Global Ecology and Conservation 12 (2017) 215–226

Contents lists available at ScienceDirect

Global Ecology and Conservation

journal homepage: http://www.elsevier.com/locate/gecco

Original Research Article

Predicting species distribution combining multi-scale drivers

Alice Fournier [a, *], Morgane Barbet-Massin [a], Quentin Rome [b], Franck Courchamp [a]

[a] Ecologie, Systématique et Evolution, Univ. Paris-Sud, CNRS, AgroParisTech, Université Paris-Saclay, Paris, 91400 Orsay, France
[b] UMS 2006 Patrimoine Naturel — AFB, CNRS, MNHN — Muséum national d'Histoire naturelle, CP50, 57 Rue Cuvier, 75235, Paris Cedex 05, France

ARTICLE INFO

Article history:
Received 28 September 2017
Received in revised form 9 November 2017
Accepted 10 November 2017
Available online 6 December 2017

Keywords:
Climatic variables
Spatial scale
Environmental filtering
Habitat
SDM
Yellow-legged hornet

ABSTRACT

Species Distribution Models (SDMs) are often used to predict the potential range of invasive species. Unfortunately, most studies do not evaluate variables relevance before selecting them to fit their models. Moreover, multiple variables such as climate and land use may drive species distribution at different spatial scales but most studies either use a single type of drivers, or combine multiple types without respecting their operating scale. We propose a three steps framework to overcome this limitation. First, use SDMs to select the most relevant climatic variables to predict a given species distribution, at continental scale. Then, characterize the species-habitat relationships, at a local scale, to produce species and area specific habitat filters. Finally, combine both information, each obtained at a relevant scale, to refine climatic predictions according to habitat suitability. We illustrate this framework with 14,794 Asian hornet (*Vespa velutina nigrithorax*) records. We show that integrating multiple drivers, while still respecting their scale of effect, produced a potential range 55.9% smaller than that predicted using the climatic model alone, suggesting a systematic overestimation in many published predictions. This general framework illustrated by a well-documented invasive species is applicable to other taxa and scenarios of future climate and land-cover changes.

1. Introduction

Species distribution models (SDMs) are regularly used to generate predictions of species presence; they rely on correlations between environmental variables and geo-localized species records to understand the environmental variables that drive species presences (Blach-Overgaard et al., 2010) and delineate potential species distributions (Araújo and Guisan, 2006; Araújo and Peterson, 2012). An extensive number of environmental datasets are available nowadays for fitting SDMs; but only a limited number of them should be included when running SDMs. Indeed, although increasing the number of predictors increases the chance of having ecologically relevant ones, it also inflates the risk of overfitting the model (Merow et al., 2014) and of collinearity issues between variables (Dormann et al., 2013a). Restricting the number of variables and choosing only the most appropriate ones for a species is thus crucial to maximize the performance of SDMs and the accuracy of the predictions (Araújo and Guisan, 2006; Araújo and Peterson, 2012; Barbet-Massin and Jetz, 2014; Braunisch et al., 2013). Ideally, this choice should rely on expert knowledge concerning the ecological requirements of the species, but such knowledge is

* Corresponding author.
E-mail address: alice.fournier@u-psud.fr (A. Fournier).

https://doi.org/10.1016/j.gecco.2017.11.002
2351-9894/

hardly ever available. Predictor selection thus remains a key challenge when running SDMs (Araújo and Guisan, 2006), and despite its proven and important effect on the predictions, has received too little attention (Austin and Van Niel, 2011; Syphard and Franklin, 2009; Tulloch et al., 2016).

Another challenge of SDMs, linked to the selection of the most appropriate variables, is to combine multiple variables that may drive species distribution at different spatial scales. Niche modelling has always lacked unifying theories and methods to bring together multi-scale drivers, which reduces their accuracy and their appropriateness for conservation planning. Notably, studies integrating climatic and land use variables at different scales remain extremely rare, despite evidences of their necessity (Sirami et al., 2016). Most biodiversity scenarios focus on climatic models and fail to integrate other environmental filters, which also have significant impacts on species distributions (Sirami et al., 2016; Titeux et al., 2016), or mix all predictors at the same resolution (Bucklin et al., 2014; Gallardo et al., 2015). These filters may be related to climate, topography, primary production or land use (Milbau et al., 2009). Furthermore, many studies have now come to the agreement that species–environment relationships are strongly scale dependent; species presence results from an interplay between climate, which governs their distributions at continental scales (Blach-Overgaard et al., 2010), and for instance habitat, which drives species' occupancy at finer spatial resolution (Luoto et al., 2007; Monceau and Thiéry, 2017; Virkkala et al., 2005). Each of these drivers, acting as a filter that shapes the species distribution as a special spatial scale, must thus be identified and included in the model at the appropriate resolution (Cabra-Rivas et al., 2015; Luoto et al., 2007; Milbau et al., 2009; Pearson et al., 2004, 2003). Very few studies have provided theoretical bases to bring together climate and habitats predictors in a hierarchical manner (Kelly et al., 2014; Milbau et al., 2009; Pearson et al., 2003; Sirami et al., 2016). They represent valuable theoretical starting points, but methods carefully selecting the regional habitat filters to be considered and put this theory into practice are needed to improve the accuracy of climate-driven models (Thuiller et al., 2004; Zhu et al., 2017). Therefore, a key challenge now is to develop a practical method to integrate multi-scale predictors and capture more accurately their environmental niche (Virkkala et al., 2005).

We present here a framework to select the most relevant climatic variables, at a global scale, build species-specific habitat suitability filters, at a local scale, and combine both information to produce refined suitability maps. We illustrate this framework with the case of the invasive Asian hornet, *Vespa velutina nigrithorax*. The crucial step of variable selection has never been addressed to predict this species potential range, and previous predictions have been obtained with climatic variables only. The Asian hornet arrived accidentally in France from China in 2004 (Arca et al., 2015), and since then has invaded almost the whole of France and other European countries (Spain, Portugal, Italy, Belgium Germany and UK). Due to its predatory behavior on insects, and particularly honey bees (Monceau et al., 2014), there are great concerns about its potential impacts on the native biodiversity (Choi et al., 2012), on the beekeeping industry and on pollination services overall (Monceau et al., 2014). It is necessary to develop tools to better predict its future invasion range and help stop its spread and reduce its impacts efficiently. The method is generalizable to any other taxa, and can be used to combine as many environmental layers as needed, as it provides tools to deal with both continuous and categorical layers.

2. Materials & methods

2.1. Distributional data

We used the GPS records from the INPN biodiversity database (http://inpn.mnhn.fr), maintained by the French National Natural History Museum, and based on a participative science program (http://frelonasiatique.mnhn.fr), as in Barbet-Massin et al. (2013). The database totaled to 14,794 records of Asian hornet colonies, from the invaded range (France, Italy, Germany, Spain & Portugal, Belgium), spanning from 2004 to 2016. We did not include presence data from the native range for two reasons. Firstly, this invasion results from a single introduction event (Arca et al., 2015), of one female only, that gave rise to the whole European population of Asian hornets. A single organism cannot encompass the whole population genetic and phenotypic diversities. Furthermore, altered species– climate relationships during invasion are recurrent for insects, due to their ability to respond quickly to novel environments (Hill et al., 2017). The niche conservation assumption is highly unlikely to be met and the niche overlap between native and invasive ranges expected to be poor (Early and Sax, 2014; Medley, 2010). In this context, a conservative framework is favored, which consists in using only invasive occurrences to build the model. A recent study on the Asian hornet showed that the predictive accuracy of the SDM was significantly better when models were calibrated with invasive data only, excluding native data (Barbet-Massin et al., unpublished results). Secondly, there were only 68 occurrences available from its native range, some of which of uncertain and unequal quality compared to the invasive range data. For these reasons, we preferred to rely on invasion occurrences only.

2.2. Environmental data

We used a set of 19 climatic variables (averaged from 1950 to 2000) available from the worldclim database (http://www.worldclim.org/) at 2.5 arc min (~4 km) resolution (Hijmans et al., 2005). These variables represent a combination of means, extremes, variability and seasonality of temperature and precipitation data that are known to influence species distribution (Root et al., 2003).

We used the CORINE (Coordination of Information on the Environment) Land Cover 2006 dataset to study the suitability of each habitat for the hornet (The European Environment Agency (EEA), 2010). This dataset is characterized by its high-spatial