Forest Ecology and Management 430 (2018) 498–508

Contents lists available at ScienceDirect

Forest Ecology and Management

journal homepage: www.elsevier.com/locate/foreco

Usability of citizen science observations together with airborne laser scanning data in determining the habitat preferences of forest birds

L. Mononen[a,b,*], A.-P. Auvinen[a], P. Packalen[c], R. Virkkala[a], R. Valbuena[d], I. Bohlin[e], J. Valkama[f], P. Vihervaara[a]

[a] Finnish Environment Institute (SYKE), Joensuu, Oulu, and Helsinki Offices, P.O. Box 140, 00251 Helsinki, Finland
[b] Department of Geographical and Historical Studies, University of Eastern Finland (UEF), P.O. Box 111, 80101 Joensuu, Finland
[c] School of Forest Sciences, University of Eastern Finland (UEF), P.O. Box 111, 80101 Joensuu, Finland
[d] University of Cambridge, Department of Plant Sciences, Forest Ecology and Conservation, Downing Street, Cambridge CB2 3EA, United Kingdom
[e] Swedish University of Agricultural Sciences (SLU), Department of Forest Resource Management, SE-901 83 Umeå, Sweden
[f] Finnish Museum of Natural History (Luomus), P.O. Box 17, FI-00014, University of Helsinki, Finland

ARTICLE INFO

Keywords:
Citizen Science (CS)
Airborne laser scanning (ALS)
Light detection and ranging (LiDAR)
Forest bird habitats
Habitat modelling

ABSTRACT

Citizens' field observations are increasingly stored in accessible databases, which makes it possible to use them in research. Citizen science (CS) complements the field work that must necessarily be carried out to gain an understanding of any of bird species' ecology. However, CS data holds multiple biases (e.g. presence only data, location error of bird observations, spatial data coverage) that should be paid attention before using the data in scientific research.

The use of Airborne Laser Scanning (ALS) enables investigating forest bird species' habitat preferences in detail and over large areas. In this study the breeding time habitat preferences of 25 forest bird species were investigated by coupling CS observations together with nine forest structure parameters that were computed using ALS data and field plot measurements. Habitat preferences were derived by comparing surroundings of presence-only observations against the full landscape. Also, in order to account for bird observation location errors, we analysed several buffering alternatives.

The results correspond well with the known ecology of the selected forest bird species. The size of a bird species' territory as well as some behavioural traits affecting detectability (song volume, mobility etc.) seemed to determine which bird species' CS data could be analysed with this approach. Especially the habitats of specialised species with small or medium sized territories differed from the whole forest landscape in the light of several forest structure parameters. Further research is needed to tackle issues related to the behaviour of the observers (e.g. birdwatchers' preference for roads) and characteristics of the observed species (e.g. preference for edge habitats), which may be the reasons for few unexpected results.

Our study shows that coupling CS data with ALS yield meaningful results that can be presented with distribution figures easy to understand and, more importantly, that can cover areas larger than what is normally possible by means of purpose-designed research projects. However, the use of CS data requires an understanding of the process of data collection by volunteers. Some of the biases in the data call for further thinking in terms of how the data is collected and analysed.

1. Introduction

Citizen science (CS) involves the collaboration of professionals and non-professionals in scientific research. During the past decades citizen participation has become a common practice in collecting ecological data for environmental monitoring (Conrad and Hilchey, 2011; Dickinson et al., 2012). Due to their detectability and the high level of ornithological expertise among non-professionals, birds are among the species groups of which CS observations hold the most potential to be used in research. There are several well-established procedures for sampling birds that contain elements of CS (e.g. Sullivan et al., 2009; Laaksonen and Lehikoinen, 2013). Observation schemes involving a

* Corresponding author at: Finnish Environment Institute (SYKE), Joensuu, Oulu, and Helsinki Offices, P.O. Box 140, 00251 Helsinki, Finland and Department of Geographical and Historical Studies, University of Eastern Finland (UEF), P.O. Box 111, 80101 Joensuu, Finland.
E-mail address: laura.mononen@ymparisto.fi (L. Mononen).

https://doi.org/10.1016/j.foreco.2018.08.040
Received 12 June 2018; Received in revised form 14 August 2018; Accepted 21 August 2018

strong CS component have been used to study the timing of migration (Jonzén et al., 2006; Saino et al., 2010; Lehikoinen et al., 2013). However, few if any previous attempts have been made to use CS data in connection with remote sensing data in studying the habitat preferences of forest birds.

Information on habitat characteristics has improved both in terms of accuracy and spatial extent over the past century along with the development of remote sensing (RS) techniques (Kerr and Ostrovsky, 2003; Cohen and Goward, 2004). In particular, habitat and species distribution modelling have benefited a great deal from the development of RS, which supplements or sometimes even replaces traditional field work (Pettorelli et al., 2014). Airborne Laser Scanning (ALS) provides three-dimensional information which greatly advances the spatial analysis of habitat structures (e.g. Lefsky et al., 2002; Hill and Thomson, 2005; Davies and Asner, 2014; Valbuena et al., 2017) and helps detecting changing patterns of habitat use in a changing climate (Melin et al., 2014). Since human activity has a strong effect on the structural complexity of forests (e.g. Brokaw and Lent, 1999), ALS derived information has been acknowledged valuable for biodiversity assessments (Vierling et al., 2008). ALS parameters assist in the detection of those species that depend on or benefit from the structural heterogeneity of canopy structure (Goetz et al., 2007; Vierling et al., 2008; Palminteri et al., 2012).

Birds are a species group that respond to environmental changes relatively promptly (Barbet-Massin et al., 2012; Frishkoff et al., 2014; Virkkala and Lehikoinen, 2014) which makes them, along with several other characteristics (including ecological traits such as position in a food chain and non-ecological traits such as popularity), good indicators of biodiversity (Butchart et al., 2010; Gregory and van Strien, 2010). Traditionally, forest bird-habitat relations have been studied in the field by measuring certain habitat variables (e.g. tree species, height, diameter etc.) and connecting these measurements with bird observations (see e.g. MacArthur and MacArthur, 1961, Wiens, 1989a and references therein). This kind of analysis produces detailed information of species' habitat selection, but can usually be carried out only over relatively small areas. Importantly, there is also stochasticity in species occurrence, which makes difficult to extrapolate results of bird-habitat studies conducted at fine scales (Wiens et al., 1987; Haila et al., 1996; Virkkala and Rajasärkkä, 2006). Bird-habitat relationships must therefore be studied on different spatial scales (Wiens et al., 1987; Wiens, 1989b). On the other hand, ALS and other RS datasets connected with species observations can be used over large areas, and thus they are not susceptible to small scale variation in the occurrence of bird species. In fact, the Group on Earth Observations Biodiversity Observation Network (GEO BON) has identified the potential of remote sensing and in situ data combinations to contribute for extensive and cost-efficient biodiversity monitoring (GEO BON, 2015). Using high quality CS observation could greatly advance this goal.

Several previous studies have proven the capability of ALS derived parameters to predict the species-richness of habitats (reviewed in Simonson et al., 2014) and, more recently, the differentiation of diversity among habitats (e.g. Zellweger et al., 2017). Further, some studies have successfully examined the specific species-habitat relations by using ALS, but these have focused only on a few habitat indicators or species or both (e.g. Graf et al., 2009; Goetz et al., 2010; Hagar et al., 2014; Melin et al., 2016). To date, only few papers have examined the habitat preferences of multiple forest song birds with ALS. E.g. in Hinsley et al. (2009) and Müller et al. (2009) the observation data was surveyed by professional ecologists. To our knowledge, no previous studies have examined the use of CS data as rigorously (but see Vihervaara et al., 2015). In this study the relatively high number of species (25) was achieved by using CS data – collecting such a large dataset over such a large area would have been out of our reach by means of a purpose-designed research project.

In this study we explore the extent to which CS data can be used to assess the habitat preferences of forest birds, and identify potential pitfalls when doing so. We use positioned observations from 25 forest bird species and nine ALS derived parameters to: (I) Explore whether CS observations in connection with ALS based forest structure parameters can provide information that is in line with the known ecological characteristics (e.g. habitat preferences) of the bird species included in the study, (II) Investigate which bird species' habitats could be best modelled by using the combination of CS and ALS data, and (III) Examine which forest structure parameters are most suitable for predicting bird species' habitats in this connection. In order to facilitate the replicability of the method, the low-density ALS data were used as today they are typically acquired at national scale. However, we used a selection of ALS derived forest parameter layers that can be computed with field plot measurements to examine whether they can offer more detailed or complementary information for research. The potential applications of our study relate not only to the field of animal ecology, but can also help in determining where to focus conservation activities (Rose et al., 2015). Potentially, the combination of CS observations and ALS data could enable us to cover areas as large as administrative regions or even nations, and, in the future, also to model and predict the occurrence of species of conservation interest.

2. Materials and methods

2.1. Study area

The study area is located in Southern Finland in the Lake Vanajavesi catchment area of 3000 km^2 (Fig. 1). The area belongs to the southern boreal taiga vegetation zone and the landscape is dominated by boreal forests. The majority of the forests are commercially managed. Large lakes and small rivers, agricultural areas and wetlands are also typical in the area. The study area was selected based on the available ALS data and CS bird observations.

2.2. Species observations

In our study, we included species of conservation concern, such as the European Union's Birds Directive Species (Annex I) and redlisted species in Finland, species preferring old-growth or mature forests, and species of herb-rich, lush, and deciduous forests (see Vihervaara et al., 2015). We also included species occurring in boreal agricultural-forest mosaics. The forest bird observation data were acquired from two sources; (i) Bird Atlas data from the database of the Finnish Museum of Natural History and (ii) faunistic observations from the Tiira database maintained by BirdLife Finland (Valkama et al., 2011). Both of these data have been collected by mostly non-professional volunteers and, although the Bird Atlas was more goal-oriented and structured, can be described with good reason as citizen science. All observations were recorded in years 2006–2012. In the Tiira database bird observations are reported at a resolution of 1 x 1 m, and we included all bird observations in the Atlas with a resolution of 10 × 10 m or less. In practice the resolution of bird observations in the Tiira database is slightly worse, so that these data sets are compatible in accuracy. For each species only the observations during their known breeding time were included in the data. In the end, 25 bird species were included and the numbers of observations per species vary between 31 and 355 (Table 1).

2.3. ALS data

ALS data were collected on May–June 2008 using Optech ALTM GEMINI laser scanning system. The ALS data point is offset at most by four years from the bird data (2006–2012). In a previous study a gap of this size between the acquisition of ALS and CS data was found to be of marginal impact on the results (Vierling et al., 2014). The area was measured from an altitude of 2000 m above ground level using half angle of 20°. This resulted in a swath width of 1 450 m and a nominal