Forest Ecology and Management 410 (2018) 104–113

Contents lists available at ScienceDirect

Forest Ecology and Management

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

Optimization of forest sampling strategies for woody plant species distribution modelling at the landscape scale

Rubén G. Mateo*, Aitor Gastón, María José Aroca-Fernández, Santiago Saura[1], Juan Ignacio García-Viñas

MONTES (ETSI de Montes, Forestal y del Medio Natural), Universidad Politécnica de Madrid, Ciudad Universitaria s/n, 28040 Madrid, Spain

ARTICLE INFO

Keywords:
Ecological modelling
Forest landscapes
Forest management tools
Sampling resource optimization
Sampling design
Forest inventory

ABSTRACT

Forest management increasingly needs the support of species distribution models (SDMs). However, different challenges remain to be addressed before the practical use and generalization of these models in the design of management measures in forest ecosystems. Due to the limited resources that are typically available for forest management, including the forest inventory phase, it is necessary to optimize the sampling approaches. Opportunistic sampling may be one strategy to reduce sampling costs, but the accuracy and suitability of the assessments derived from this sampling remain yet poorly addressed, particularly in forest landscapes. On the other hand, different forestry applications require landscape-wide estimates but still with fine resolution (ideally meters) as to guide management interventions. We, here, assess in detail the performance of the classical regular (systematic) sampling strategy and three different opportunistic samplings approaches along roads and tracks in a forest-dominated Biosphere Reserve (≈15,000 ha) in central Spain. We use specifically gathered field data in different inventories with sampling intensities of about 1–2 plots/km^2. We compare, for each sampling strategy, the resultant consensus species distribution models for 28 woody plant species (trees and shrubs) developed at a spatial resolution of 25 m. We found that SDMs were reliable (AUC $>$ 0.75) for 20 species out of 28 using either an opportunistic or/and systematic sampling. In general, opportunistic sampling was more efficient than systematic sampling, resulting in SDMs with a comparable accuracy for a lower inventory cost or in more accurate SDMs for the same sampling effort. This was mainly because the four sampling strategies adequately captured the environmental variability of the study area, and because plots in the opportunistic sampling were located at a sufficient distance from tracks to avoid potential edge effects. The minimum sample size and the species ecology should, however, be evaluated with caution in each case. We conclude that opportunistic sampling along tracks may be a practical and cost-effective option for SDMs in forest landscapes, as long as the density and spatial arrangement of tracks are sufficient to cover the full range of environmental conditions. This prerequisite could be evaluated, before the inventory phase, using available statistical approaches and spatial layers, potentially allowing for a remarkable saving of sampling resources in forest management planning. Species distribution models calibrated at the landscape extent with fine resolution could hence become ever more powerful and cost-effective tools for forest management and planning in both planted and natural forests.

1. Introduction

Species distribution models (SDMs) are powerful methods in forest management and plant species ecology research (Henderson et al., 2014). Primarily, they can deliver potential species distribution and vegetation maps at different scales (Guisan et al., 2017), which are essential in forest management and conservation planning. SDMs support the understanding, for example, of plant biodiversity patterns (Mateo et al., 2016; D'Amen et al., 2017), potential effects of invasive plant species (Petitpierre et al., 2012; Mateo et al., 2015) and invasive forest pathogens (Václavík et al., 2010), climate change effects on plant species (Thuiller et al., 2008; Engler et al., 2011) and forest distribution (Moreno-Fernández et al., 2016; van der Maaten et al., 2017), endangered species management (Rovzar et al., 2016), ecological forest

*Abbreviations:*SDMs, species distribution models

* Corresponding author.
E-mail addresses: rubeng.mateo@gmail.com (R.G. Mateo), aitor.gaston@upm.es (A. Gastón), pepa.aroca@gmail.com (M.J. Aroca-Fernández), santiago.saura@upm.es (S. Saura), juanignacio.garcia@upm.es (J.I. García-Viñas).
[1] Current address: European Commission, Joint Research Centre (JRC), Directorate D: Sustainable Resources, Via E. Fermi 2749, I-21027 Ispra, VA, Italy.

https://doi.org/10.1016/j.foreco.2017.12.046
Received 9 October 2017; Received in revised form 22 December 2017; Accepted 28 December 2017

restoration (Gastón et al., 2014), etc.

Different methodological aspects can affect the performance of these models (Guisan and Zimmermann, 2000; Mateo et al., 2011), such as the minimum sample size required (Wisz et al., 2008; Mateo et al., 2010b), parameterization (Moreno-Amat et al., 2015), or pseudo-absences generation (Wisz and Guisan, 2009; Mateo et al., 2015). Significant advances have been made in the last years on those methodological questions (Guisan et al., 2017). However, two of these essential topics need more exploration before the application of these models in forest management can be generalized. First, SDMs are usually fitted on data achieved through sampling strategies not expressly planned for modelling (Loiselle et al., 2008) or management goals (opportunistic data). On the other hand, classical forest inventories are commonly done through regular-systematic sampling strategies (Mello et al., 2015). Are SDMs developed with opportunistic or systematic sampled data comparable and reliable for forest management purposes? Second, the bulk of SDMs studies are defined over large areas (regional and continental extent) with a coarse resolution (1–50 km). This is still far from the use of SDMs at the landscape extent and fine grain resolution (meters), which is required to support forest management interventions (Gastón and García-Viñas, 2010). Indeed, numerous forestry applications rely on the correct evaluation of models developed at the landscape scale, such as the selection of species for reforestation (Gastón et al., 2014); accurate evaluation of climate change effects (Randin et al., 2009); management of plant diversity patterns (Dubuis et al., 2011) and endangered species (Rovzar et al., 2016).

Databases derived from opportunistic sampling strategies (Graham et al., 2004; Brotons et al., 2007) are classically achieved from natural history collections (Araújo and Williams, 2000). These databases have several benefits: quantity, accessibility, and extensive geographic and temporal scales (Garcillán and Ezcurra, 2011). Conversely, the primary constraint of these databases is that they do not derive from specific prearranged samplings. The data are usually surveyed near tracks or attractive botanical zones (Schulman et al., 2007). Consequently, opportunistic databases are low-cost; nevertheless they could be spatially biased (Jones, 2011). For example, a sampling strategy following tracks could be related to spatial bias, which can lead to climatic or ecological bias (Hortal et al., 2008), which in turn could affect the SDMs performance (Hirzel and Guisan, 2002). However, it is also possible that a sampling strategy following tracks adequately covers the study area (no spatial bias), thus containing the climatic variability and the environmental conditions wholly. Different authors (Kadmon et al., 2004; Loiselle et al., 2008) deal with this issue at coarse resolutions and regional scale, and they concluded that it is achievable to obtain reliable SDMs from opportunistic sampling strategies. Natural history collections are also unlikely to have the spatial resolution and density of observations required to conduct forest management and planning. Therefore, a more detailed sampling strategy at the landscape scale should be defined to guide forestry applications.

Species distribution modelling aimed at small geographic regions and a fine resolution was infrequently conducted until recently, but nowadays many research teams are developing SDMs in small areas and fine resolutions (D'Amen et al., 2015, 2017; Scherrer et al., 2017). This is primarily because of the new accessibility to high resolution (meters) GIS free data, such as digital elevation models, forest maps, remote sensing or LIDAR images. Nevertheless, the factors encompassed in species distribution modelling are not compatible among different scales (Mateo et al., 2017) and the results from previous methodological studies conducted at larger areas and coarser resolutions (see above) might not be consistent with this local scale and the related forestry applications. Therefore, before the generalization of the application of landscape species distribution modelling to forest and woodland management, additional methodological studies and insights are needed.

Gathering data at local-landscape level is, however, inherently costly in time and resources. Considering the limited resources that governments or management agencies can usually allocate to implement forest and woodland management, resources should be optimized when capturing the data in the field and during the modelling process. In this regard, comparisons of opportunistic low-cost sampling strategies with systematic (regular) sampling are needed. Systematic sampling has been shown to be one of the most precise sampling procedures (Hirzel and Guisan, 2002), but the cost and time required for both sampling strategies are very different. Here, we investigate and compare, for 28 woody species, the reliability of SDMs obtained under two different sampling strategies at the landscape scale: systematic (more resource demanding) and opportunistic along tracks and roads (less resource demanding). The initial hypothesis is that systematic sampling could be necessary to more accurately represent fine-scale species distributions in the forest landscape (Edwards et al., 2006; Braunisch and Suchant, 2010). However, we hypothesize too that opportunistic sampling could derive reliable SDMs if the sampling intensity (minimum sample size, Mateo et al., 2010b) is appropriate, and if the accessible areas by roads and tracks do not include significant spatial biases, so that they capture the range of environmental conditions in the study area.

2. Materials and methods

2.1. Study area

The study was developed at the *Sierra del Rincón* Biosphere Reserve (41°03′N 3°29′W, central Spain). The extension of the area is 15,231 ha, elevation ranges from 670 to 2,200 m, mean temperature ranges from 4.9 to 13.2 °C, and rainfall ranges from 650 to 1,300 mm/year. Soils are derived from siliceous substrates, mostly from gneiss, schists, quartzites, and slates. A mosaic of forests, shrublands and grasslands cover the Reserve. Forests are mainly dominated by oaks (*Quercus pyrenaica*) and pines (*Pinus sylvestris*, *P. pinaster*, *P. nigra*), the latter as a consequence of the afforestation programs of the 1950s. The Reserve includes the *Montejo* beech forest, which is the SW distribution limit of *Fagus sylvatica*.

2.2. Sampling strategies and distribution data

We carried out and compared two different sampling strategies: (1) a systematic sampling, and (2) an opportunistic sampling along tracks and roads. All sampling plots were positioned precisely using a GPS, and presence/absence of all woody (trees and shrubs) species was recorded at woodland (forest and shrublands) areas in circular plots with a ten meters radius. Plots at pastures, inaccessible and non-natural vegetation areas (rocks, crops, villages, etc.) were discarded. Finally, 28 woody species (nine trees and 19 shrubs, see Appendix A) present in a minimum of five plots were used for modelling. They represent typical genus of Mediterranean plant species, for example: *Adenocarpus* sp., *Calluna* sp., *Cistus* sp., *Crataegus* sp., *Cytisus* sp., *Erica* sp., *Fraxinus* sp., *Genista* sp., *Halimium* sp., *Helichrysum* sp., *Ilex* sp., *Juniperus* sp., *Lavandula* sp., *Pinus* sp., *Prunus* sp., *Quercus* sp., *Salix* sp., *Santolina* sp., *Sorbus* sp., and *Thymus* sp.

The systematic sampling was designed *a priori* following a regular lattice over the complete study area with vertices separated by 1,000 m, coinciding with those of the 4th Spanish National Forest Inventory. A total of 132 plots were sampled (Fig. 1a). It was an exhaustive sampling routine that involved two persons performing fieldwork for approximately seven weeks. This sampling strategy requires a more significant time of displacement walking, in some cases through steep terrain, in order to arrive at the plots than an opportunistic sampling.

An opportunistic sampling along tracks and roads was achieved independently; plots were established at 1,000 m road or tracks intervals, and the starting point at every road-track was set up at 500 m of the road-track sector. However, unlike systematic sampling, the