GfÖ
GfÖ Ecological Society of Germany, Austria and Switzerland
Basic and Applied Ecology 20 (2017) 75–85
Basic and Applied Ecology
www.elsevier.com/locate/baae

Using citizen science data to estimate climatic niches and species distributions

Patrícia Tiago[a,b,*], Henrique M. Pereira[b,c,d], César Capinha[b,e]

[a]*CE3C, Centre for Ecology, Evolution and Environmental Changes, Faculdade de Ciências da Universidade de Lisboa, 1749-016 Lisbon, Portugal*
[b]*Cátedra Infraestruturas de Portugal-Biodiversidade, CIBIO/InBIO, Universidade do Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal*
[c]*German Centre for Integrative Biodiversity Research (iDiv), Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany*
[d]*Institute of Biology, Martin Luther University Halle-Wittenberg, Am Kirchtor 1, 06108 Halle, Saale, Germany*
[e]*Global Health and Tropical Medicine (GHTM), Instituto de Higiene e Medicina Tropical (IHMT), Universidade Nova de Lisboa (UNL), Rua da Junqueira 100, 1349-008 Lisboa, Portugal*

Received 20 August 2016; received in revised form 31 March 2017; accepted 1 April 2017
Available online 8 April 2017

Abstract

Opportunistic citizen data documenting species observations – i.e. observations collected by citizens in a non-standardized way – is becoming increasingly available. In the absence of scientific observations, this data may be a viable alternative for a number of research questions. Here we test the ability of opportunistic species records to provide predictions of the realized distribution of species and if species attributes can act as indicators of the reliability and completeness of these data. We use data for 39 reptile and amphibian species across mainland Portugal as a case study. We collected distribution data from two independent sources: a national citizen-science project and a scientific. We measure and compare the climatic niche width of the species as represented by each of the two data sources. Generalized linear mixed models (GLMM) were used to relate a set of response variables describing the species' morphology, life-history, communication, type of locomotion, habitat and geographic distribution, to observed differences in niche widths. We also performed species distribution models (SDMs) for each of the two types of data using generalized additive models. We found that 12 species had more than 50% of their climate niche covered by citizen science data. Results from GLMMs suggested that the number of grid cells in which a species occurs and its use of forest habitat were positively related to the comprehensiveness of the sampling of climatic niches by citizen science data. Variation in the p of SDMs for both types of data (as measured by the true skill statistic; TSS) was highly similar but SDMs from citizen science data had an overall lower performance. Nevertheless a few species achieved good predictions (TSS > 0.6) using these data. We conclude that species observations in citizen science projects can provide accurate predictions of species

*Corresponding author at: CE3C, Centre for Ecology, Evolution and Environmental Changes, Faculdade de Ciências da Universidade de Lisboa, 1749-016 Lisbon, Portugal.
E-mail address: patricia.tiago@gmail.com (P. Tiago).

http://dx.doi.org/10.1016/j.baae.2017.04.001

realized distributions, however, efforts should be made to identify the conditions under which these data are more likely to provide reliable representations of the species niches.

Keywords: Citizen science; Opportunistic species observations; Climatic niche; Species distribution models

Introduction

Currently there is a large interest in citizen science – i.e. the engagement of non-professionals in scientific research (Miller-Rushing, Primack, & Bonney 2012) – marked by a strong increase on citizen science programs (Dickinson et al. 2012). The scope of these programs is wide, covering research areas such as conservation biology and biodiversity monitoring, which are using citizen science programs to collect large amounts of species distribution data allowing to fill existing gaps of information (e.g. Pereira & Cooper 2006; Danielsen et al. 2009; Danielsen, Burgess, Jensen, & Pirhofer-Walzl 2010; Pereira et al. 2010; Szabo, Vesk, Baxter, & Possingham 2010; Dickinson et al. 2012). These citizen science projects can collect data in a similar way to scientific surveys, i.e. following strict protocols. In these cases, the main difference between the two types of survey concerns the general lack of experience of the citizen scientists which can lead to taxonomic misidentifications, reducing data accuracy. Alternatively, citizen science surveys of species distributions can take place through the collection of opportunistic data, data collected by non-standardized methods, with no sampling design and no systematic protocol (Dickinson, Zuckerberg, & Bonter 2010). These later programs can cover wide spatial extents and often provide a large number of records (Chandler et al. 2012, 2017). These opportunistic records have the same problem of lack of taxonomic expertise of the participants and can, in many cases, be spatially and temporally biased (Beck, Ballesteros-Mejia, Nagel, & Kitching 2013; Higa et al. 2014). Bias in species observations provided by citizen-science programs may hinder the usefulness of these records in ecological research. Importantly, the sampling effort of opportunistic records is generally not known but it can vary widely over time (Dickinson et al. 2010; Snäll, Kindvall, Nilsson, & Pärt 2011) and across space (Dickinson et al. 2010; van Strien, van Swaay, & Termaat 2013) and between and within taxonomic groups (Kéry et al. 2010; Snäll et al. 2011).

Species distribution models (SDMs) – also known as ecological niche models or habitat selection models – are now widely used in ecological and evolutionary research (Elith & Leathwick 2009; Kozak, Graham, & Wiens 2008). These models relate data on species distributions with spatial environmental data in order to estimate locations where the species could occur (Elith & Leathwick 2009). The questions they allow to address are wide-ranging and include how climate change may modify biodiversity patterns (Thuiller et al. 2008), where invasive species may become established (Capinha & Anastácio 2011), where the hotspots of endangered species are located (Godown & Peterson 2000), which areas should be prioritized for conservation (Chen & Peterson 2002) or which locations are suitable for species translocations or cultivation (Jovanovic, Arnold, & Booth 2000; Cunningham, Anderson, & Walsh 2002). SDMs rely on two types of data, species distribution data and environmental data. While the latter is now widely available at high spatial resolution and for wide spatial extents (Kozak et al. 2008), mainly due to large-scale mapping and modelling projects (e.g. Higa et al. 2014; Levinsky et al. 2013), the geographic distribution of many species still remains poorly known (Scheffers, Joppa, Pimm, & Laurance 2012). In this context, it is relevant to understand if species observation records coming from citizen science projects are useful for inferring species distributions, and if so, under which conditions these records are more or less reliable.

In this work we assess whether opportunistic citizen science databases are viable data sources to model species distributions and test if species attributes can indicate the reliability and completeness of the opportunistic distribution data. We use amphibians and reptiles records from the BioDiversity4All database (www.biodiversity4all.org), a country-wide citizen science project in Portugal. We use herptiles (i.e. reptiles and amphibians) because many of these species tend to be cryptic and pass unnoticed and also because the prejudice associated with this group can affect the observations recorded in a citizen science project, as several of these species are feared and despised by many people (Price & Dorcas 2011). These characteristics contribute to a distributional data shortfall – as opposed, for instance, to a few other conspicuous and 'attractive' groups such as birds for which distributional data is more abundant. We use opportunistic citizen science records to measure the climatic niche width of 39 herptiles. We then compare these niche widths to the ones obtained using records from an, independent, long-term, scientific atlas and test for species traits and characteristics of the species distributions as indicators of the differences found. Finally, we also build projections of species distribution models based on each of the two distinct sources of species records and compare their predictive performances.

Materials and methods

In this work, we perform three main analyses to assess the merits of opportunistic citizen science records of species observations. In the first analysis, we measure the climatic