Assessing the performance of object-oriented LiDAR predictors for forest bird habitat suitability modeling

Anouk Glad1,2, Björn Reineking1, Marc Montadert3, Alexandra Depraz2 & Jean-Matthieu Monnet1

1University Grenoble Alpes, Irstea, LESSEM, Grenoble 38000, France
2Groupe Tétras Jura, 9 impasse du tacon, Les Bouchoux 39370, France
3ONCFS, 90 impasse Les Daudes, Sévrier 74320, France

Abstract

Habitat suitability models (HSMs) are widely used to plan actions for species of conservation interest. Models that will be turned into conservation actions need predictors that are both ecologically pertinent and fit managers’ conceptual view of ecosystems. Remote sensing technologies such as light detection and ranging (LiDAR) can describe landscapes at high resolution over large spatial areas and have already given promising results for modeling forest species distributions. The point-cloud (PC) area-based LiDAR variables are often used as environmental variables in HSMs and have more recently been complemented by object-oriented (OO) metrics. However, the efficiency of each type of variable to capture structural information on forest bird habitat has not yet been compared. We tested two hypotheses: (1) the use of OO variables in HSMs will give similar performance as PC area-based models; and (2) OO variables will improve model robustness to LiDAR datasets acquired at different times for the same area. Using the case of a locally endangered forest bird, the capercaillie (Tetrao urogallus), model performance and predictions were compared between the two variable types. Models using OO variables showed slightly lower discriminatory performance than PC area-based models (average D_AUC = 0.032 and 0.01 for females and males, respectively). OO-based models were as robust (absolute difference in Spearman rank correlation of predictions ≤ 0.21) or more robust than PC area-based models. In sum, LiDAR-derived PC area-based metrics and OO metrics showed similar performance for modeling the distribution of the capercaillie. We encourage the further exploration of OO metrics for creating reliable HSMs, and in particular testing whether they might help improve the scientist–stakeholder interface through better interpretability.

Introduction

Habitat suitability models (HSMs), also known as environmental niche models, are statistical models that relate species occurrence to environmental factors (biotic and/or abiotic) (Guisan and Zimmermann 2000). HSMs quantify the species’ response to environmental variables and spatially predict the probability of occurrence of the target species. Both results can be used to plan conservation actions in favor of a target species (Guisan and Thuiller 2005; Johnson and Gillingham 2005; Franklin 2009): the response to the environmental variables can help characterize the optimal habitat and the distribution maps can be used to target areas for prospective conservation projects (Franklin 2009). Regardless of the modeling method chosen, the use of environmental predictors that are ecologically relevant is essential to obtain reliable results and to best explain the species distribution in the landscape (Johnson and Gillingham 2005; Fourcade et al. 2018).

Advances in remote sensing technologies are making them powerful tools for improving HSMs (He et al. 2015). One of these methods is light detection and ranging (LiDAR), which provides highly precise data on the
three-dimensional structure of the environment at high resolution over large areas (Vierling et al. 2008). LiDAR has been used in HSMs for a wide variety of species including forest birds, which are sensitive to variations in vegetation structure (Bergen et al. 2007, 2009). The use of LiDAR vegetation structure metrics has given promising results by significantly improving a model’s performance, either alone or when combined with other predictors (Graf et al. 2009; Farrell et al. 2013; Bae et al. 2014).

Broadly, one can distinguish two types of variables extracted from LiDAR point clouds: point-cloud (PC) area-based metrics and object-oriented (OO) metrics. PC area-based metrics summarize the characteristics of the point cloud within a given region of interest (typically a square pixel). Variables such as the standard deviation of the 10th percentile (Zellweger et al. 2013), the proportion of echoes above 5 m (Melin et al. 2016) or height skewness of the returns located between 1.5 m and 5 m (Kortmann et al. 2018) have been used to describe vegetation. However, these metrics may be difficult to interpret or communicate because they are not used as field metrics by managers (such as the percentage of vegetation cover by layer and the number of trees). Accordingly, recent efforts have been made by scientists to propose concrete recommendations for managers from PC area-based results by adapting the predictor description and designation (e.g., the proportion of echoes expressed as the proportion of vegetation in shrubs or canopy layers) (Melin et al. 2018).

Advances in image processing have promoted the use of OO methods with the objective of overcoming PC area-based limitations (Benz et al. 2004). Contrary with the area-based metrics, object-related metrics are defined by known common characteristics (structure, shape, texture), describing real-world objects such as buildings, roads and trees. Thus, with LiDAR data, different object types can be defined by their three-dimensional characteristics and be extracted from the raw point clouds. This type of approach has been widely used as a forest management tool by exploiting single tree detection methods (Othmani et al. 2013; Munoz et al. 2015). Several recent studies used OO metrics to produce easier interpretable variables such as gaps, percentage of deciduous trees and edge length (Rechsteiner et al. 2017; Kortmann et al. 2018), with a view toward management advice for species conservation. To date, the efficiency of these two types of metrics (PC area-based and OO) in extracting pertinent structural information on forest structure for modeling forest species distribution has not been evaluated.

The capercaillie (Tetrao urogallus) is a species classified as Least Concern worldwide and in Europe (Storch 2007). However, the species is classified as vulnerable in France and populations occurring in the Jura massif (France and Switzerland) are considered threatened (Storch 1994; Sachot 2002). The contraction of its occurring range is mainly due to habitat loss and alteration (Storch 2007; Mikoláš et al. 2017). In France, capercaillie favors old mixed forests constituted of a mosaic of habitats (Sachot et al. 2003) with a heterogeneous cover for both canopy and understory layers (Graf et al. 2009). The species is strictly dependent on a high proportion of moderate canopy density favoring the growth of a key food resource: the bilberry (Vaccinium myrtillus) (Storch 1993). Thus, the encroachment of the understory layer resulting in a poor proportion of bilberry in the habitat is leading to habitat degradation (Sachot et al. 2003). In addition, it is thought that according to capercaillie seasonal needs, its home range should contain mixed patches of young forest stages (predator shelter), clear-cuts (brood rearing) and patches of clear canopy density with fir (Abies alba) (Rolstad 1988; Rolstad et al. 1988; Storch 1994; Sachot 2002). In this context, obtaining capercaillie distribution predictions at the regional scale holds great promise to optimize and plan conservation actions. Numerous studies on this species have already been conducted in other countries to produce such species distribution maps (Sachot et al. 2006; Braunisch and Suchant 2007; Graf et al. 2007; Teuscher et al. 2013). More recently, in order to improve models, derived predictors from both LiDAR PC area-based and OO metrics have been used (Graf et al. 2009; Zellweger et al. 2013; Melin et al. 2016; Kortmann et al. 2018).

To explore the efficiency of OO LiDAR metrics as environmental variables, we evaluated two hypotheses on their performance relative to PC area-based models using the capercaillie as a case study: (H1) OO Lidar metrics show similar HSMs performance as PC area-based models, and (H2) OO variables improve model transferability robustness to a different LiDAR dataset, because the structure of an object is less dependent on acquisition characteristics such as point density or flight season (leaf-on/leaf-off) (Tiede et al. 2007; Gaulton and Malthus 2010). We compared model performance between the two types of variables for capercaillie winter observations in the French Jura, and the similarity between the predicted maps was assessed. Then the robustness of the model predictions to changes in the LiDAR campaign was evaluated for each variable type using a second LiDAR dataset over a 4.4-km² overlapping area, acquired about 1.5 years later than the first LiDAR dataset.

**Materials and Methods**

**Study area**

The study area is located in the east of France, in the Jura massif, located within the Ain, Jura and Doubs
departments (Fig. 1). The landscape is composed of a mosaic of small urban areas, pastures, forests, and fields. The massif is composed of a low plateau (elevation ranging from 400 m to 700 m a.s.l.) and a high plateau (elevation ranging from 700 m to 1620 m a.s.l.). On the lower plateau, forests are mainly deciduous-dominated forests composed of beech (*Fagus sylvatica*), ash (*Fraxinus excelsior*), oak (*Quercus petraea*), and spruce (*Picea abies*). The high plateau is characterized mainly by coniferous-dominated forest, composed of mixed spruce, fir and beech above 1000 m, and by the presence of a grassland subalpine landscape at the highest altitudes (IGN national forest inventory 2005–2009). The climate is continental with temperature varying between −5.9°C and 21.1°C and mean annual precipitation of 1187 mm (https://donneespubliques.meteo.fr/).

**LiDAR datasets**

The study area was covered by two airborne LiDAR surveys (Fig. 1) using a Riegl LMS Q680i system (RIEGL Inc., Horn, Austria). The system was operated with a maximum half-scan angle of 30°, and full wave-form digitization. The first LiDAR campaign, hereafter referred as the “Ain study area,” was conducted in autumn 2014 and covered 626 km². An average point density of 21.3 points/m² with a vertical error of 10 cm was obtained. The second LiDAR campaign, hereafter referred as the “Jura study area,” was conducted in summer 2016 and covered a surface of 431 km². An average point density of 18 points/m² with a vertical error of 10 cm was obtained. These two campaigns overlap over an area of 4.4 km², in the Ain department.

**Capercaillie (*Tetrao urogallus*) dataset**

Long-term capercaillie winter surveys were organized between 2007 and 2018 by the “Groupe Tétras Jura,” a nongovernmental organization. Observers surveyed forests known as favorable for capercaillie, by navigating through the focus areas according to observers’ preferences (data from 2007–2015). The implementation of a new survey protocol since 2016 for a large-scale capercaillie genetic survey involved the collection of droppings following a
standardized predefined path. Transect trajectories were separated by 80 m, and observers were requested to stay within 20 m of their assigned transect. Using a Global Navigation Satellite System receiver (Garmin 64s GLO-NASS receiver and 62s WAAS receiver), observers recorded their survey tracks and capercaillie signs (feces, prints, and feathers) and locations. Commercial-grade handheld GNSS receivers have an accuracy of around 10 m, 3 m for models with WAAS capacity. However, while surveys usually involved one observer, in some cases two to three observers were searching at the same time, leading to observations recorded at a distance of the main observer track (mean distance to tracks, 6.4 ± 11 m). The total length of survey tracks was 1206 km. Sex was assigned to each observation when possible based on feathers or dropping size. All observations without an assigned sex were removed and the 379 remaining observations were used for the analysis (207 males and 172 females) in order to create separate models for each sex due to their differences in habitat use (Thiel et al. 2007).

Spatial scale choice

The spatial scale is known to be an essential aspect for understanding the habitat use of individuals and populations (McGarigal et al. 2016). We therefore created models at multiple spatial scales. Six scales were chosen among three hierarchical orders: home range (56 ha, 27.5 ha and 15.8 ha), patch (1.8 ha and 0.81 ha), and micro-habitat (0.31 ha) (Meyer and Thuiller 2006). For each scale, the mean (indicating the general habitat characteristic) and the standard deviation (indicating the horizontal spatial heterogeneity of each variable) of LiDAR metrics (presented in detail in the next section) were calculated for each pixel over the whole study area using a circular moving window of 35, 25, 19, 7, 5, and 3 pixels in diameter. This task was done with the R package rgrass7 and the r.neighbors function (Bivand et al. 2016).

LiDAR-derived predictors

Raw point clouds were normalized with the lasnormalize function (tin algorithm option) implemented in lidR package v2.0.0 (Roussel and Auty 2016). Both PC area-based and OO metrics were calculated using the R packages lidR and lidaRtRee (https://gitlab.irstea.fr/jean-matthieu.monnet/lidaRtRee) using the normalized point clouds (Fig. 2).

Point-cloud area-based metrics

Point-cloud area-based metrics were calculated for 25-m × 25-m pixels. Six variables were selected with the aim to capture habitat components that are important for capercaillie (Table 1). (1) The relative density of the canopy between 10 m and 20 m was chosen as a proxy for the canopy cover (Næsset 2004; Ståhl et al. 2010). (2) To take into account horizontal heterogeneity, the standard deviation (SD) of the 20-30 m relative canopy density and the SD of the penetration ratio between 2 and 5 m, calculated for the different scales (among pixels within a circular window of the size of the given scale, e.g., 1.8 ha) were selected (Graf et al. 2009; Bae et al. 2014). (3) The mean penetration ratio between 2 m and 5 m was selected as a variable for shrub characteristics (Bae et al. 2014). (4) Aiming to take into account unfavorable low vegetation area (non-forest) the height 25th quantile (Q25) was selected (Latifi et al. 2015). (5) The vertical heterogeneity of the vegetation (Kortmann et al. 2018) was quantified by the Simpson index for canopy height. This index was calculated among the quantile height values (10, 25, 50, 75, 90% quantile strata) using the function diversity of the R package vegan (Oksanen et al. 2016) ($S = 1−\sum p_i^2$, where $p_i$ is the proportional abundance of returns in stratum $i$). The analyses were performed using each of the selected variables (mean or SD) at the six different spatial scales selected.

Object-oriented metrics

Two types of objects were extracted: trees (Kortmann et al. 2018) and gaps. All objects were segmented on the canopy height model over the 0.25-m$^2$ pixel (0.5 m × 0.5 m). Concerning trees, we used a tree-top detection method (method #1 in Eysn et al. 2015) where the minimum height for local maxima detection was set at 5 m. After the extraction of individual trees, metrics such as tree density (number of trees per ha) were derived over pixels with a surface covering 625 m$^2$ (25 m × 25 m).

Gaps were selected among areas characterized by a height lower than 1 m surrounded by trees whose height was less than twice the gap width and with a minimum surface of 25 m$^2$ (Fig. 3). The aim here was to propose functional gaps in accordance with the target species requirements. Such gaps need to allow sufficient light input to the ground for the development of bilberry, the key capercaillie food resource. Indeed, in the case of gaps surrounded by tall trees the ground remains in the shade, limiting the growth of species that need an intermediate irradiance input (Parlane et al. 2006).

From this basic gap map, habitat metrics were derived such as percentage of gaps by pixel (625-m$^2$ surface) with different surface categories: small gaps are objects with a 25- to 200-m$^2$ surface area, medium-sized gaps are objects with a 200- to 1000-m$^2$ surface area and open areas are objects with a surface >1000 m$^2$. This means, for example, that if 50% of the pixel is composed of medium-sized
gap objects (the objects can be larger than the pixel), then
the proportion of medium-sized gaps will be 50% in that
pixel (Fig. 4).

The metrics were chosen to represent different habitat
components as presented with the PC area-based metrics
(Table 1). (1) The density of the canopy was represented
by the density of trees higher than 10 m. (2) The hori-
zontal heterogeneity was represented by two variables: the
SD of the density of trees higher than 20 m and the pro-
portion of small gaps. (3) The density of small trees was
represented by the density of trees between 5 m and
10 m (number of trees per ha). (4) The presence of gaps
in the canopy was represented by the proportion of med-
ium-sized gaps. (5) The presence of non-forest area was
represented by the proportion of open areas. (6) To re-
present the vertical heterogeneity, the Gini Index of tree
heights was also calculated using the Gini function from
the R package reldist

\[ 1 - \sum_{k=1}^{n} (X_k - X_{k-1})(Y_k + Y_{k-1}), \]

where \( X_k \) is the cumulated proportion of the population
variable and \( Y_k \) the cumulated proportion of the tree
height sum variable, \( n \) is the number of strata: 4 for OO
(5–10 m; 10–20 m, 20–30 m, 30–60 m) and 7 for PC
(0.5–1 m; 1–2 m; 2–5 m; 5–10 m; 10–20 m; 20–30 m;
30–60 m) (Handcock 2016). The analyses were performed
using each of the selected variables (mean or SD) at the
six different spatial scales selected after discussion with
managers and from the literature (Storch 2002; Zellweger
et al. 2013) (Table 2).

Other metrics
To take into account the potential effect of disturbances
on capercaillie (Thiel et al. 2008), the distance to the clos-
est ski trails (Groupe Tétras Jura dataset) was used to
quantify this effect in all models. In the Jura massif the

Figure 2. Visualization of the calculated light
detection and ranging (LiDAR) metrics.
main winter activity is cross-country skiing, which is believed to cause disturbances because of the many trails penetrating into the forest units. Furthermore, the trails may also be used for other activities (snowshoeing, sledging, dog sleds) due to their easy access.

Species distribution modeling methods

Models were created for each sex and scale separately. Here, we used a recent reimplementation of Maxent, building on the equivalence of Maxent with an infinitely weighted logistic regression using the R package maxnet (Phillips et al. 2017). The spatial sampling bias generated by the uneven sampling of the study area by the observers (with a large majority of unsystematic sampling) was corrected by a targeted background point method (Phillips et al. 2009). Ten thousand background points were randomly located along the observer’s sampling trajectories (giving a density of 8 points per 1 km), where distances between each background point and trajectories followed a logspline distribution fitted to capercaillie observation data to account for the effect of multiple observers and detection distances. All explanatory variables were standardized to the 0 mean and standard deviation of 1 before modeling. In addition, the pairwise Pearson correlation between explanatory variables was evaluated prior to each analysis. When the absolute correlation value was higher than 0.7, one of the variables was dropped (Dormann et al. 2013), except for variables chosen to represent the interaction between landscape horizontal heterogeneity ($\text{sd}$ of canopy density, 20–30 m; $\text{sd}$ of tree $>20$ m density) and variables describing the general landscape type (e.g., open area, forest; Q25 and proportion of open areas), which where both retained. For all models, both linear and quadratic terms were selected.

Comparison of models results

Model performance

First, the performance of each model was evaluated using the area under the curve (AUC), which corresponds to the probability that a randomly drawn occurrence point has a higher predicted occurrence probability than a randomly drawn background point. The AUC was evaluated on a 10-fold cross-validation where the mean and standard error were calculated using the R package cvAUC (LeDell et al. 2015).

Variable contributions and responses

The contribution of each variable to determining model predictions was estimated by comparing the model coefficients. All absolute coefficient values associated with the same variable (linear, quadratic or product) were added to obtain the final contribution value. The model coefficient of product variables was divided by two before being added to the linear and quadratic values of each

Table 1. LiDAR-extracted variable name and description for the two types of metrics: PC and OO.

| Type of metrics | Metric name                        | Description                                                                                                                                 |
|-----------------|-----------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------|
| PC              | Relative density of the 10–20-m canopy | Point relative density for the height layer 10- to 20-m                                                                                   |
|                 | Relative density of the 20–30-m canopy | The standard deviation of the canopy density for the height layer 20- to 30-m                                                             |
|                 | Penetration ratio 2–5 m            | Penetration ratio standard deviation for the 2- to 5-m height layer                                                                       |
|                 | Q25                               | 25th Quantile of height of vegetation and unclassified points                                                                            |
|                 | Simpson index                      | Simpson index for canopy height                                                                                                          |
| OO              | Tree density (5–10 m)              | Density of trees shorter than 10 m but taller than 5 m (number/ha)                                                                         |
|                 | Tree density (>10 m)               | Density of trees taller than 10 m (number/ha)                                                                                             |
|                 | Tree density (>20 m)               | Standard deviation of density of trees taller than 20 m (number/ha)                                                                       |
|                 | Gini index                         | Tree height Gini index                                                                                                                   |
|                 | Open area                          | Proportion in the pixel that contains a grassland object (surface $<1000 \text{ m}^2$ and vegetation height $<1 \text{ m}$) (%)            |
|                 | Medium gap                         | Proportion in the pixel that contains a large gap object from 200 $\text{ m}^2$ to 1000 $\text{ m}^2$ (Height is $<1 \text{ m}$ and half height of surrounding trees is less than half of the gap width) (%) |
|                 | Small gap                          | Standard deviation of the proportion in the pixel that contains an object large gap from 25 $\text{ m}^2$ to 200 $\text{ m}^2$ (Height is $<1 \text{ m}$ and half the height of surrounding trees is less than half the gap width) (%) |

PC, point-cloud; OO, object-oriented; LiDAR, light detection and ranging.
Figure 3. The selection process of the gap objects.

Figure 4. The calculation of the proportion of gap variables for three different surfaces (small gaps 25–200 m², medium gaps (200–1000 m²) and open areas (>1000 m²).
Performance of Object-Oriented LiDAR Metrics

Table 2. Mean AUC (and standard error) obtained with a 10-fold cross-validation for each scale and model type.

| Scale  | Female AUC | Male AUC |
|--------|------------|----------|
|        | PC         | OO       | PC         | OO       |
| 0.31 ha| 0.75 (0.02)| 0.72 (0.02)| 0.75 (0.01)| 0.75 (0.01) |
| 0.81 ha| 0.76 (0.02)| 0.73 (0.02)| 0.75 (0.01)| 0.78 (0.01) |
| 1.8 ha | 0.77 (0.02)| 0.73 (0.02)| 0.78 (0.01)| 0.77 (0.01) |
| 15.8 ha| 0.74 (0.02)| 0.73 (0.02)| 0.79 (0.02)| 0.78 (0.01) |
| 27.5 ha| 0.75 (0.02)| 0.72 (0.02)| 0.8 (0.01)| 0.78 (0.02) |
| 56 ha | 0.76 (0.02)| 0.71 (0.02)| 0.79 (0.01)| 0.74 (0.02) |

AUC, area under the curve; PC, point-cloud; OO, object-oriented.

variable. In addition, the variables associated with a contribution value below 5% were dropped from the analysis of the results, keeping only the most important variables to facilitate visibility. The response curves were calculated for the three most contributing LiDAR-based variables (except the distance to ski trails variable).

Predictions

Predictions of the probability of occurrence were calculated over the Ain study area. Then, using the models trained with observations and LiDAR data from the Ain study area, predictions were calculated using the second LiDAR dataset over the overlapping area. Afterwards, Spearman rank correlation coefficients between PC area-based and OO predicted maps were calculated at each scale over the Ain study area. To assess the robustness of model predictions across different LiDAR campaigns, Spearman rank correlation coefficients between predictions by type (PC area-based or OO) were examined over the overlapping surface of the two LiDAR campaigns. All predicted maps were generated using the exponential (also called “raw”) output.

Results

Model performance

Cross-validated model performance across variable types, spatial scales, and sexes was evaluated using the AUC index (Table 2). Models showed moderate to good performance according to AUC (0.71–0.8; standard errors ≤0.02). Overall, OO-based models had a lower performance than PC area-based models; on average, the OO model AUC was lower by ~0.032 for females, and ~0.01 for males. The largest differences in AUC between OO and PC variables, ΔAUC = 0.05, were found at the coarsest spatial scale of 56 ha. The spatial scale had a substantial effect on model performance (ΔAUC 0.02–0.05). However, there was no consistently best spatial scale for males or females across the two variable types. Models for females showed poorer performance than models for males in all cases (maximum AUC difference for the same spatial scale: 0.05 for PC area-based and 0.07 for OO-based models).

Variable contributions and response

In both PC area-based and OO models, and at all scales for both sexes, the variable distance to ski trails was highly contributive, between 20% and 33% (Fig. 5). The response curve showed an optimal distance of 1 km from ski trails for both sexes (Supporting Information Figure S1). For OO models, the Gini Index, the sd of tree (>20 m) density and the proportion of open areas contributed the most at each scale and for both sexes. The tree (>10 m) density and sd of the small gaps contributed only at the 0.31-ha and 0.81-ha scales for both sexes. Regarding the PC area-based models, the Simpson Index, the 10- to 20-m canopy density, the Q25 and the sd of tree density contributed the most at each scale and for both sexes.

We calculated response curves of the three most contributing LiDAR variables for each model type (Figs. 6 and 7). (1) The response curves of the Gini Index and the Simpson Index showed a monotonously increasing suitability across scales, with the exception of models for males at the fine scale (0.31 ha and 0.81 ha), where models indicated an optimum at intermediate values for the Gini Index. (2) The response of the Q25 variable showed similar quadratic responses between sexes, except for two fine scales (0.31 ha and 0.81 ha), where a negative linear response was observed for females. An understory height from 2 to 5 m was more favorable. (3) The response to the 10- to 20-m canopy density for females showed that an optimum ratio of 0.2 (number of points in a layer divided by the total number of points) was preferred. At large spatial scales, ratios of 0 were also suitable. The responses were highly different for males, at the smallest scales showing a more suitable habitat for either low or high canopy density, whereas an optimum at the 0.2-ratio was observed at large scales, with a similar response at low density values, as for females. (4) Regarding the sd of tree (>20 m) density from the OO-type models, males and females had different responses. For males, a negative linear response was observed for two scales (0.31 ha and 56 ha). The other responses indicated values between 50 trees/ha and 70 trees/ha as unsuitable for the species, whereas extreme lower and higher values were more suitable. For females, the exact opposite response was observed at larger scales with more suitable habitat.
predicted for the same values. At the smallest scales, the observed response showed very slight variations, and overall were favorable. (5) For the proportion of open areas variable, a higher percentage of open area was unsuitable for males at small scales. At 15.8 ha and 27.5 ha, a positive response was found, whereas at the large 56-ha scales, an optimum at 50% open area was found. For females, a linear negative response was observed for two scales (0.31 ha and 0.81 ha), but a positive response was observed for the other scales. For the largest 56-ha scale, an optimum at 50% was observed.

Predictions

Predictions were calculated for each scale and sex with the two types of metrics (Fig. 8) to evaluate the similarity between OO and PC model predictions. The Spearman rank correlation coefficients between PC area-based and OO predictions ranged from 0.51 to 0.80 (Table 3). The lowest correlations between predictions, for both female and male models, were found at the largest scale (56 ha), with correlations <0.69. Higher correlation values were observed for male models at each scale.

In the second step, the robustness of the model to changes of the LiDAR survey (the two LiDAR surveys were done in two different seasons, which can influence point densities and ratios depending on changes in tree foliage density) was evaluated separately for each metric type. The correlation ranged between 0.45 and 0.77 for females and 0.40 and 0.70 for males. OO-based models were overall more robust, showing overall higher correlation values than PC area-based models for both females and males (Table 4). In sum, OO models were as robust (absolute difference in correlation ≤0.02) or more robust than PC area-based models.

Discussion

The aim of this study was to evaluate the use of OO LiDAR-extracted variables to predict the distribution of a locally endangered forest bird, the capercaillie. To achieve this objective, we compared model performance and robustness using two types of LiDAR-extracted variables: commonly used PC area-based and OO metrics. The hypothesis that models with OO variables perform better than or similarly to PC area-based models for the distribution of capercaillie was not confirmed by our observations. Overall, OO models performed slightly less well than PC area-based models for both sexes, particularly at larger spatial scales. However, the influence of scale was highlighted by both model types, showing variations in model performance.
A high contribution and a similar response shape were observed for each variable describing the heterogeneity of the vertical structure (the PC area-based metric Simpson Index and the OO metric Gini Index), highlighting the importance of forest stand heterogeneity for the species. Furthermore, this result confirms the equivalence between these two variable types in our models. However, concerning the other variables, a general pattern was not observed. In PC area-based models, the canopy density and height contributed the most, indicating the importance of the cover and the vertical structure, whereas in OO models the proportion of open areas and the SD of tree (>20 m) density were the most contributing, indicating here the importance of the horizontal heterogeneity. The response curves showed that the less contributing variables were also different between scales and sexes for both PC area-based and OO models, making it difficult to extract general conclusions from these results.

A consistency of the predictions across scales was not observed, with a correlation between PC area-based and OO maps in some cases low for models of females, in particular at the 56-ha scale (0.51) and the 27.5-ha scale (0.58). These results indicate that in some cases the models substantially diverge in spatial predictions, resulting in uncertainty regarding model accuracy, possibly because the models created from PC area-based and OO variables may capture different characteristics of the environment, leading to differences in predictions. Most particularly, the extraction of understory objects remains a challenge due to the difficulty finding the vertical separation between small and tall trees; the extraction of such objects was therefore not addressed in this study. Information on

Figure 6. Male response curves.
the understory is provided by two PC area-based variables in our models: the 2- to 5-m penetration ratio (mean and SD) and the Q25. In contrast, in OO models, the minimum height for tree-top detection was 5 m and the detection of trees under or near higher trees is known to be a limitation of the maxima detection (65% omission rate, for multi-layered forest) (Eysn et al. 2015). Yet, understory is known to constitute an essential component of the capercaillie habitat, which may explain why OO models performed worse than PC area-based models. Nevertheless, recent progress in LiDAR processing may soon fill this gap and allow the extraction of objects below the canopy (Hamraz et al. 2017). In light of these results, the optimal habitat for capercaillie should include trees of diverse heights from 5 m to 20 m, at both small and large scales. In addition, at the small scale the presence of open areas with a surface higher than 1000 m² is unfavorable for the species in winter (avoid large clear-cutting), but the presence of such habitat type can be maintained at the large scale at less than 25% of the surface. In addition, the presence of vegetation below 5 m should be favored because high Q25 values are not favorable.

Also, the role of feature types (here quadratic and linear) in models may greatly influence the predictions, in particular for extreme values. The response curves presented for the three main contributing variables in some cases differed greatly between sexes and scales. In some cases, this might be due to the limited shapes of response curves that can be represented by quadratic features. The response curves should therefore be interpreted with caution, for example in the case of the influence of ski trails. The use of more flexible, but adequately constrained, response curves using new features (e.g., hinges, splines) could reduce this effect (Ranc et al. 2017).
There was some evidence that models using OO variables had similar or higher robustness than PC area-based models, particularly at finer spatial scales. Yet, the overall robustness to another LiDAR dataset was moderate, with 17 cases with values below 0.7 of 24 models. These observations indicate that despite the use of relative PC area-based variables (relative density and ratio), which are less sensitive to the scanning parameters (Kerän et al. 2016), and the use OO-based variables (gaps and trees), differences between LiDAR campaigns may still require separate models to be more accurate than a global model (Chasmer et al. 2006; Hopkinson 2007).

We observed that PC and OO metrics can both be used as proxy for vegetation structures and each type has its own limitation (better understory estimation for PC, higher robustness for OO). In addition to the results presented in this study, one of the potential advantages in using OO metrics as predictors in HSMs is that they allow a better comprehension and transmission of results. The type of object and its structural definition can be defined closely in collaboration with managers regarding their needs and expectations. Close collaboration and cooperation is known to be a key point of successful conservation actions in order to take into account the specificities of each project (Arlettaz et al. 2010; Leidner and Buchanan 2018). Furthermore, conservation decisions, in particular in forest ecosystems, are made not only with scientists and local managers, but also considering the opinions of private owners or elected representatives, which highlights the need for variables that will be widely understood. Using OO metrics, remote sensing scientists can now propose pertinent variables that both

Table 3. Spearman rank correlation coefficients between PC area-based and OO predictions over the Ain study area.

| Scale    | Female | Male |
|----------|--------|------|
| 0.31 ha  | 0.75   | 0.78 |
| 0.81 ha  | 0.74   | 0.78 |
| 1.8 ha   | 0.74   | 0.78 |
| 15.8 ha  | 0.69   | 0.77 |
| 27.5 ha  | 0.58   | 0.80 |
| 56 ha    | 0.51   | 0.69 |

PC, point-cloud; OO, object-oriented.

Table 4. Spearman rank correlation coefficients between predictions using two different LiDAR datasets for the same study region.

| Scale  | Female | Male |
|--------|--------|------|
|        | PC     | OO   |
|        | PC     | OO   |
| 0.31 ha | 0.65   | 0.77 |
| 0.81 ha | 0.61   | 0.73 |
| 1.8 ha  | 0.56   | 0.76 |
| 15.8 ha | 0.52   | 0.66 |
| 27.5 ha | 0.46   | 0.67 |
| 56 ha   | 0.55   | 0.56 |

PC, point-cloud; OO, object-oriented; LiDAR, light detection and ranging.
represent the target species requirements and are likely meaningful for managers. While more progress still needs to be made to propose tools to facilitate visualization of the results (Bailey et al. 2018), we believe that using OO variables is a first step that will help fill the gap between scientists and managers.

**Conclusion**

The use of LiDAR-extracted object-oriented variables had similar model performance and robustness to changes within LiDAR campaigns characteristics. A strong variation in terms of performance was observed in scales and sex, indicating that the type of variable used is not the only important component influencing model accuracy in this case study. This result shows that this type of metric can be a competitive asset in habitat suitability modeling. We encourage the further exploration of object-oriented metrics for creating reliable habitat suitability models to other species and locations, in particular to test whether they might help improve the scientist–stakeholder interface through better interpretability.

**Acknowledgements**

The first LiDAR campaign was funded by the European Territorial Cooperation program INTERREG IV France-Switzerland 2007–2013 and the Ain department. The Conseil régional Bourgogne Franche-Comté funded the second LiDAR campaign and field missions. Anouk Glad’s PhD grant was provided by the Ain department as part of the “Plan départemental d’actions en faveur des tétraonidés.” The authors would also like to thank the Groupe Tétars Jura for providing the field dataset gathering observations from volunteers and partner institutions (ONCFS, ONF, LPO, CRPF, ADEFOR, PNR du Haut Jura, RNN du Lac de Remoray et RNN de la Haute Chaîne du Jura).

**Conflict of Interest**

The authors declare no conflict of interest.

**References**

Arlettaz, R., M. Schaub, J. Fournier, T. S. Reichlin, A. Sierro, J. E. Watson, et al. 2010. From publications to public actions: when conservation biologists bridge the gap between research and implementation. *Bioscience* **60**, 835–842.

Bae, S., B. Reineking, M. Ewald, and J. Mueller. 2014. Comparison of airborne lidar, aerial photography, and field surveys to model the habitat suitability of a cryptic forest species—the hazel grouse. *Int. J. Remote Sens.* **35**, 6469–6489.

Bailey, H., E. Hazen, B. Mate, S. J. Bograd, L. Irvine, D. M. Palacios, et al. 2018. Lessons learned from WhaleWatch. Pp. 229–273 in A. K. Leidner and G. M. Buchanan, eds. *Satellite Remote Sensing for Conservation Action: Case Studies from Aquatic and Terrestrial Ecosystems*. Cambridge University Press, Cambridge.

Benz, U. C., P. Hofmann, G. Willhauck, I. Lingenfelder, and M. Heynen. 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS J. Photogramm. Remote Sens.* **58**, 239–258.

Bergen, K. M., A. M. Gilboy, and D. G. Brown. 2007. Multi-dimensional vegetation structure in modeling avian habitat. *Ecol. Inform.* **2**, 9–22.

Bergen, K. M., S. J. Goetz, R. O. Dubayah, G. M. Henebry, C. T. Hunsaker, M. L. Imhoff, et al. 2009. Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for lidar and radar spaceborne missions. *J. Geophys. Res. Biogeosci.* **114**, G00E06–G00E06.

Bivand, R., R. Krug, M. Neteler, and S. Leworutzki. (2016). rgrass7: *Interface Between GRASS 7 Geographical Information System and R*. R package version 0.1-10. Available from: https://CRAN.R-project.org/package=rgrass7

Braunisch, V., and R. Suchant. 2007. A model for evaluating the habitat potential of a landscape for capercaillie *Tetrao urogallus*: a tool for conservation planning. *Wildl. Biol.* **13**, 21–33.

Chasmer, L., C. Hopkinson, B. Smith, and P. Treitz. 2006. Examining the influence of changing laser pulse repetition frequencies on conifer forest canopy returns. *Photogramm. Eng. Remote Sens.* **72**, 1359–1367.

Dormann, C. F., J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, et al. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **36**, 27–46.

Eysn, L., M. Hollaus, E. Lindberg, F. Berger, J.-M. Monnet, M. Dalponte, et al. 2015. A benchmark of lidar-based single tree detection methods using heterogeneous forest data from the alpine space. *Forests* **6**, 1721–1747.

Farrell, S. L., B. A. Collier, K. L. Skow, A. M. Long, A. J. Campomizzi, M. L. Morrison, et al. 2013. Using LiDAR-derived vegetation metrics for high-resolution, species distribution models for conservation planning. *Ecosphere* **4**, 1–8.

Fourcade, Y., A. G. Besnard, and J. Secondi. 2018. Paintings predict the distribution of species, or the challenge of selecting environmental predictors and evaluation statistics. *Glob. Ecol. Biogeogr.* **27**, 245–256.

Franklin, J. 2009. Moving beyond static species distribution models in support of conservation biogeography. *Divers. Distrib.* **16**, 321–330.

Gaulton, R., and T. J. Malthus. 2010. LiDAR mapping of canopy gaps in continuous cover forests: a comparison of canopy height model and point cloud based techniques. *Int. J. Remote Sens.* **31**, 1193–1211.
Leidner, A. K., and G. M. Buchanan. 2018. Operational conservation remote sensing. Pp. 301–317 in A. K. Leidner and G. M. Buchanan, eds. Satellite Remote Sensing for Conservation Action: Case Studies from Aquatic and Terrestrial Ecosystems. Cambridge University Press, Cambridge.

McGarigal, K., H. Y. Wan, K. A. Zeller, B. C. Timm, and S. A. Cushman. 2016. Multi-scale habitat selection modeling: a review and outlook. *Landsc. Ecol.* 31, 1161–1175.

Melin, M., L. Mehtäralo, J. Miettinen, S. Tossavainen, and P. Packalen. 2016. Forest structure as a determinant of grouse brood occurrence – An analysis linking LiDAR data with presence/absence field data. *For. Ecol. Manage.* 380, 202–211.

Melin, M., S. Hinsley, R. Broughton, P. Bellamy, and R. Hill. 2018. 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.* 33, 895–910.

Meyer, C. B., and W. Thuiller. 2006. Accuracy of resource selection functions across spatial scales. *Divers. Distrib.* 12, 288–297.

Mikoláš, M., M. Tejkal, T. Kuemmerle, P. Griffiths, M. Svoboda, T. Hlásky, et al. 2017. Forest management impacts on capercaillie (*Tetrao urogallus*) habitat distribution and connectivity in the Carpathians. *Landsc. Ecol.* 32, 163–179.

Munoz, A., J. Bock, J. Monnet, J. Renaud, A. Jolly, and C. Riond. 2015. Évaluation par validation indépendante des prédictions des paramètres forestiers réalisées à partir de données LiDAR aéroporté. *Rev. Fr. Photogrammétrie Télédétection* 211, 81–92.

Nesset, E. 2004. Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scand. J. For. Res.* 19, 164–179.

Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, et al. 2016. vegan: Community Ecology Package. R package version 2.4-2. Available from: https://CRAN.R-project.org/package=vegan.

Othmani, A., L. F. L. Y. Voon, C. Stolz, and A. Piboule. 2013. Single tree species classification from Terrestrial Laser Scanning data for forest inventory. *Pattern Recognit. Lett.* 34, 2144–2150.

Parlane, S., R. Summers, N. Cowie, and P. Van Gardingen. 2006. Management proposals for bilberry in Scots pine woodland. *For. Ecol. Manag.* 222, 272–278.

Phillips, S. J., M. Dudík, J. Elith, C. H. Graham, A. Lehmah, J. Leathwick, et al. 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecol. Appl.* 19, 181–197.

Phillips, S. J., R. P. Anderson, M. Dudík, R. E. Schapire, and M. E. Blair. 2017. Opening the black box: an open-source release of Maxent. *Ecography* 40, 887–893.

Ranc, N., L. Santini, C. Rondinini, L. Boitani, F. Poitevin, A. Angerbjörn, et al. 2017. Performance tradeoffs in target-group bias correction for species distribution models. *Ecography* 40, 1076–1087.

Rechsteiner, C., F. Zellweger, A. Gerber, F. T. Breiner, and K. Bollmann. 2017. Remotely sensed forest habitat structures improve regional species conservation. *Remote Sens. Ecol. Conserv.* 3, 247–258.
Rolstad, J. 1988. Autumn habitat of capercaillie in southeastern Norway. J. Wildl. Manag. 747–753.
Rolstad, J., P. Wegge, and B. B. Larsen. 1988. Spacing and habitat use of capercaillie during summer. Can. J. Zool. 66, 670–679.
Roussel, J.-R., and D. Auty. 2016. lidR: Airborne LiDAR Data Manipulation and Visualization for Forestry Applications. R package version 1.0.0. Available from: https://github.com/Jean-Romain/lidR.
Sachot, S. 2002. Viability and Management of an Endangered capercaillie (Tetrao urogallus) Metapopulation. Doctoral dissertation. University of Lausanne.
Sachot, S., N. Perrin, and C. Neet. 2003. Winter habitat selection by two sympatric forest grouse in western Switzerland: implications for conservation. Biol. Conserv. 112, 373–382.
Sachot, S., N. Perrin, and C. Neet. 2006. Viability and management of an endangered capercaillie (Tetrao urogallus) metapopulation in the Jura Mountains, Western Switzerland. Biodivers. Conserv. 15, 2017–2032.
Stähle, G., S. Holm, T. G. Gregoire, T. Gobakken, E. Næsset, and R. Nelson. 2010. Model-based inference for biomass estimation in a LiDAR sample survey in Hedmark County, Norway. Can. J. For. Res. 41, 96–107.
Storch, I. 1993. Habitat selection by capercaillie in summer and autumn: is bilberry important? Oecologia 95, 257–265.
Storch, I. 1994. Habitat and survival of capercaillie Tetrao urogallus nests and broods in the Bavarian Alps. Biol. Conserv. 70, 237–243.
Storch, I. 2002. On spatial resolution in habitat models: can small-scale forest structure explain capercaillie numbers? Conserv. Ecol. 6, 6.
Storch, I. 2007. Grouse, Status Survey and Conservation Action Plan 2006–2010. Gland, Switzerland: IUCN; Fordingbridge, UK: World Pheasant Association.
Teuscher, M., R. Brandl, B. Foerster, T. Hothorn, S. Roesner, and J. Mueller. 2013. Forest inventories are a valuable data source for habitat modelling of forest species: an alternative to remote-sensing data. Forestry 86, 241–253.
Thiel, D., C. Unger, M. Kéry, and L. Jenni. 2007. Selection of night roosts in winter by capercaillie Tetrao urogallus in Central Europe. Wildl. Biol. 13, 73–86.
Thiel, D., S. Jenni-Eiermann, V. Braunisch, R. Palme, and L. Jenni. 2008. Ski tourism affects habitat use and evokes a physiological stress response in capercaillie Tetrao urogallus: a new methodological approach. J. Appl. Ecol. 45, 845–853.
Tiede, D., S. Lang, and B. Maier. 2007. Transferability of a tree-crown delineation approach using region-specific segmentation. Proceedings of XIII Brazilian Remote Sensing Symposium. pp. 21–26.
Vierling, K. T., L. A. Vierling, W. A. Gould, S. Martinuzzi, and R. M. Clawges. 2008. Lidar: shedding new light on habitat characterization and modeling. Front. Ecol. Environ. 6, 90–98.
Zellweger, F., V. Braunisch, A. Baltensweiler, and K. Bollmann. 2013. Remotely sensed forest structural complexity predicts multi species occurrence at the landscape scale. For. Ecol. Manag. 307, 303–312.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. (A) Response curves to the major contributing variable distance do ski runs (female). (B) Response curves to the major contributing variable distance do ski runs (male).