

An efficient method to exploit LiDAR data in animal ecology

Simone Ciuti^{1,2}  | Henriette Tripke¹ | Peter Antkowiak¹ | Ramiro Silveyra Gonzalez¹ | Carsten F. Dormann¹ | Marco Heurich^{3,4}

¹Department of Biometry and Environmental System Analysis, University of Freiburg, Tennenbacher Straße 4, Freiburg, Germany

²School of Biology and Environmental Science, University College Dublin, Belfield, Ireland

³Department of Conservation and Research, Bavarian Forest National Park, Freyunger Straße 2, Grafenau, Germany

⁴Chair of Wildlife Ecology and Wildlife Management, University of Freiburg, Tennenbacher Straße 4, Freiburg, Germany

Correspondence

Simone Ciuti
Email: simone.ciuti@ucd.ie

Handling Editor: David Warton

Abstract

1. Light detection and ranging (LiDAR) technology provides ecologists with high-resolution data on three-dimensional vegetation structure. Large LiDAR datasets challenge predictive ecologists, who commonly simplify point clouds into structural attributes (namely LiDAR-based metrics such as canopy height), which are used as predictors in ecological models, potentially with loss of relevant information.
2. We illustrate an efficient alternative approach to reduce the dimensionality of LiDAR data that aims at minimal data filtering with no *a priori* assumptions on the ecology of the target species. We first fit the ecological model exploiting the full variability in the LiDAR point cloud, then we explain the results using post-modelling LiDAR-data classification for ecological interpretation only. This is the classical logic of explorative, hypothesis generating and predictive statistics, rather than testing specific vegetation-structural hypotheses.
3. First, we reduce the dimensionality of the LiDAR point cloud by principal component analysis (PCA) to fewer predictors. Second, we show that LiDAR-PCs are capable to outperforming commonly used environmental predictors in ecological modelling, including LiDAR-based metrics. We exemplify this by modelling red deer (*Cervus elaphus*) and roe deer (*Capreolus capreolus*) resource selection in the Bavarian Forest National Park, Germany. After fitting the ecological model, we provide an interpretation of the information included in LiDAR-PCs, which allows users to draw conclusions whenever using them as predictors. We make use of the PCA rotation matrix and post-modelling data classification, and document deer selection for understorey vegetation at unprecedented fine scale.
4. Our approach is the first attempt in animal ecology to avoid the use of LiDAR-based metrics as model predictors, but rather generate principal components able to capture most of the LiDAR point cloud variability. Our study demonstrates that LiDAR-PCs can boost ecological models. We envision a potential use of LiDAR-PCs in several applications, particularly species distribution and habitat suitability models. We demonstrate an application of our approach by building suitability maps for both deer species, which can be used by practitioners to visualize model spatial predictions and understand the type of forest structures selected by deer.

KEYWORDS

habitat modelling, LiDAR, principal component analysis, red deer, resource selection functions, roe deer, satellite telemetry

1 | INTRODUCTION

Light detection and ranging (LiDAR) is an active remote sensing method that measures distances to the earth surface by sending light from an aircraft in the form of laser pulses, which are recorded as backscattered light after hitting on any surface structure. Three-dimensional spatial coordinates are calculated for each pulse at the location where it hits a surface, thus generating a “point cloud.” LiDAR was designed to retrieve information useful for planning dams, road construction, steering cruise missiles, forest inventories (Lim, Treitz, Wulder, St-Onge, & Flood, 2003), and, more recently, LiDAR has been used to gather three-dimensional vegetation structure for ecologists (Neumann et al., 2015; Vierling, Vierling, Gould, Martinuzzi, & Clawges, 2008).

Light detection and ranging enables precise measurement of the three-dimensional structure of ecosystems across spatial scales from tree branches to entire landscapes (Vierling et al., 2008). A plethora of studies on birds, bats, nonflying mammals and invertebrates have shown three-dimensional structure to be the key driver of animal ecology and diversity (Davies & Asner, 2014; Simonson, Allen, & Coomes, 2014), including species distributions, behaviour and resource selection (Davies, Tambling, Kerley, & Asner, 2016; Flaspohler et al., 2010; Vierling et al., 2008).

Light detection and ranging point clouds are a conceptual and technical challenge to applied ecologists. To reduce file size and number of dimensions, the most common approach used so far is to compute LiDAR-based metrics and then use those as predictors in an ecological model. In other words, LiDAR datasets can be interrogated at different grain sizes and produce different metrics for different ecological questions (Davies & Asner, 2014; Simonson et al., 2014). Flying animals clearly move and live in three-dimensional space, for instance and LiDAR-derived metrics that supposedly affect movement in flight have to be selected accordingly (Davies & Asner, 2014). LiDAR-based metrics can be extracted from the point cloud at different grid cell size, i.e. area-based LiDAR metrics, and for individual trees following crown segmentation, i.e. individual tree recognition (Simonson et al., 2014).

Light detection and ranging-based metrics such as understorey density, canopy vertical distribution, canopy height and cover have been shown to be related to animal ecology across taxonomic groups (see table 1 in Davies & Asner, 2014; see table 1 in Simonson et al., 2014). Classification of LiDAR returns in height percentiles, fractional cover or forest density classes are an example of LiDAR-derived metrics deployed in large herbivorous ecology (Lone, Loe, et al., 2014; Lone, van Beest, et al., 2014; Melin et al., 2014; Nijland, Nielsen, Coops, Wulder, & Stenhouse, 2014). Depending on the ecological question, researchers may decide to limit the use of LiDAR data describing vegetation within a certain height threshold, for example two metres above the ground for large herbivores, which is directly related to the feeding ecology of the target species (Ewald, Dupke, Heurich, Müller, & Reineking, 2014; Lone, van Beest, et al., 2014). This approach, however, subsets LiDAR data based on what the researcher thinks an animal is looking for, rather than allowing the data to suggest a model based on where the animals have been recorded.

We thus illustrate an alternative “fit first, explain later”-philosophy for avoiding loss of information with high-dimensional predictors. This approach implies a limited pre-processing of LiDAR point cloud data (i.e. dimensionality reduction with insignificant loss of variability), which is necessary to make the data usable in ecological applications, but avoids data filtering. We first fit the ecological model exploiting the full variability in the LiDAR point cloud data, and then we explain and interpret the results using post-modelling LiDAR data classification only for ecological interpretation. This is the classical logic of explorative, hypothesis generating and predictive statistics, rather than testing specific vegetation-structural hypotheses.

We describe our approach to handle LiDAR data in three main steps.

First, we show that the principal component analysis (PCA) is a simple but effective method to condense most of the variability in LiDAR point cloud into few predictors (1st goal).

Second, we show that principal component axes derived from LiDAR data are capable to outperforming commonly used environmental and habitat predictors (e.g. habitat maps, LiDAR-based metrics) in ecological predictions (2nd goal). We exemplify this by analysing red deer (*Cervus elaphus*) and roe deer (*Capreolus capreolus*) resource selection in the Bavarian Forest National Park (BFNP), Germany, although this approach similarly applies to other ecological applications. The use of PCA is not new in the remote sensing community, where it is routinely used to reduce the data dimensionality or extract information on individual tree crown when using LiDAR point clouds and/or hyperspectral data (Fayad et al., 2014; La, Eo, Chang, & Kim, 2015; Onojeghwo & Blackburn, 2011). Our work, however, is the first attempt in animal ecology to avoid the use of LiDAR-based metrics, but rather generate principal components as predictors able to capture most of the LiDAR point cloud variability and boost the performance of ecological models.

Third, after fitting the ecological model, we discuss our results by providing an interpretation of the information included in principal axes derived from LiDAR data (3rd goal), which allows users (from theoretical ecologists to practitioners) to make inferences and draw final conclusions, whenever using LiDAR principal components as predictors.

2 | MATERIALS AND METHODS

2.1 | Using PCA to capture LiDAR data variability (1st goal)

We collected full waveform LiDAR data for the BFNP (240 km²). The aerial survey was conducted under leaf-on conditions by a contractor (Milan-Flug-GmbH) on 24–26.07.2012. LiDAR data were recorded by a Riegl-Q680i (operating at a pulse frequency of 350 KHz) laser scanner mounted on the aircraft. Specific details on LiDAR survey: height above the ground: 650 m; flight speed: 36 m/s; overlap: 50%; laser wavelength: 1,550 nm; scan frequency: 37.500/s; pulse frequency: 300.000/s; FOV: ±30°; beam dispersion: <0.5 mrad.

During pre-processing, we applied a decomposition of the LiDAR waveforms (i.e. discretized and then discarded) using Gaussian functions (Wagner, Ullrich, Ducic, Melzer, & Studnicka, 2006). The number of beams was $15/\text{m}^2$, whereas the number of recorded reflections after waveform decomposition was $40/\text{m}^2$. We converted the raw data to spd format with the open-source Sorted Pulse Data library, SPDLib (Bunting, Armston, Clewley, & Lucas, 2013; Bunting, Armston, Lucas, & Clewley, 2013). Ground returns were first classified with the progressive morphology filter (Zhang et al., 2003), and then by the multi-scale curvature algorithm (Evans & Hudak, 2007), both implemented in the SPDLib. We thus used a natural neighbour algorithm to interpolate ground elevation values for each return. We obtained a normalized three-dimensional LiDAR point cloud with the position of all returns relative to ground level by subtracting ground elevation from the return coordinates.

We subdivided the LiDAR point cloud at the finest resolution determined by the accuracy of the laser sensor: vertical subdivisions were made every 0.5 m above ground, whereas horizontal resolution was initially set to 1 m^2 . We thus created voxels (volume elements) of $1 \times 1 \times 0.5\text{ m}^3$. We counted the proportion of returns in each voxel (i.e. the number of returns as a proportion of the total number in the entire vertical column to which the voxel belongs) and stored all voxel counts with the same height above ground in a raster layer. Consequently, one hundred layers were created representing all the height levels from ground to the maximum tree height of 50 m in the surveyed area. The lowest layer with returns below 0.5 m was discarded due to the inability in distinguishing between vegetation and ground returns (Heurich, Fischer, Knoerzer, & Krzystek, 2008). The rasterized dataset thus comprised 99 layers of $33,000 \times 31,000$ grid cells, each layer including same height voxels and each grid cell representing the proportion of returns in one voxel.

We did not take arbitrary decisions in defining the vertical resolution (i.e. voxel height), because it was determined by the laser-sensor accuracy. Selecting a coarser vertical resolution would have speeded up our computations, which we avoided, however, in order to keep the entire variability in the point cloud data. A different approach is needed for the horizontal resolution, which is the size of the grid cell that can affect the way we represent the forest structures. On the one hand, under dense canopy cover, our starting 1-m^2 resolution may be a rather small grain size, resulting in few laser returns from the understorey vegetation, leading to several empty voxels and sparse matrices. On the other hand, setting a larger horizontal resolution would correspond to a generalization of details and imply a loss of fine-scale information. Thus, the horizontal resolution has to be defined via sensitivity analysis by inspecting increasingly coarser grid cell size (details below).

Return proportions in vertically adjacent voxels are highly collinear. We thus ran a PCA to generate uncorrelated principal components (PCs) able to harness the variability in the 99 layers and represent the vertical vegetation structure. This analysis is achieved by performing an eigenvalue decomposition on the correlation matrix of the data. Given that X is the original matrix of data points, the transformed data matrix X' is generated as $X' = \lambda X$, with λ being the matrix of the eigenvalues

of the correlation matrix or rotation matrix. In practice, we performed a singular value decomposition of the centred and scaled data matrix, which leads to a higher numerical accuracy (Mardia, Kent, & Bibby, 1980).

We repeated this procedure at different horizontal resolutions by aggregating laser returns from the original $1 \times 1\text{ m}^2$ raster to the new horizontal extents. The optimal horizontal resolution is here defined as the one that can explain the greatest amount of LiDAR variability with fewest PCs and, at the same time, preserve fine detail by keeping grid cell size low. This can be achieved via sensitivity analysis by screening the curves of cumulative variance explained by PCAs carried out on different horizontal aggregation levels. Note that this approach is expected to define different optimal horizontal resolutions depending on the vegetation complexity of the surveyed areas as well as on the pulse density of the LiDAR dataset, and hence this sensitivity analysis is an intrinsic part of the overall analysis.

Once the spatial resolution is defined, the corresponding PCA produces as many axes as the number of dimensions in the dataset (in our case $n = 99$). We suggest reducing this number while retaining enough PCs to embrace most of LiDAR data variability. We selected all PCs that explain a proportion of variance equal or higher than $1/n$ (Jolliffe, 2002). We will refer to them as LiDAR-PCs hereafter. We visualized key steps described so far in Figure 1. The handling of large LiDAR datasets may be challenging and requires computer memory and power, which may not be accessible to all ecologists and practitioners. We provided as Appendix S1 an overview of shortcuts that may be necessary in case of limited computer power and full details on the sensitivity analysis for the horizontal resolution.

2.2 | Fit first, explain later: time to fit (2nd goal)

We compared the performance of LiDAR-PCs with alternative environmental covariates using deer satellite telemetry data collected in the BFNP (study site details in Appendix S2) and modelling deer habitat selection.

2.2.1 | Roe and red deer telemetry data

We obtained telemetry data for female red and roe deer (Ciuti et al., 2017) in the BFNP (details in Appendix S2). Both deer species were fit with GPS-GSM collars (VECTRONIC Aerospace, Germany). Accuracy field tests carried out with static GPS collars revealed a high overall position acquisition rate (96.7%) and accuracy (10 m; Stache, Löttker, & Heurich, 2012). We used deer telemetry data collected in spring-summer 2010–2012, when the conditions were similar to those recorded by LiDAR (summer 2012). Our four datasets included relocations at 15-min (11,956 locations, $n = 10$ females) and 1-h (5,290 locations, $n = 10$ females) sampling rate for red deer, and 1-h (6,678 locations, $n = 17$ females) and 4-h (1,415 locations, $n = 15$ females) sampling rate for roe deer respectively. Fine-scale resource selection patterns can be affected by sampling rate (Boyce, 2006; Thurfjell, Ciuti, & Boyce, 2014); we thus ran separate analyses on these four datasets.

2.2.2 | Building deer resource selection functions

A resource selection function (RSF) is defined as any statistical model designed to estimate the relative probability of an animal selecting a resource unit relative to alternative possible available resource units (Manly, McDonald, Thomas, McDonald, & Erickson, 2002). To avoid misconception, selection in RSFs is clearly based on used and available resource units, and not on used and unused ones. For every GPS-location where deer were relocated, we associated 50 random locations to represent alternatively available resource units (details on why and how we selected 50 random locations are fully reported in Appendix S2). Random available locations were associated to each used location in a way that available locations were actually reachable by the deer within the time interval defined by the fix rate.

We first fit GLMMs with binomial response variable (1: used; 0: available) to estimate model parameters using the package “LME4” in R (Bates, Mächler, Bolker, & Walker, 2015). The stratum—i.e. each deer-used location and its 50 random available locations—was nested within the deer identity as random intercept in the mixed model. We then calculated RSFs based on model parameters β (excluding the intercept) estimated by the GLMMs in the following equation: $w(x) = \exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)$, where $w(x)$ is the relative probability of selection and x is a covariate vector of length n (Manly et al., 2002). We compared (using Akaike’s information criterion) alternative resource selection model structures containing either LiDAR-PCs or other environmental variables as predictors.

2.2.3 | Fitting alternative resource selection model structures

All models contained a set of default predictors taking the form $sun_elev + sun_elev^2 + rugged + rugged^2 + dist_trails + dist_trails^2 + solrad + sun_elev \times dist_trails + sun_elev \times solrad$.

Sun_elev (elevation of the sun with respect to the horizon) is a proxy for time of the day. *Rugged* is the terrain ruggedness derived from the digital elevation model and is a proxy for terrain accessibility and human disturbance (e.g. steeper terrains less accessible to humans). *Dist_trails* represents the distance to the closest hiking trail, again a proxy for human disturbance, whereas *solrad* is a categorical proxy for microclimate conditions (see Table S1 for additional details). These predictors allowed taking into account variation in selection patterns as a function of time of the day, human disturbance and microclimate conditions respectively. Quadratic terms for continuous predictors were included to allow for nonlinear relationships in selection patterns. The interaction $sun_elev \times dist_trails$ was included to allow variation in the selection for roads as a function of the time of the day (e.g. assuming less humans on roads at night), whereas $sun_elev \times solrad$ was included to depict the variation in selection of sites with different microclimatic conditions as a function of the time of the day and related temperature. The models differed in the environmental predictors unique to each model (described below and in Table S1). When the environmental predictor was categorical, then the model also included a two-way interaction with *sun_elev* to depict variation in selection for the environmental predictor as a function of time of the day. Likewise, when the predictor

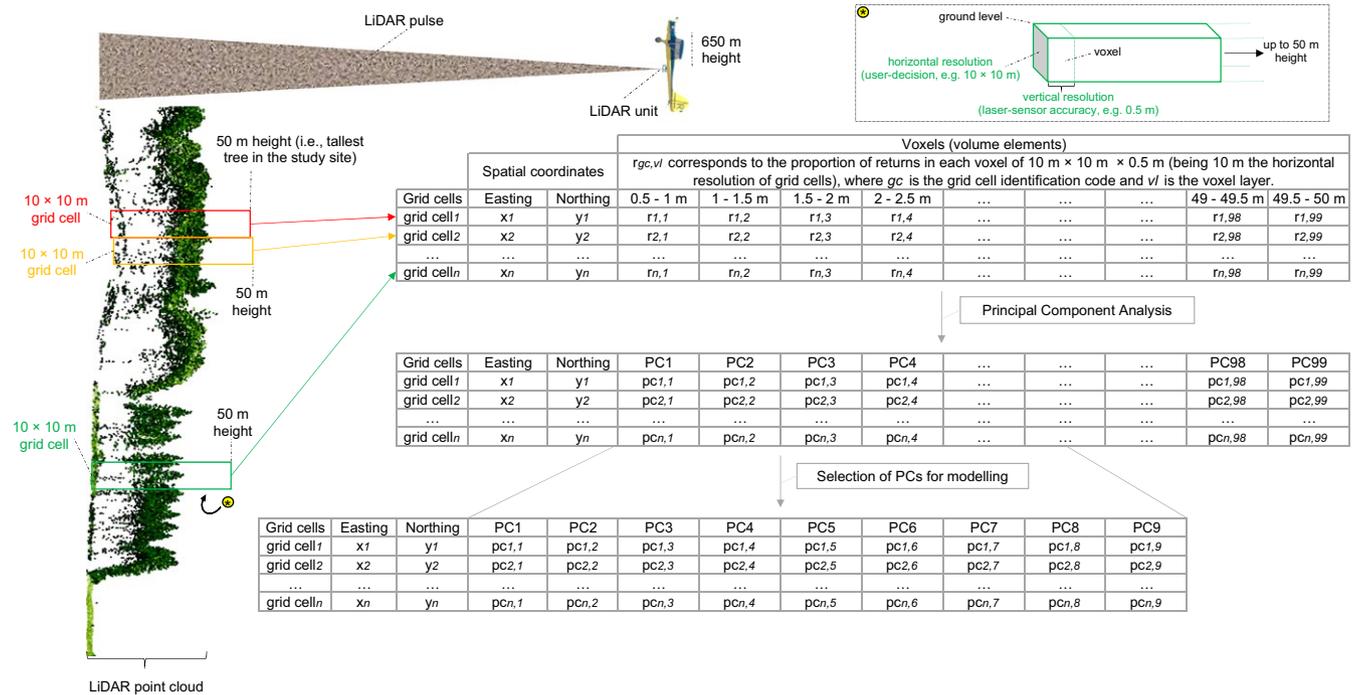


FIGURE 1 Main steps involved in the reduction in dimensionality of light detection and ranging (LiDAR) point cloud data into fewer principal components (PCs). In this example with LiDAR data from the Bavarian Forest National Park, Germany, horizontal resolution was set to 10 m, leading to the selection of the first nine PCs responsible for carrying more than 90% of the variability in the point cloud

was continuous, we included a quadratic term to account for nonlinear effects as well as a two-way interaction with *sun_elev*.

We fit seven alternative GLMMs. Model 1 included only the default predictors. Models 2–7 included alternative environmental predictors which were satellite-, aerial image- or LiDAR-derived covariates (Table S1) respectively. Model 2 included a photo-interpreted habitat map of the BFNP from year 2012 (*map2012*). This map was commissioned by the BFNP to a contractor (Geoplana), with a resolution below 5 m. Model 3 included the NDVI (*ndvi*) as the main environmental predictor. We derived the monthly NDVI from Landsat-bands 5 (TM) and 7 (ETM+). We downloaded one satellite image for each month across the monitoring period (spring-summer 2010–2012) with cloud cover lower than 10%. We computed NDVI from band 3 (red) and 4 (near infrared). Models 4 and 5 included the classical LiDAR-based metrics used in large herbivorous ecology. Model 4 included as the main predictor the mean height of vegetation (*MeanHeight*) from all returns of the point cloud within a grid cell (Ackers, Davis, Olsen, & Dugger, 2015; Farrell et al., 2013). Model 5 included the fractional vegetation cover of three different height strata as predictors. Following the approach of Ewald et al. (2014), we calculated the understory (*UStory*), midstory (*MStory*) and overstory vegetation (*OStory*). Each predictor described the proportion of returns within their height strata and captured information of the vegetation density (Lone, Loe, et al., 2014; Lone, van Beest, et al., 2014). Finally, we built the alternative models 6 and 7 using LiDAR-PCs. On the basis of the results of the sensitivity analysis on LiDAR horizontal resolution, we decided on using two different resolutions and created a model for each of them (see Section 3). Thus, we incorporated a different number of predictors (number of PCs) in model 9 and 10.

Predictors included in each model structure were not collinear (Pearson $|r_p| \ll .7$, sensu (Zuur, Ieno, & Smith, 2007)). Specifically, default predictors used as standard inputs to all GLMMs were practically uncorrelated (Pearson $|r_p| < .15$).

2.3 | Fit first, explain later: time to explain (3rd goal)

The use of LiDAR-PCs poses some challenges in result interpretation and accessibility to practitioners and applied scientists, and we need to provide a way to understand their meaning. We thus provided the following items that we use in the discussion for interpretation:

1. Rotation matrix plots showing how LiDAR return height classes (thus, vegetation structure) load on each principal component;
2. reclassification of PCs' values depending on the vegetation classes defined by the most accurate vegetation map available for the area, meaning that the reader can translate the deer-selected PCs into selection for habitat classes meaningful to humans;
3. vegetation cross sections ($300 \times 10 \text{ m}^2$) to better visualize the relationship between LiDAR-PC-values, vegetation type and LiDAR point cloud structure. This step also demonstrates how LiDAR-PCs are able to capture fine-scale details in vegetation structure usually not detectable (thus lost) in high-resolution vegetation maps;
4. examples of suitability maps (based on the top-ranking models) for both deer species, which can be used by practitioners to visualize

model spatial predictions, understand the relationship between PCs and vertical forest structure, and manage their landscapes accordingly.

Data handling and analyses were carried out using R 3.3.2 (R Core Team, 2016).

3 | RESULTS

3.1 | Using PCA to capture LiDAR data variability (1st goal)

The sensitivity analysis for the optimal horizontal resolution (Appendix S1) showed that grid cell sizes lower than 25 m^2 ($5 \text{ m} \times 5 \text{ m}$) did not allow the PCA to effectively reduce the complexity of the data (e.g. more than 40 principal components required to explain 90%

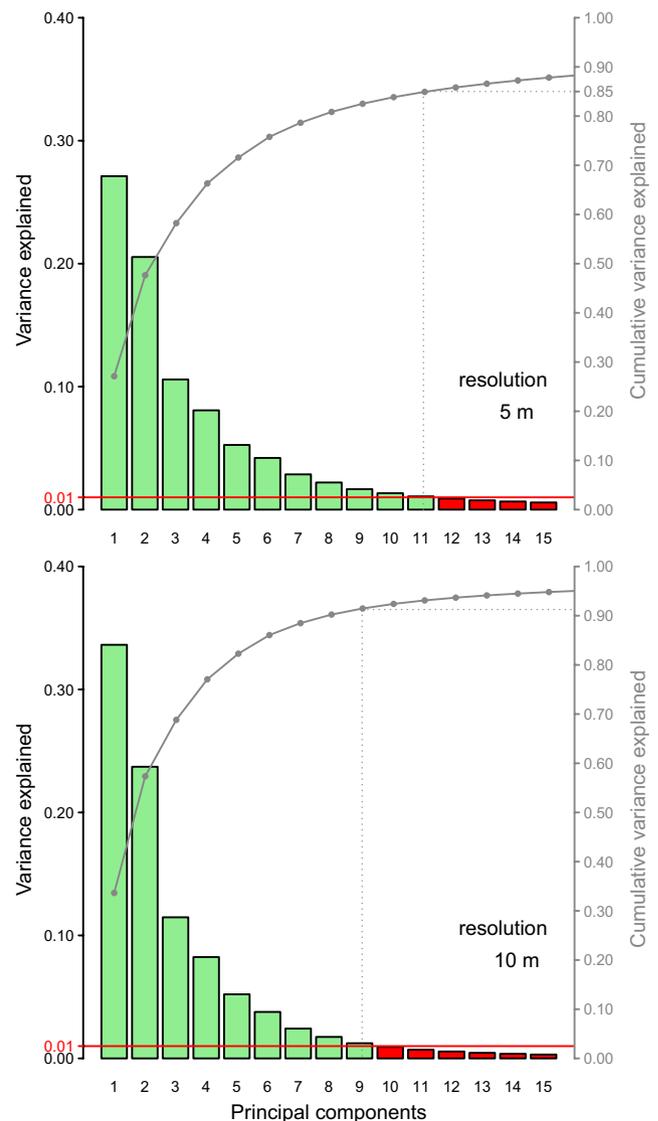


FIGURE 2 Variance and cumulative variance explained in light detection and ranging point cloud data by the first 15 principal components (PCs) as predicted by principal component analysis. PCs able to explain more than 1% of variance (histograms above the 0.01 threshold) were deployed as predictors in deer resource selection models

TABLE 1 Comparison of generalized linear mixed-effect models (different deer species and fix rates) explaining the variability in resource selection based on the Akaike's information criterion

No.	Model	Estimated parameters	Red deer		Red deer		Roe deer		Roe deer	
			15-min	Rank	1-h	Rank	1-h	Rank	4-h	Rank
1	No vegetation	12	354.8	7	507.1	7	792.6	7	396.0	7
3	NDVI ^a	15	333.9	6	454.0	6	752.2	6	367.0	6
4	Mean vegetation height ^b	15	287.9	5	393.8	5	300.1	5	128.1	5
2	Habitat classification map BFNP ^c	42, 42, 42, 38 ^d	229.4	4	231.2	4	169.9	3	55.1	3
5	Fractional vegetation cover ^b	21	1.5	2	45.1	2	179.9	4	80.0	4
6	LiDAR-PCs (5 m horizontal resolution) ^b	45	40.8	3	74.4	3	71.8	2	33.0	2
7	LiDAR-PCs (10 m horizontal resolution) ^b	39	0	1	0	1	0	1	0	1

Model 1 includes default predictors but no environmental ones. Models 2–7 include alternative environmental predictors, which were ^asatellite-, ^bLiDAR- or ^caerial image-derived covariates. See Table S1 for full details on predictors.

^dEstimated parameters vary depending on the levels of the categorical predictor.

of variability). We stopped our sensitivity analysis at the resolution 10 m × 10 m (100 m²), when we could capture most of the variability in LiDAR data with an affordably low number of principal components. This was our main PCA grid cell size (Figure 2; Figure S1). We also picked the arbitrary horizontal resolution 5 m × 5 m (25 m²) as a comparison (Figure 2; Figure S1). We selected all PCs explaining more than 1.0% of the original variance as predictors in ecological modelling. This approach yielded 11 PCs for the 5 m resolution (capable of capturing 85% of the total variability included in LiDAR data), and nine PCs for the 10 m resolution (91% of cumulative variance explained; Figure 2).

3.2 | Fit first, explain later: time to fit (2nd goal)

Models including LiDAR-PCs (10 m resolution) outperformed any other model including alternative environmental covariates in predicting habitat selection in deer (Table 1). In roe deer (both fix rates) and red deer (1-h fix rate only), 10-m-resolution LiDAR-PCs improved model fit by least 33.0 points in AIC compared to other LiDAR-derived metrics, and even greater differences in alternative models. In red deer only (15-min fix rate), LiDAR-derived fractional vegetation cover had predictive performances comparable to 10-m-resolution LiDAR-PCs (Table 1). Five-metre resolution LiDAR-PCs had weaker predictive ability than 10-m-resolution LiDAR-PCs (Table 1). We reported in Table S2 additional alternative models with further environmental predictors not described here, which again were outperformed by 10-m-resolution LiDAR-PCs.

Parameter estimates for the four best models selected in Table 1 are reported in Table S3. Both deer species selection was driven by terrain ruggedness, distance to trails and vegetation structure represented by LiDAR-PC1-PC5 as a function of time of the day, with minor deviation from the general pattern in few cases depending on species and/or fix rate.

With respect to default predictors, roe and especially red deer selected for more gentle terrain (Figure S2), whereas both species were less likely to be relocated closer to trails during the day (Figure S2; Table S3). According to parameters estimated for LiDAR-PCs, both deer species strongly selected for the environmental characteristics described by the first five PCs (Table S3), which alone carry more than 80% of the entire variability in LiDAR data (Figure 2). Selection patterns for PC1-5 by deer are depicted in Figure 3 (1-h fix rates) and Figure S3 (red deer 15-min and roe deer 4-h fix rates). Based on the relationship between vegetation height and PC loading depicted by rotation matrixes (Figure 3), the first five PCs were able to capture the understorey vegetation occurring in different vegetation context, i.e. PC1: negative loadings, within open areas; PC2: positive loadings, within 25–30 m forest height; PC3-4: positive loadings, within 15–25 m forest height; PC5, negative loadings, within 15–20 and 30–35 forest height. Remarkably, both deer species strongly selected for those sites containing understorey vegetation (Figure 3, Figure S3), with variation depending on the time of the day and the forest structure, for example understorey vegetation within open areas depicted by PC1 mostly selected at twilight and night vs. understorey vegetation within forest patches depicted by PC5 mostly selected during the day (Figure 3; Figure S3).

4 | DISCUSSION

We provided an empirical illustration of how LiDAR data performance as environmental predictors can be greatly improved (2nd goal) by embracing their variability with simple PCAs (1st goal). In contrast to previous research (reviewed in Davies & Asner, 2014; Simonson et al., 2014), we did not filter LiDAR data by creating

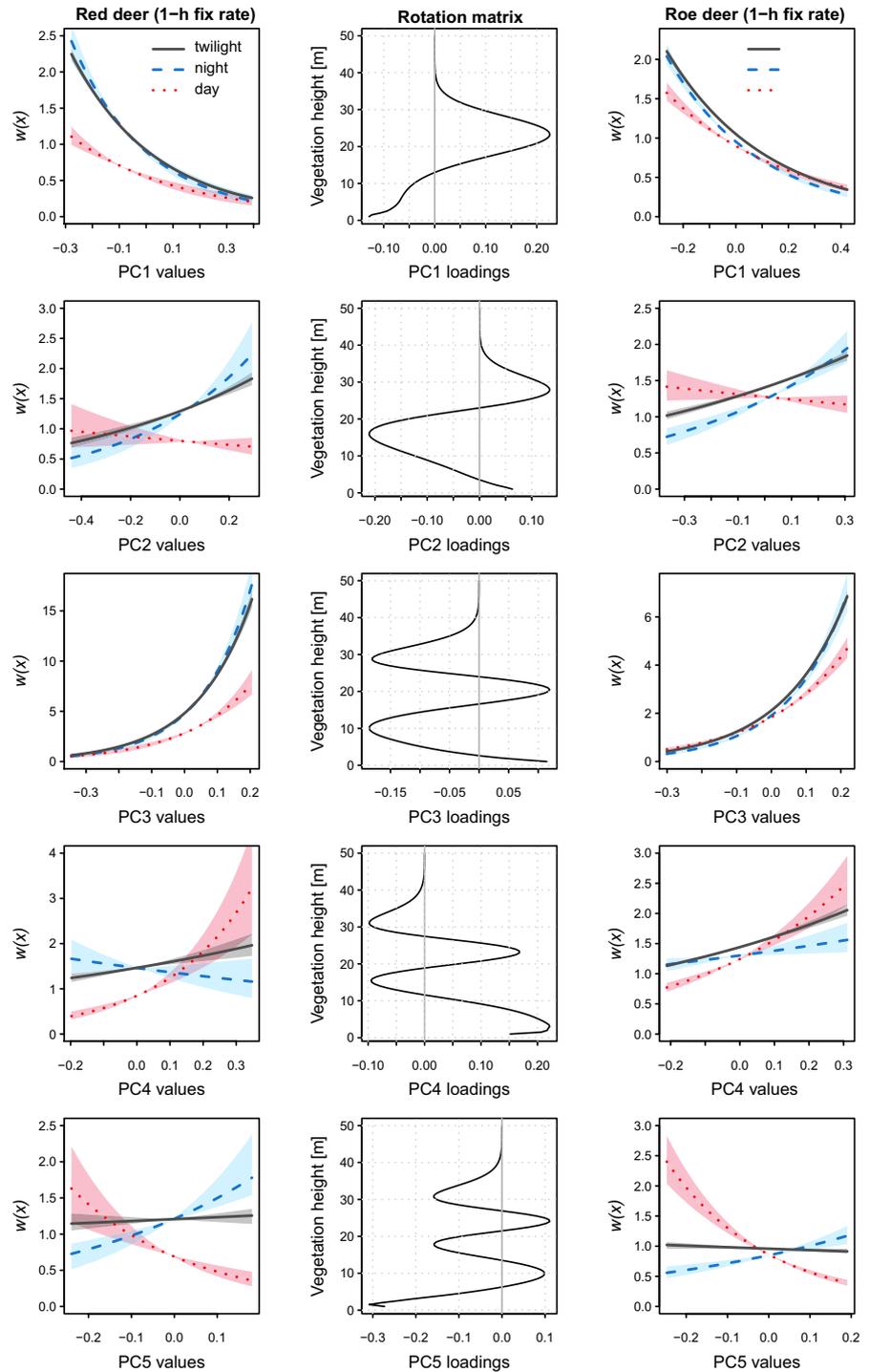


FIGURE 3 Relative probability of selection $w(x)$ for vegetation characteristics (light detection and ranging point cloud data, 10 m resolution) described as principal axes (PC1–PC5) depending on the time of the day by red deer (left column) and roe deer (right column) as predicted by resource selection functions (RSFs) for the 1-hourly fix rates. Rotation matrices (central column) describe the loading of vegetation height (from laser returns, in metres) on each principal component axis. Light conditions are based on sun elevation, simplified to twilight, night and day. RSFs were built using parameters estimated by top ranked generalized linear mixed-effect model selected in Table 1. Shaded areas are 95% conditional CIs

metrics based on *a priori* assumptions linked to the ecological problem under scrutiny. Previous research in large herbivores showed how LiDAR-based metrics may have comparable predictive performance compared to non-LiDAR environmental predictors (Lone, Loe, et al., 2014; Lone, van Beest, et al., 2014), whereas our LiDAR-PCs significantly outperformed all other environmental predictors.

If the aim of a study is to predict animal resource selection or distribution, then our LiDAR-derived PCs are suitable candidate predictors able to nicely embrace environmental variability which, combined with

other covariates, can guarantee the best predicting performances, no matter what the ecological interpretation of LiDAR-PCs is. We thus envision a potential use of LiDAR-PCs in species distribution models (Elith & Leathwick, 2009; Warren & Seifert, 2011). However, if the aim of a study is to explain observed ecological patterns using LiDAR-PCs, then it is time to provide the interpretation of the items that can be extracted from our analyses. These items can be used by behavioural ecologists to undertake experiments with respect to observed behavioural patterns, or by applied ecologists and practitioners to take actions in the landscapes they manage.

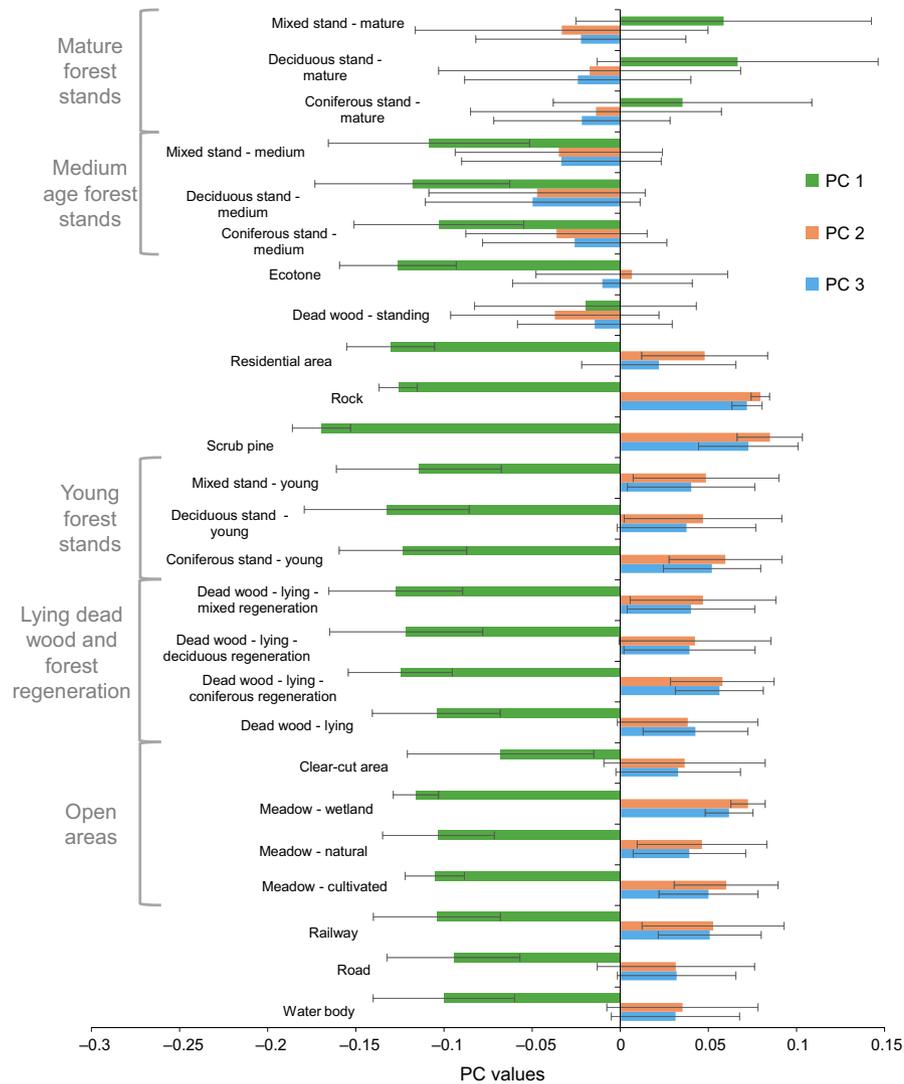


FIGURE 4 Principal component (PC) values ($M \pm SD$) for each level of the categorical land-cover map built from the interpretation of aerial images of the Bavarian Forest National Park, Germany. The figure depicts the first three PCs, explaining two-thirds of the overall variability in light detection and ranging point cloud data (10 m horizontal resolution)

4.1 | Fit first, explain later: time to explain (3rd goal)

Rotation matrix plots reported in Figure 3 depict the relationship between vegetation height and PC loadings. The understory vegetation captured by negative loadings of PC1 within open areas, for instance differs to that represented within forest areas by positive and negative loadings of PC2-4 and PC5 respectively. The result is that our RSF models could depict fine-scale deer selection for such different understory vegetation depending on the LiDAR returns recorded along the entire vertical structure of the forest. Whereas deer selected for understory vegetation within open areas mostly at night and twilight (Figure 3, PC1), they were more likely to use understory vegetation within forest during the day (Figure 3, PC5), thus balancing the need to access to forage and reduce the risk of being spotted, for example by humans (Dupke et al., 2016). Because we avoided subsetting LiDAR returns occurring below an arbitrary threshold (e.g. point cloud data limited to the first 2 m above the ground: Ewald et al., 2014; Lone, van Beest, et al., 2014), we could distinguish between structurally similar ground vegetation which occurred in different ecological contexts, such as the vegetation within open areas (Figure 3, PC1) vs. the

understorey vegetation beneath forest of different heights (Figure 3, PC2-PC5). Our LiDAR-PC approach was therefore able to disentangle the selection for understory vegetation at unprecedented fine scale.

The ability of each PC in explaining the variance in the 10-m-resolution LiDAR point cloud decreases from 34% in PC1 to less than 15% in PC5 and even less than 1% in PC9 (Figure 2). This means that higher ranking PCs explain major patterns in the LiDAR data (e.g. PC1, open areas vs. forest, see rotation matrix in Figure 3), whereas lower ranking PCs capture complex fine-scale differences in vegetation structure (e.g. PC4, forest stands with and without understory vegetation). Both deer species strongly selected for the environmental variability embraced by PC1-5, confirming the importance of decomposing the three-dimensional complexity of the forest into multiple principal components.

Interpretation of selection patterns, however, does not have to rely on single PCs. The statement that deer select for negative values of PC1 and thus open habitats is not wrong, in principle, but it clearly oversimplifies a much more sophisticated behaviour. Selection for PCs should be interpreted overall, at least for those first PCs that were strongly selected for by deer in our empirical example. We make use of Figure 4 to explain this feature of LiDAR-derived PCs values. Each habitat type depicted by

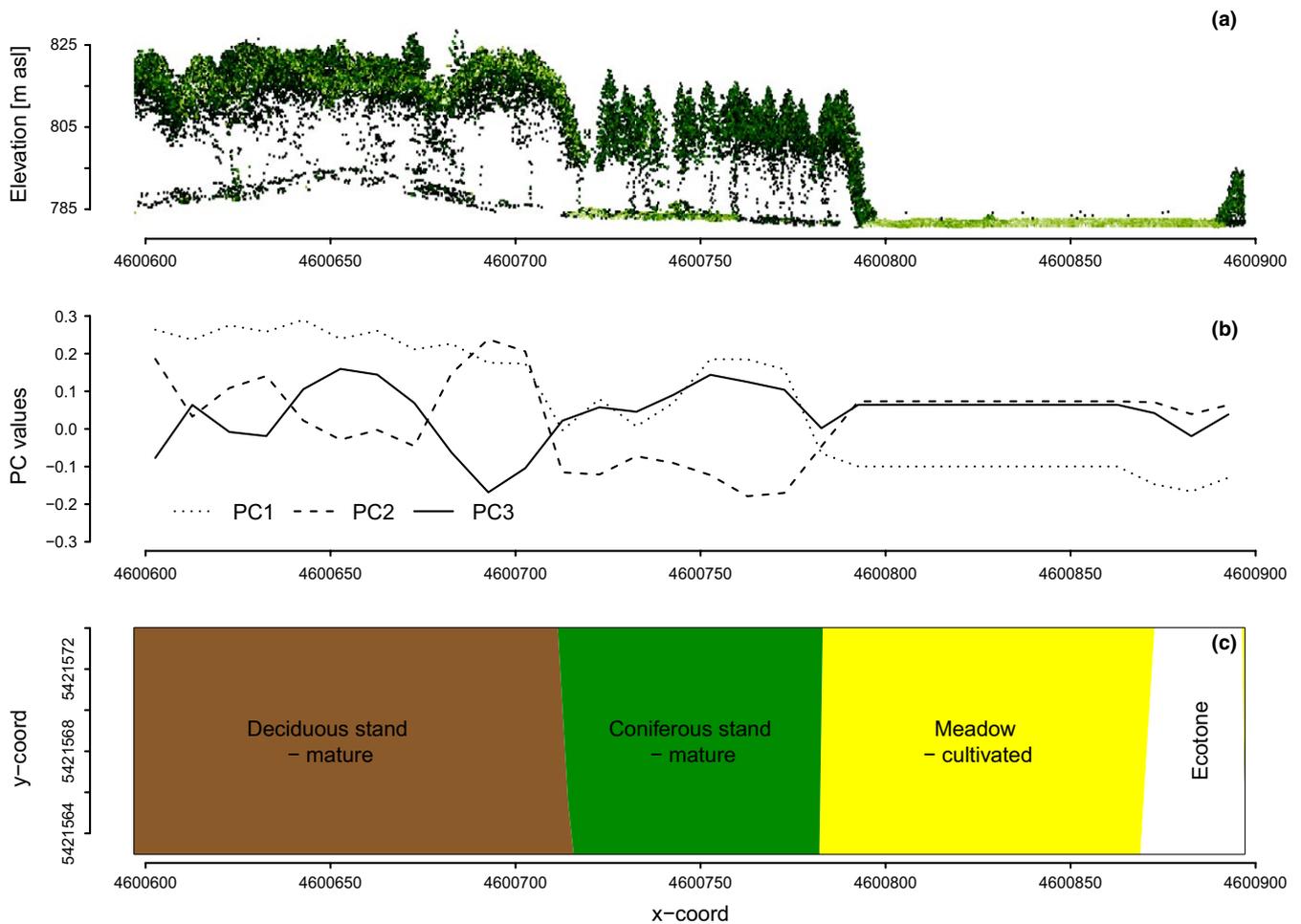


FIGURE 5 Comparison between (a) light detection and ranging point cloud data, (b) PC (principal component) values and (c) land-cover classification for a $300 \times 10 \text{ m}^2$ cross section within the Bavarian Forest National Park, Germany

the best land-cover map available for our study site can be described by a unique combination of PC values. Mature forest patches, for instance are characterized by positive values of PC1 and negative values of PC2 and PC3, whereas ecotones are characterized by negative values of PC1 as well as PC2 and PC3 values close to zero. In terms of ecological interpretation, this means that when deer are recorded to select for a certain combination of PC values, this pattern corresponds to the selection of a given habitat type. Clearly, this exercise is performed only for data interpretation, whereas the model has been trained with unclassified data.

Researchers also have the option to further improve result interpretation by fitting cluster analyses, random forests and several classification methods of their choice on significantly selected LiDAR-PCs (Breiman, 2001; Zuur et al., 2007), and interpret the results afterwards, for example by validation in the field.

Vegetation cross sections randomly selected in the study area (Figure 5; Figures S4 and S5) can help the reader to better appreciate how the variability included in PCs can be lost using common land-cover maps, even when the latter are of the highest achievable quality (<5 m resolution in our case). Variation in the vegetation structure in a mature deciduous stand (Figure 5), for instance is captured by varying combination of PC1, PC2 and PC3, whereas the land-cover map does not have

this level of detail (i.e. polygon assuming homogenous vegetation inside its area). The high standard deviation in Figure 4 indeed reflects the high variability in PCs within each “homogeneous” land-cover map category.

Finally, we reported in Figure 6 an example of suitability maps for both deer species, which can be used to visualize a model’s spatial predictions and understand the type of forest structure selected by deer. This type of map can be used by practitioners, to identify important spots of the landscapes they wish to manage, and to make decision based on conservation and management objectives. RSF scores predicted by our models can also be plotted against other environmental covariates, which can improve our understanding of the structure of the understorey vegetation selected by deer. This may be achieved by depicting, for example the relationships between RSF scores and fractional vegetation cover of the understorey vegetation (Figure S6).

4.2 | Cautionary notes to the readers: spatio-temporal autocorrelation and the use of PCs in regression analysis

We ask the readers to pay attention to our informative supplementary material before reproducing our methodology. Firstly, telemetry data

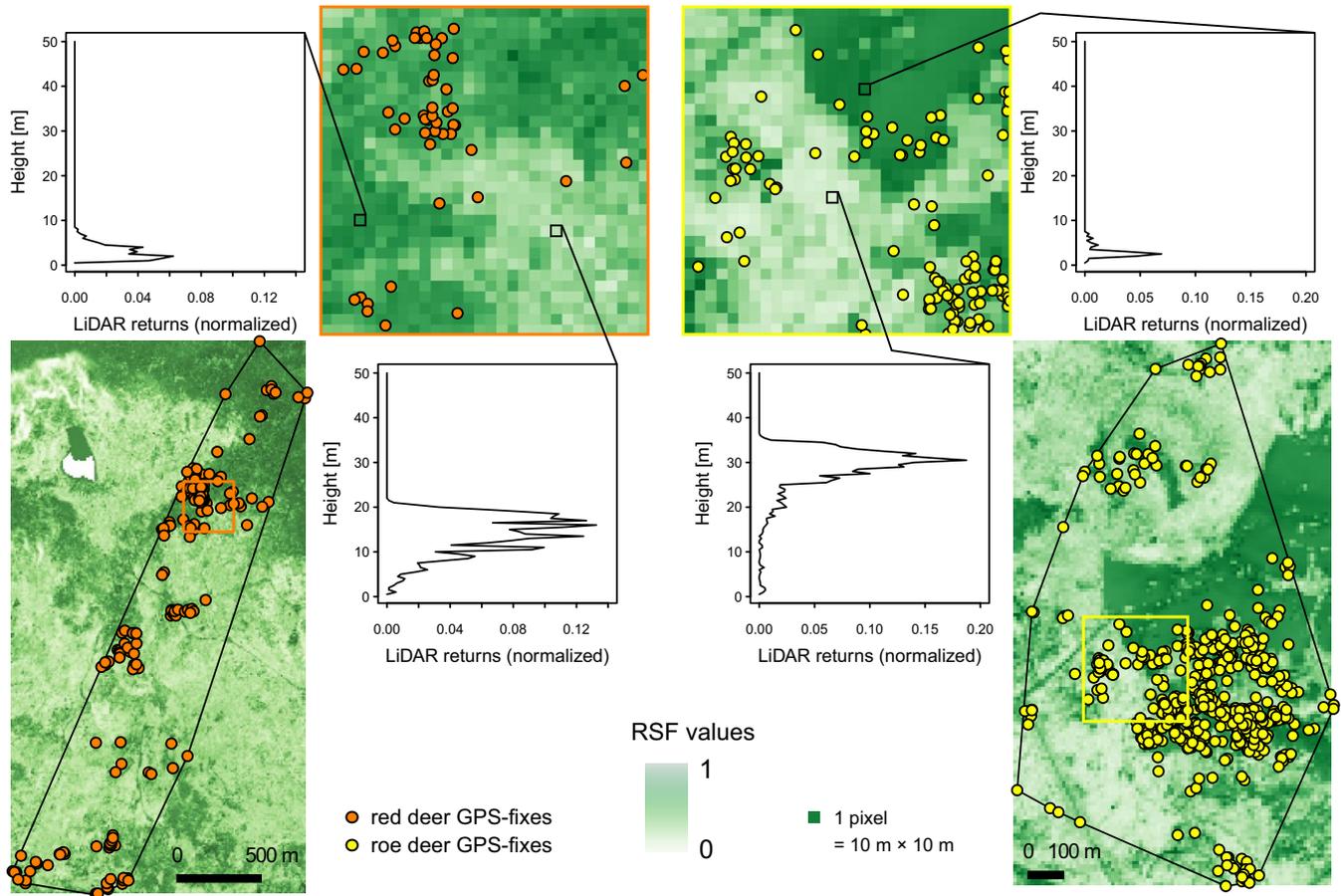


FIGURE 6 Relative probability of selection (rescaled resource selection function [RSF] values) predicted by best ranked models in the Bavarian Forest National Park for red deer “Burgl” and roe deer “Gitze” tracked with satellite telemetry (1-h fix rate, spring–summer locations surrounded by minimum convex polygons). Scenarios for both animals are for twilight, although practitioners can vary parameters plugged into the RSF equation and predict different scenarios. Small insets depict the relationship between predicted relative probability of selection and actual vegetation structure (i.e. light detection and ranging returns) for $10 \times 10 \text{ m}^2$ grid cells in which selection was high (dark green) and low (light green) for both species. White patch in the red deer sector (left) corresponds to a lake (RSF = 0)

(like those used in this study) typically are temporally and spatially correlated, and this will often be transferred to model residuals and not lead to valid inferences. Researchers need to use diagnostic tools to check for autocorrelation—which we did not find, see Appendix S2, although this is more the exception rather than the rule—and, whenever residual autocorrelation is found, include spatio-temporal structures in their models. See full discussion on this matter at the end of Appendix S2 along with significant literature cited. Second, note our cautionary reminder on the use of some (but not all) principal component axes in regression analysis (Appendix S3).

4.3 | LiDAR-PCs: a new tool in the ecological modelling toolbox

We did not include alternative environmental predictors within the LiDAR models because building the best possible model was beyond the scope of this paper. After screening for collinearity (Dormann et al., 2013), LiDAR and non-LiDAR derived predictors should be used as predictors in the one model to further improve predictive and explanatory performances. Our suitability model could

be slightly improved by adding NDVI and other spatio-temporal variable indexes such those derived by airborne hyperspectral data (Wang, Franklin, Guo, & Cattet, 2010). Hyperspectral data are an exciting prospect because they have been successfully applied in recording information regarding critical plant traits, discriminating tree species in landscapes (Nagendra & Rocchini, 2008) and estimating nutritional value of ground vegetation for ungulates (Schweiger et al., 2015).

We showed that the horizontal resolution matters (10 m vs. 5 m) and this should be carefully evaluated by users running sensitivity analysis prior to fitting ecological models. Note that this may depend on the ecological question under scrutiny and the scale at which the ecological process occurs (Schneider, 2001). Also, our results may suggest that LiDAR data have stronger predictive ability when their horizontal resolution (10 m) matches that of satellite-telemetry relocation error (10 m), where it is much less affected by noise due to resolution mismatch (Frair et al., 2010).

Animal resource selection studies and species distribution models are strongly affected by the scale of the analyses and the resolution of environmental predictors (Boyce, 2006; Seo, Thorne,

Hannah, & Thuiller, 2009). Uncertainties in SDMs can indeed result from the quality of environmental data (Seo et al., 2009), and surrogate environmental predictors may yield biased species habitat suitability estimates (Heikkinen et al., 2014), which can be improved by fine-scale LiDAR data (Nijland et al., 2014). The particularly detailed selection patterns that we described in our study (e.g. for understorey vegetation) is an example of how we can improve our understanding of fine-scale habitat selection patterns in large herbivores, with the potential to tackle more specific hypotheses on their foraging ecology and anti-predator strategies within multi-predator landscape of fear (Lone, Loe, et al., 2014). Herein we stress the importance to primarily fitting ecological models able to exploit LiDAR data variability, which can secondarily be manipulated after fitting the model in order to improve the interpretation and understanding of observed patterns.

In conjunction with GPS telemetry data, LiDAR-PCs provide an excellent chance to assess how animal movement is affected by vegetation structure. Therewith, new insights and finer details about how animals make decisions based on habitat structure can be achieved. Davies and Asner (2014) strongly recommended that future animal ecology studies will benefit by placing higher priority on metrics produced by LiDAR such as measurements related to vertical structural heterogeneity and complexity. This is the best way to proceed if the goal of the research is testing specific vegetation-structural hypotheses. However, we suggest our approach as the alternative to the use of LiDAR-based metrics whenever ecological modellers need to retain the full variability in point clouds, and boost predictive ecological models such as species distribution and habitat suitability models.

AUTHORS' CONTRIBUTIONS

S.C., H.T., P.A., C.F.D. and M.H. conceived the ideas and designed the methodology. M.H. gathered the data. H.T., P.A. and S.C. analysed the data. M.H., R.S.G., S.C. and C.F.D. contributed with additional data and analysis tools. S.C. led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

DATA ACCESSIBILITY

Red and roe deer telemetry data deposited in the Dryad Digital Repository <https://datadryad.org/resource/doi:10.5061/dryad.4t18d> (Ciuti et al., 2017)

ORCID

Simone Ciuti  <http://orcid.org/0000-0003-1052-9509>

REFERENCES

Ackers, S. H., Davis, R. J., Olsen, K. a., & Dugger, K. M. (2015). The evolution of mapping habitat for northern spotted owls (*Strix occidentalis caurina*): A comparison of photo-interpreted, landsat-based, and lidar-based

- habitat maps. *Remote Sensing of Environment*, 156, 361–373. <https://doi.org/10.1016/j.rse.2014.09.025>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Boyce, M. S. (2006). Scale for resource selection functions. *Diversity and Distributions*, 12, 269–276. <https://doi.org/10.1111/ddi.2006.12.issue-3>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Bunting, P., Armston, J., Clewley, D., & Lucas, R. M. (2013). Sorted pulse data (SPD) library-Part II: A processing framework for LiDAR data from pulsed laser systems in terrestrial environments. *Computers and Geosciences*, 56, 207–215. <https://doi.org/10.1016/j.cageo.2013.01.010>
- Bunting, P., Armston, J., Lucas, R. M., & Clewley, D. (2013). Sorted pulse data (SPD) library. Part I: A generic file format for LiDAR data from pulsed laser systems in terrestrial environments. *Computers and Geosciences*, 56, 197–206. <https://doi.org/10.1016/j.cageo.2013.01.019>
- Ciuti, S., Tripke, H., Antkowiak, P., Silveyra Gonzalez, R., Dormann, C. F., & Heurich, M. (2017). Data from: An efficient method to exploit LiDAR data in animal ecology. *Dryad Digital Repository*, <https://datadryad.org/resource/doi:10.5061/dryad.4t18d>
- Davies, A. B., & Asner, G. P. (2014). Advances in animal ecology from 3D-LiDAR ecosystem mapping. *Trends in Ecology & Evolution*, 29, 681–691. <https://doi.org/10.1016/j.tree.2014.10.005>
- Davies, A. B., Tambling, C. J., Kerley, G. I. H., & Asner, G. P. (2016). Effects of vegetation structure on the location of lion kill sites in African Thicket. *PLoS ONE*, 11, e0149098. <https://doi.org/10.1371/journal.pone.0149098>
- Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., ... Björn, R. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36, 27–46. <https://doi.org/10.1111/j.1600-0587.2012.07348.x>
- Dupke, C., Bonenfant, C., Reineking, B., Hable, R., Zeppenfeld, T., Ewald, M., & Heurich, M. (2016). Habitat selection by a large herbivore at multiple spatial and temporal scales is primarily governed by food resources. *Ecography*, 40, 1014–1027. <https://doi.org/10.1111/ecog.02152>
- Elith, J., & Leathwick, J. R. (2009). Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, 40, 677. <https://doi.org/10.1146/annurev.ecolsys.110308.120159>
- Evans, J. S., & Hudak, A. T. (2007). A multiscale curvature algorithm for classifying discrete return LiDAR in forested environments. *IEEE Transactions on Geoscience and Remote Sensing*, 45, 1029–1038. <https://doi.org/10.1109/TGRS.2006.890412>
- Ewald, M., Dupke, C., Heurich, M., Müller, J., & Reineking, B. (2014). LiDAR remote sensing of forest structure and GPS telemetry data provide insights on winter habitat selection of European roe deer. *Forests*, 5, 1374–1390. <https://doi.org/10.3390/f5061374>
- Farrell, S. L., Collier, B. a., Skow, K. L., Long, a. M., Campomizzi, a. J., Morrison, M. L., ... Wilkins, R. N. (2013). Using LiDAR-derived vegetation metrics for high-resolution, species distribution models for conservation planning. *Ecosphere*, 4, 1–18. <https://doi.org/10.1890/es12-000352.1>
- Fayad, I., Baghdadi, N., Bailly, J. S., Barbier, N., Gond, V., El Hajj, M., ... Bourguin, B. (2014). Canopy height estimation in French Guiana with LiDAR ICESat/GLAS data using principal component analysis and random forest regressions. *Remote Sensing*, 6, 11883–11914. <https://doi.org/10.3390/rs61211883>
- Flaspohler, D. J., Giardina, C. P., Asner, G. P., Hart, P., Price, J., Lyons, C. K., & Castaneda, X. (2010). Long-term effects of fragmentation and fragment properties on bird species richness in Hawaiian forests. *Biological Conservation*, 143, 280–288. <https://doi.org/10.1016/j.biocon.2009.10.009>
- Frair, J. L., Fieberg, J., Hebblewhite, M., Cagnacci, F., DeCesare, N. J., & Pedrotti, L. (2010). Resolving issues of imprecise and habitat-biased

- locations in ecological analyses using GPS telemetry data. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 365, 2187–2200. <https://doi.org/10.1098/rstb.2010.0084>
- Heikkinen, R. K., Bocedi, G., Kuussaari, M., Heliöla, J., Leikola, N., Pöyry, J., & Travis, J. M. J. (2014). Impacts of land cover data selection and trait parameterisation on dynamic modelling of species' range expansion. *PLoS ONE*, 9, e108436. <https://doi.org/10.1371/journal.pone.0108436>
- Heurich, M., Fischer, F., Knoerzer, O., & Krzystek, P. (2008). Assessment of digital terrain models (DTM) from data gathered with airborne laser scanning in temperate European Beech (*Fagus sylvatica*) and Norway Spruce (*Picea abies*) Forests. *Photogrammetrie, Fernerkundung, Geoinformation*, 6, 473–488.
- Jolliffe, I. T. (2002). *Principal component analysis*. New York, NY: Springer.
- La, H. P., Eo, Y. D., Chang, A., & Kim, C. (2015). Extraction of individual tree crown using hyperspectral image and LiDAR data. *KSCE Journal of Civil Engineering*, 19, 1078–1087. <https://doi.org/10.1007/s12205-013-1178-z>
- Lim, K., Treitz, P., Wulder, M., St-Onge, B., & Flood, M. (2003). LiDAR remote sensing of forest structure. *Progress in Physical Geography*, 27, 88–106. <https://doi.org/10.1191/0309133303pp360ra>
- Lone, K., Loe, L. E., Gobakken, T., Linnell, J. D. C., Odden, J., Remmen, J., & Mysterud, A. (2014). Living and dying in a multi-predator landscape of fear: Roe deer are squeezed by contrasting pattern of predation risk imposed by lynx and humans. *Oikos*, 123, 641–651. <https://doi.org/10.1111/more.2014.123.issue-6>
- Lone, K., van Beest, F. M., Mysterud, A., Gobakken, T., Milner, J. M., Ruud, H.-P., & Loe, L. E. (2014). Improving broad scale forage mapping and habitat selection analyses with airborne laser scanning: The case of moose. *Ecosphere*, 5, 1–22. <https://doi.org/10.1890/es14-00156.1>
- Manly, B. F. J., McDonald, L. L., Thomas, D. L., McDonald, T. L., & Erickson, W. P. (2002). *Resource selection by animals. Statistical design and analysis for field studies*, 2nd edn. Dordrecht/Boston/London: Kluwer Academic Publishers.
- Mardia, K. V., Kent, J. T., & Bibby, J. M. (1980). *Multivariate analysis (probability and mathematical statistics)*. London: Academic Press.
- Melin, M., Matala, J., Mehtätalo, L., Tiilikainen, R., Tikkanen, O. P., Maltamo, M., ... Packalen, P. (2014). Moose (*Alces alces*) reacts to high summer temperatures by utilizing thermal shelters in boreal forests – An analysis based on airborne laser scanning of the canopy structure at moose locations. *Global Change Biology*, 20, 1115–1125. <https://doi.org/10.1111/gcb.2014.20.issue-4>
- Nagendra, H., & Rocchini, D. (2008). High resolution satellite imagery for tropical biodiversity studies: The devil is in the detail. *Biodiversity and Conservation*, 17, 3431–3442. <https://doi.org/10.1007/s10531-008-9479-0>
- Neumann, W., Martinuzzi, S., Estes, A. B., Pidgeon, A. M., Dettki, H., Ericsson, G., & Radeloff, V. C. (2015). Opportunities for the application of advanced remotely-sensed data in ecological studies of terrestrial animal movement. *Movement Ecology*, 3, 8. <https://doi.org/10.1186/s40462-015-0036-7>
- Nijland, W., Nielsen, S. E., Coops, N. C., Wulder, M. a., & Stenhouse, G. B. (2014). Fine-spatial scale predictions of understory species using climate- and LiDAR-derived terrain and canopy metrics. *Journal of Applied Remote Sensing*, 8, 083572. <https://doi.org/10.1117/1.jrs.8.083572>
- Onojeghuro, A. O., & Blackburn, G. A. (2011). Optimising the use of hyperspectral and LiDAR data for mapping reedbed habitats. *Remote Sensing of Environment*, 115, 2025–2034. <https://doi.org/10.1016/j.rse.2011.04.004>
- R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Schneider, D. C. (2001). The rise of the concept of scale in ecology. *BioScience*, 51, 545–553. [https://doi.org/10.1641/0006-3568\(2001\)051\[0545:TROTCO\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2001)051[0545:TROTCO]2.0.CO;2)
- Schweiger, A. K., Risch, A. C., Damm, A., Kneubuhler, M., Haller, R., Schaepman, M. E., & Schutz, M. (2015). Using imaging spectroscopy to predict above-ground plant biomass in alpine grasslands grazed by large ungulates. *Journal of Vegetation Science*, 26, 175–190. <https://doi.org/10.1111/jvs.12214>
- Seo, C., Thorne, J. H., Hannah, L., & Thuiller, W. (2009). Scale effects in species distribution models: Implications for conservation planning under climate change. *Biology Letters*, 5, 39–43. <https://doi.org/10.1098/rsbl.2008.0476>
- Simonson, W. D., Allen, H. D., & Coomes, D. A. (2014). Applications of airborne lidar for the assessment of animal species diversity. *Methods in Ecology and Evolution*, 5, 719–729. <https://doi.org/10.1111/2041-210X.12219>
- Stache, A., Löttker, P., & Heurich, M. (2012). Red deer telemetry: Dependency of the position acquisition rate and accuracy of GPS collars on the structure of a temperate forest dominated by European beech. *Silva Gabreta*, 18, 35–48.
- Thurfjell, H., Ciuti, S., & Boyce, M. S. (2014). Applications of step-selection functions in ecology and conservation. *Movement Ecology*, 2, 4. <https://doi.org/10.1186/2051-3933-2-4>
- Vierling, K. T., Vierling, L. A., Gould, W. A., Martinuzzi, S., & Clawges, R. M. (2008). Lidar: Shedding new light on habitat characterization and modeling. *Frontiers in Ecology and the Environment*, 6, 90–98. <https://doi.org/10.1890/070001>
- Wagner, W., Ullrich, A., Ducic, V., Melzer, T., & Studnicka, N. (2006). Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60, 100–112. <https://doi.org/10.1016/j.isprsjprs.2005.12.001>
- Wang, K., Franklin, S. E., Guo, X. L., & Cattet, M. (2010). Remote sensing of ecology, biodiversity and conservation: A review from the perspective of remote sensing specialists. *Sensors*, 10, 9647–9667. <https://doi.org/10.3390/s101109647>
- Warren, D. L., & Seifert, S. N. (2011). Ecological niche modelling in MaxEnt: The importance of model complexity and the performance of model selection criteria. *Ecological Applications*, 21, 335–342. <https://doi.org/10.1890/10-1171.1>
- Zhang, K., Chen, S. C., Whitman, D., Shyu, M. L., Yan, J., & Zhang, C. (2003). A progressive morphological filter for removing nonground measurements from airborne LIDAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 41, 872–882. <https://doi.org/10.1109/TGRS.2003.810682>
- Zuur, A., Ieno, E. N., & Smith, G. M. (2007). *Analysing ecological data*. New York, NY: Springer-Verlag.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Ciuti S, Tripke H, Antkowiak P, Silveyra Gonzalez R, Dormann CF, Heurich M. An efficient method to exploit LiDAR data in animal ecology. *Methods Ecol Evol*. 2017;00:1–12. <https://doi.org/10.1111/2041-210X.12921>