D.; Gossner, M.M.; Renner, S.C.; Alt, F.; Arndt, H.; Baumgartner, V.; Binkenstein, J.; et al. Biodiversity at multiple trophic levels is needed for ecosystem multifunctionality. *Nature* **2016**, *536*, 456. [CrossRef]
8. Nagendra, H.; Lucas, R.; Honrado, J.P.; Jongman, R.H.G.; Tarantino, C.; Adamo, M.; Mairota, P. Remote sensing for conservation monitoring: Assessing protected areas, habitat extent, habitat condition, species diversity, and threats. *Ecol. Indic.* **2013**, *33*, 45–59. [CrossRef]
9. Noss, R. Indicators for Monitoring Biodiversity—A Hierarchical Approach. *Conser. Biol.* **1990**, *4*, 355–364. [CrossRef]
10. Wiens, J.A. Landscape Ecology as a Foundation for Sustainable Conservation. *Landsc. Ecol.* **2009**, *24*, 1053–1065. [CrossRef]
11. Haase, P.; Tonkin, J.D.; Stoll, S.; Burkhard, B.; Frenzel, M.; Geijzendorffer, I.R.; Häuser, C.; Klotz, S.; Kühn, I.; McDowell, W.H.; et al. The next generation of site-based long-term ecological monitoring: Linking essential biodiversity variables and ecosystem integrity. *Sci. Total Environ.* **2018**, *613–614*, 1376–1384. [CrossRef] [PubMed]
12. Jetz, W.; McGeoch, M.; Robert, G.; Ferrier, S.; Beck, J.; Costello, M.; Fernández, M.; Geller, G.; Keil, P.; Merow, C.; et al. Essential biodiversity variables for mapping and monitoring species populations. *Nat. Ecol. Evol.* **2019**, *3*, 539–551. [CrossRef] [PubMed]
13. Pereira, H.M.; Ferrier, S.; Walters, M.; Geller, G.N.; Jongman, R.H.G.; Scholes, R.J.; Bruford, M.W.; Brummitt, N.; Butchart, S.H.M.; Cardoso, A.C.; et al. Essential Biodiversity Variables. *Science* **2013**, *339*, 277–278. [CrossRef]
14. Müller, J.; Brandl, R. Assessing biodiversity by remote sensing in mountainous terrain: The potential of LiDAR to predict forest beetle assemblages. *J. Appl. Ecol.* **2009**, *46*, 897–905. [CrossRef]
15. Kissling, W.D.; Ahumada, J.A.; Bowser, A.; Fernandez, M.; Fernández, N.; García, E.A.; Guralnick, R.P.; Isaac, N.J.B.; Kelling, S.; Los, W.; et al. Building essential biodiversity variables (EBVs) of species distribution and abundance at a global scale. *Biol. Rev.* **2018**, *93*, 600–625. [CrossRef]
16. Proença, V.; Martin, L.J.; Pereira, H.M.; Fernandez, M.; McRae, L.; Belnap, J.; Böhm, M.; Brummitt, N.; García-Moreno, J.; Gregory, R.D.; et al. Global biodiversity monitoring: From data sources to Essential Biodiversity Variables. *Biol. Conserv.* **2017**, *213*, 256–263. [CrossRef]
17. Pettorelli, N.; Wegmann, M.; Skidmore, A.; Mucher, S.; Dawson, T.P.; Fernandez, M.; Lucas, R.; Schaepman, M.E.; Wang, T.; O'Connor, B.; et al. Framing the concept of satellite remote sensing essential biodiversity variables: Challenges and future directions. *Remote Sens. Ecol. Conserv.* **2016**, *2*, 122–131. [CrossRef]
18. Froidevaux, J.S.P.; Zellweger, F.; Bollmann, K.; Jones, G.; Obrist, M.K. From field surveys to LiDAR: Shining a light on how bats respond to forest structure. *Remote Sens. Environ.* **2016**, *175*, 242–250. [CrossRef]
19. Macarthur, R.; Macarthur, J.W. On Bird Species Diversity. *Ecology* **1961**, *42*, 594–598. [CrossRef]
20. Martins, A.C.M.; Willig, M.R.; Presley, S.J.; Marinho-Filho, J. Effects of forest height and vertical complexity on abundance and biodiversity of bats in Amazonia. *Forest Ecol. Manag.* **2017**, *391*, 427–435. [CrossRef]
21. Melin, M.; Hinsley, S.A.; Broughton, R.K.; Bellamy, P.; Hill, R.A. Living on the edge: Utilising lidar data to assess the importance of vegetation structure for avian diversity in fragmented woodlands and their edges. *Landsc. Ecol.* **2018**, *33*, 895–910. [CrossRef]
22. Rocchini, D.; Balkenhol, N.; Carter, G.A.; Foody, G.M.; Gillepsie, T.W.; He, K.S.; Kark, S.; Levin, N.; Lucas, K.; Luoto, M.; et al. Remotely sensed spectral heterogeneity as a proxy of species diversity: Recent advances and open challenges. *Ecol. Inform.* **2010**, *5*, 318–329. [CrossRef]
23. Asner, G.P.; Martin, R.E. Spectranomics: Emerging science and conservation opportunities at the interface of biodiversity and remote sensing. *Glob. Ecol. Conserv.* **2016**, *8*, 212–219. [CrossRef]
24. Mairota, P.; Cafarelli, B.; Labadessa, R.; Lovergine, F.; Tarantino, C.; Lucas, R.; Nagendra, H.; Didham, R. Very high resolution Earth Observation features for monitoring plant and animal community structure across multiple spatial scales in protected areas. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *37*, 100–105. [CrossRef]
25. Nagendra, H.; Rocchini, D. High resolution satellite imagery for tropical biodiversity studies: The devil is in the detail. *Biodivers. Conserv.* **2008**, *17*, 3431–3442. [CrossRef]
26. Wallis, C.I.B.; Brehm, G.; Donoso, D.A.; Fiedler, K.; Homeier, J.; Paulsch, D.; Suessenbach, D.; Tiede, Y.; Brandl, R.; Farwig, N.; et al. Remote sensing improves prediction of tropical montane species diversity but performance differs among taxa. *Ecol. Indic.* **2017**, *83*, 538–549. [CrossRef]

27. Pettorelli, N.; Laurance, W.; O'Brien, T.; Wegmann, M.; Nagendra, H.; Turner, W. Satellite remote sensing for applied ecologists: Opportunities and challenges. *J. Appl. Ecol.* **2014**, *51*, 839–848. [CrossRef]
28. Lefsky, M.A.; Cohen, W.B.; Parker, G.G.; Harding, D.J. Lidar Remote Sensing for Ecosystem Studies. *Bioscience* **2002**, *52*, 19. [CrossRef]
29. Müller, J.; Vierling, K., Assessing Biodiversity by Airborne Laser Scanning. In *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies*; Maltamo, M., Næsset, E., Vauhkonen, J., Eds.; Springer: Dordrecht, The Netherlands, 2014; pp. 357–374. [CrossRef]
30. Clawges, R.; Vierling, K.T.; Vierling, L.A.; Rowell, E. The use of airborne LiDAR to assess avian species diversity, density, and occurrence in a pine/aspen forest. *Remote Sens. Environ.* **2008**, *112*, 2064–2073. [CrossRef]
31. Flashpohler, D.J.; Giardina, C.P.; Asner, G.P.; Hart, P.; Price, J.; Lyons, C.K. Long-term effects of fragmentation and fragment properties on bird species richness in Hawaiian forests. *Biol. Conserv.* **2010**, *143*, 280–288. [CrossRef]
32. Goetz, S.; Steinberg, D.; Dubayah, R.; Blair, B. Laser remote sensing of canopy habitat heterogeneity as a predictor of bird species richness in an eastern temperate forest, USA. *Remote Sens. Environ.* **2007**, *108*, 254–263. [CrossRef]
33. Jung, K.; Kaiser, S.; Böhm, S.; Nieschulze, J.; Kalko, E.K.V. Moving in three dimensions: Effects of structural complexity on occurrence and activity of insectivorous bats in managed forest stands. *J. Appl. Ecol.* **2012**, *49*, 523–531. [CrossRef]
34. Lesak, A.A.; Radeloff, V.; Hawbaker, T.; Pidgeon, A.M.; Gobakken, T.; Contrucci, K. Follow publication Modeling forest song bird species richness using LiDAR-derived measures of forest structure. *Remote Sens. Environ.* **2011**, *115*, 2823–2835. [CrossRef]
35. Müller, J.; Moning, C.; Bässler, C.; Heurich, M.; Brandl, R. Using airborne laser scanning to model potential abundance and assemblages of forest passerines. *Basic Appl. Ecol.* **2009**, *10*, 671–681. [CrossRef]
36. Müller, J.; Stadler, J.; Brandl, R. Composition versus physiognomy of vegetation as predictors of bird assemblages: The role of LiDAR. *Remote Sens. Environ.* **2010**, *114*, 490–495. [CrossRef]
37. Rechsteiner, C.; Zellweger, F.; Gerber, A.; Breiner, F.T.; Bollmann, K. Remotely sensed forest habitat structures improve regional species conservation. *Remote Sens. Ecol. Conserv.* **2017**, *3*, 247–258. [CrossRef]
38. Vogeler, J.C.; Hudak, A.T.; Vierling, L.A.; Evans, J.; Green, P.; Vierling, K.T. Terrain and vegetation structural influences on local avian species richness in two mixed-conifer forests. *Remote Sens. Environ.* **2014**, *147*, 13–22. [CrossRef]
39. Zellweger, F.; Braunisch, V.; Baltensweiler, A.; Bollmann, K. Remotely sensed forest structural complexity predict multi species occurence at the landscape scale. *Forest Ecol. Manag.* **2013**, *307*, 303–312. [CrossRef]
40. Myers, N.; Mittermeier, R.A.; Mittermeier, C.G.; Da Fonseca, G.A.B.; Kent, J. Biodiversity hotspots for conservation priorities. *Nature* **2000**, *403*, 853–858. [CrossRef]
41. Basset, Y.; Cizek, L.; Cuenoud, P.; Didham, R.; Guilhaumon, F.; Missa, O.; Novotny, V.; Ødegaard, F.; Roslin, T.; Schmidl, J.; et al. Arthropod Diversity in a Tropical Forest. *Science* **2012**, *338*, 1481–1484. [CrossRef]
42. Novotny, V.; Drozd, P.; Miller, S.; Kulfan, M.; Janda, M.; Basset, Y.; Weiblen, G.D. Why are there so many species of herbivorous insects in tropcial rainforests? *Nature* **2006**, *313*, 1115–1118.
43. Peters, M.K.; Hemp, A.; Appelhans, T.; Behler, C.; Classen, A.; Detsch, F.; Ensslin, A.; Ferger, S.W.; Frederiksen, S.B.; Gebert, F.; et al. Predictors of elevational biodiversity gradients change from single taxa to the multi-taxa community level. *Nat. Commun.* **2016**, *7*, 13736. [PubMed]
44. Vehmas, M.; Eerikäinen, K.; Peuhkurinen, J.; Packalén, P.; Maltamo, M. Identification of boreal forest stands with high herbaceous plant diversity using airborne laser scanning. *Forest Ecol. Manag.* **2009**, *257*, 46–53. [CrossRef]
45. Borthagaray, A.I.; Arim, M.; Marquet, P.A. Connecting landscape structure and patterns in body size distributions. *OIKOS* **2012**, *121*, 697–710. [CrossRef]
46. Morse, D.; Lawton, J.; Dodson, M.; Williamson, M. Fractal Dimension of Vegetation and the Distribution of Arthropod Body Length. *Nature* **1985**, *314*, 731–733. [CrossRef]
47. Siemann, E.; Haarstad, J.; Tilman, D. Dynamics of plant and arthropod diversity during old field succession. *Ecography* **1999**, *22*, 406–414. [CrossRef]
48. Stanska, M.; Stanski, T. Body size distribution of spider species in various forest habitats. *Pol. J. Ecol.* **2017**, *65*, 359–370. [CrossRef]
49. Kaspari, M.; Weiser, M.D. The size–grain hypothesis and interspecific scaling in ants. *Funct. Ecol.* **1999**, *13*, 530–538. [CrossRef]
50. Sarty, M.; Abbott, K.L.; Lester, P.J. Habitat complexity facilitates coexistence in a tropical ant community. *Oecologia* **2006**, *149*, 465–473. [CrossRef]
51. Google Maps. Mt Kilimanjaro, Maxar Technologies (2020), CNES/Airbus (2020). Available online: https://www.google.de/maps/search/kilimanjaro/@-3.094907,37.282672,49189m/data=!3m1!1e3 (accessed on 22 October 2021).
52. Peters, M.; Andreas, H.; Appelhans, T.; Becker, J.; Behler, C.; Classen, A.; Detsch, F.; Ensslin, A.; Ferger, S.; Frederiksen, S.B.; et al. Climate–land-use interactions shape tropical mountain biodiversity and ecosystem functions. *Nature* **2019**, *568*, 88–92. [CrossRef]
53. Isenburg, M. LAStools—Efficient LiDAR Processing Software. Available online: http://rapidlasso.com/LAStools (accessed on 22 October 2021).
54. Woellauer, S.; Zeuss, D.; Magdon, P.; Nauss, T. RSDB: An easy to deploy open-source web platform for remote sensing raster and point cloud data management, exploration and processing. *Ecography* **2020**, *44*, 414–426. [CrossRef]
55. Jenness, J. Calculating Landscape Surface Area from Digital Elevation Models. *Wildl. Soc. Bull.* **2004**, *32*, 829–839. [CrossRef]
56. Getzin, S.; Fischer, R.; Knapp, N.; Huth, A. Using airborne LiDAR to assess spatial heterogeneity in forest structure on Mount Kilimanjaro. *Landsc. Ecol.* **2017**, *32*, 1881–1894. [CrossRef]

57. Kuhn, M. *Caret: Classification and Regression Training*; R package version 6.0-80; 2018. Available online: https://cran.r-project.org/web/packages/caret/index.html (accessed on 22 October 2021).

58. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2018.

59. Carrascal, L.M.; Galván, I.; Gordo, O. Partial least squares regression as an alternative to current regression methods used in ecology. *Oikos* **2009**, *118*, 681–690. [CrossRef]

60. Laurin, G.V.; Chen, Q.; Lindsell, J.A.; Coomes, D.A.; Frate, F.D.; Guerriero, L.; Pirotti, F.; Valentini, R. Above ground biomass estimation in an African tropical forest with lidar and hyperspectral data. *ISPRS J. Photogramm. Remote Sens.* **2014**, *89*, 49–58. [CrossRef]

61. Luo, S.; Wang, C.; Xi, X.; Pan, F.; Peng, D.; Zou, J.; Nie, S.; Qin, H. Fusion of airborne LiDAR data and hyperspectral imagery for aboveground and belowground forest biomass estimation. *Ecol. Indic.* **2017**, *73*, 378–387. [CrossRef]

62. Næsset, E.; Bollandsås, O.M.; Gobakken, T. Comparing regression methods in estimation of biophysical properties of forest stands from two different inventories using laser scanner data. *Remote Sens. Environ.* **2005**, *94*, 541–553. [CrossRef]

63. Meyer, H. *CAST: 'Caret' Applications for Spatial-Temporal Models*; R package version 0.3.1; 2018. Available online: https://cran.r-project.org/web/packages/CAST/index.html (accessed on 22 October 2021).

64. Meyer, H.; Reudenbach, C.; Hengl, T.; Katurji, M.; Nauss, T. Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. *Environ. Model. Softw.* **2018**, *101*, 1–9. [CrossRef]

65. Acebes, P.; Lillo, P.; Jaime-Gonzalez, C. Disentangling LiDAR Contribution in Modelling Species-Habitat Structure Relationships in Terrestrial Ecosystems Worldwide. A Systematic Review and Future Directions. *Remote Sens.* **2021**, *13*, 3447. [CrossRef]

66. Smart, L.S.; Swenson, J.J.; Christensen, N.L.; Sexton, J.O. Three-dimensional characterization of pine forest type and red-cockaded woodpecker habitat by small-footprint, discrete-return lidar. *Forest Ecol. Manag.* **2012**, *281*, 100–110. [CrossRef]

67. Vogeler, J.C.; Hudak, A.T.; Vierling, L.A.; Vierling, K.T. Lidar-derived Canopy Architecture Predicts Brown Creeper Occupancy of Two Western Coniferous Forests. *Condor* **2013**, *115*, 614–622. [CrossRef]

68. Davies, A.B.; Asner, G.P. Advances in animal ecology from 3D-LiDAR ecosystem mapping. *Trends Ecol. Evol.* **2014**, *29*, 681–691. [CrossRef] [PubMed]

69. Vierling, K.T.; Bässler, C.; Brandl, R.; Vierling, L.A.; Weiss, I.; Müler, J. Spinning a laser web: Predicting spider distributions using LiDAR. *Ecol. Appl.* **2011**, *21*, 577–588. [CrossRef] [PubMed]

70. Zellweger, F.; Roth, T.; Bugmann, H.; Bollmann, K. Beta diversity of plants, birds and butterflies is closely associated with climate and habitat structure. *Glob. Ecol. Biogeogr.* **2017**, *26*, 898–906. [CrossRef]

71. McCain, C.M.; Grytnes, J.A. Elevational Gradients in Species Richness. In *eLS*; American Cancer Society: Atlanta, GA, USA, 2010. [CrossRef]

72. Müller, J.; Bae, S.; Röder, J.; Chao, A.; Didham, R.K. Airborne LiDAR reveals context dependence in the effects of canopy architecture on arthropod diversity. *Forest Ecol. Manag.* **2014**, *312*, 129–137. [CrossRef]

73. Schooler, S.L.; Zald, H.S.J. Lidar Prediction of Small Mammal Diversity in Wisconsin, USA. *Remote Sens.* **2019**, *11*, 2222. [CrossRef]