Original article

Predicting species richness and distribution ranges of centipedes at the northern edge of Europe

Elisavet Georgopoulou [a, *], Per Djursvoll [b], Stylianos M. Simaiakis [c]

[a] Geological-Paleontological Department, Natural History Museum Vienna, Burgring 7, AT-1010 Wien, Austria
[b] University Museum of Bergen, PO Box 7800, NO-5020 Bergen, Norway
[c] Natural History Museum of Crete, University of Crete, Knossos Avenue, GR-71409 Heraklion, Crete, Greece

ARTICLE INFO

ABSTRACT

In recent decades, interest in understanding species distributions and exploring processes that shape species diversity has increased, leading to the development of advanced methods for the exploitation of occurrence data for analytical and ecological purposes. Here, with the use of georeferenced centipede data, we explore the importance and contribution of bioclimatic variables and land cover, and predict distribution ranges and potential hotspots in Norway. We used a maximum entropy analysis (Maxent) to model species' distributions, aiming at exploring centres of distribution, latitudinal spans and northern range boundaries of centipedes in Norway. The performance of all Maxent models was better than random with average test area under the curve (AUC) values above 0.893 and True Skill Statistic (TSS) values above 0.593. Our results showed a highly significant latitudinal gradient of increased species richness in southern grid-cells. Mean temperatures of warmest and coldest quarters explained much of the potential distribution of species. Predictive modelling analyses revealed that south-eastern Norway and the Atlantic coast in the west (inclusive of the major fjord system of Sognefjord), are local biodiversity hotspots with regard to high predictive species co-occurrence. We conclude that our predicted northward shifts of centipedes' distributions in Norway are likely a result of post-glacial recolonization patterns, species' ecological requirements and dispersal abilities.

© 2016 Elsevier Masson SAS. All rights reserved.

1. Introduction

In the last few years a range of species distribution modelling methods (see Elith et al., 2006 and references therein) have been used extensively in biogeography, evolutionary biology and conservation ecology (see Elith et al., 2006 for review). Species distribution models (SDMs) aim to promote the exploration of distribution patterns (e.g. Raes et al., 2009), assess the impact of environmental factors on species occurrence (e.g. Wollan et al., 2008), produce a spatially explicit representation of habitat suitability for certain biotas (see Guisan and Zimmermann, 2000; Elith et al., 2006 for reviews), and predict species distributions in the face of climate change and other ongoing environmental changes (e.g. Eskildsen et al., 2013). For conservation planning in particular,

SDMs require fine georeferenced data often found in museum and herbarium collections (Garcia-Milagros and Funk, 2010), and Geographical Information System (GIS) tools. One of the most popular distribution modelling algorithms to analyse empirical data is the maximum entropy (Maxent) model, a reliable, general-purpose method for predicting the potential distribution of species from presence-only records (Phillips et al., 2006; Yackulic et al., 2013). It has proven its robustness in comparison to other methods (for further details, see Elith et al., 2006) even when using a low number of sampling records (Hernandez et al., 2006). Maxent has been used in a wide variety of fields such as historical biogeography (e.g. Cordellier and Pfenninger, 2009), ecology (e.g. Graham and Hijmans, 2006; Wollan et al., 2008), and evolutionary and conservation biology (e.g. Tinoco et al., 2009; Young et al., 2009; Kassara et al., 2012).

Geographic range sizes are crucial in understanding ecological and evolutionary processes, such as migration routes, species ecological tolerance and diversity gradients (Tomašových et al., 2015 and references therein). In particular, range dynamics on the northern margins of a species distribution can be quite informative

* Corresponding author.
E-mail addresses: elisavet.georgopoulou@nhm-wien.ac.at, georgelisavet@yahoo.gr (E. Georgopoulou), per.djursvoll@um.uib.no (P. Djursvoll), ssimaiakis@yahoo.com (S.M. Simaiakis).

of species responses to climate change (e.g. Devictor et al., 2012) and can also provide insights to past distributions (e.g. Nogués-Bravo, 2009; Svenning et al., 2008). For example, the evaluation of the northern refuge hypothesis using SDMs for 22 widespread European tree species during the Last Glacial Maximum (LGM) supports the possibility that trees were more widespread in northern Europe during the LGM than previously supposed (Svenning et al., 2008). Nevertheless, the predictions of SDMs should be treated with caution when used to predict range shifts under changing climate at the expanding northern edges of a species' distribution (Eskildsen et al., 2013).

In recent years, several studies on species distribution modelling have been carried out for many organisms (e.g. Wollan et al., 2008; Eskildsen et al., 2013; Kramer-Schadt et al., 2013; Patrão et al., 2015), but, so far, not for centipedes. Centipedes are among the most important soil organisms with some 3100 species worldwide, belonging to 396 genera, 23 families and five orders (Minelli et al., 2006). They prey on a variety of sizes of invertebrates and vertebrates by means of their powerful poison-claws, and occur in habitats such as woodland leaf litter, decaying wood, hedgerows and compost heaps (Bonato and Zapparoli, 2011). Centipedes show a large variety of dispersal abilities: large scolopendromorphs, lithobiomorphs and scutigeromorphs are highly vagile, whereas smaller in size lithobiomorphs and geophilomorphs, which are strongly connected to humid and subterranean soil environment, are not. As a result of their (1) wide geographical diversification, especially in Europe, (2) leading role as soil predators in different environments (see Voigtländer, 2011 for review), and (3) moderate dispersal abilities, centipedes are intriguing model organisms for studying historical and analytical biogeography (Giribet and Edgecombe, 2006; Edgecombe and Giribet, 2007; Vélez et al., 2012; Simaiakis and Strona, 2015).

Norway covers about 13° latitude spanning 1800 km between latitudes 58° and 71° N (up to 81° N when including the island of Svalbard), and longitudes 4° and 32° E. It has a total coastline of approximately 57,000 km (Ramberg et al., 2008), a relatively complex topography and a great variety in elevation over short distances (Moen, 1999). It also has an average elevation of 500 m with many peaks over 1000 m, the highest at 2469 m (Galdhøpiggen, Oppland). Norway was covered with ice during the last ice age (Moen, 1999). Glacial erosion and deposition by ice sheets and glaciers have played a primary role to the creation of the mainland. Land began to emerge from the edge of the melting ice sheet 12,000 years ago, ca 9000 years after the end of the LGM; thus the flora and fauna are certainly less than 12,000 years old (but see Parducci et al., 2012). Because of its broken topography and elongate shape, Norway is geographically very diverse (Moen, 1999). The afore-mentioned features, as well as the temperate waters of the Gulf Stream, strongly impact the climate conditions of the country with a considerable temperature gradient from north to south, determining patterns and processes of species distributions. Furthermore, its location at the northernmost periphery of Europe makes it the distribution border of several species (Andersson et al., 2005, 2008, 2013) and of great interest when attempting to investigate biogeographical patterns. Norway acts as a "natural laboratory" (Ramberg et al., 2008) and provides a well-documented environment (it currently supplies more than 13 million species occurrence records to the GBIF Network - www.gbif.org/country/NO/publishing) that render it an ideal system for studying a wide spectrum of ecological topics, such as the role of climate and history on species ranges (see also Simaiakis et al., 2013 and references therein).

In this study the main objectives addressed are: (1) to identify the primary bioclimatic parameters and land cover types determining centipede species distributions in Norway, (2) to model and evaluate predicted species distributions based on presence-only records with emphasis on latitudinal range, centres of distribution and northern range boundaries of species in the periphery of Europe, and (3) to reveal potential biodiversity hotspots of continental importance within Norway.

2. Materials and methods

2.1. The data

A total of 1684 georeferenced centipede species records from Norway were extracted from the Norwegian Biodiversity Information Centre (Artsdatabanken, http://www.artsdatabanken.no/). Species records were not uncritically employed. We only considered species records with a sound taxonomic identification and excluded doubtful records. Such records include occurrences in greenhouses or compost sites, where species were likely transported by humans (e.g. *Lithobius lapidicola*). In addition, species names were taxonomically updated (e.g. *Geophilus insculptus* to *Geophilus alpinus*) following Minelli et al. (2006). All species selected for modelling were represented in the study area by a minimum of 10 unique occurrence records. Duplicate presences within a grid-cell of ca 1 km^2 and presences missing environmental data were automatically removed by Maxent. In total, 20 out of 25 known species from Norway with a sum of 1321 unique presences, ranging from 10 to 466 records per species were modelled (Fig. 1, see also Supplementary material Appendix A Table S1). All samples are housed in natural history museum collections (i.e., University Museum of Bergen, Tromsø University Museum, BioFokus).

Initially 20 environmental predictors were selected to model the species distributions. All bioclimatic variables (from bio1 to bio19) were downloaded from the WorldClim set of global climate layers (http://www.worldclim.org), at a spatial resolution of approximately 1 km^2, to capture the highest degree of environmental variability according to Hijmans et al. (2005). CORINE land cover layer was obtained from the European Environment Agency (2006) in a spatial resolution of 100 m^2. The land cover layer was subsequently resampled to the resolution of the bioclimatic layers. All the environmental layers were processed and clipped to the geographic extent of study area i.e., mainland Norway, in ArcMap v.10 (ESRI, 2011).

2.2. Statistical analyses

Prior to constructing Maxent models, the variance inflation factor (VIF) was used for collinearity control among the 19 bioclimatic variables. Explanatory variables with VIF values greater than 10 usually indicate strong correlations between the variables (Quinn and Keough, 2002) and, in general, are excluded from the model-building process. Variables were not arbitrarily included in the models, but were primarily chosen based on their importance on the biology of centipedes (e.g. reproduction and hibernation; see Lewis, 1981; Vedel et al., 2008; Voigtländer, 2011; Simaiakis et al., 2013). After detecting a multicollinearity effect among the selected climatic variables, the number of predictors was reduced to seven (lc: land cover, bio2: mean diurnal range, bio8: mean temperature of wettest quarter, bio9: mean temperature of driest quarter, bio10: mean temperature of warmest quarter, bio11: mean temperature of coldest quarter, bio18: precipitation of warmest quarter; abbreviations follow WorldClim code names). VIF was calculated using the R package HH v. 3.1–14 (Heiberger, 2015).

We also tested for the presence of spatial autocorrelation in our dataset. Spatial autocorrelation indicates lack of independence among observations (Legendre and Legendre, 1998) which in turn can lead to inaccurate model predictions by introducing false