ECOGRAPHY

Forum

Applying species distribution models to caves and other subterranean habitats

Stefano Mammola and Boris Leroy

S. Mammola (http://orcid.org/0000-0002-4471-9055) (stefanomammola@gmail.com), Laboratory of Terrestrial Ecosystems, Dept of Life Sciences and Systems Biology, Univ. of Torino, Torino, Italy. – B. Leroy, Unité Biologie des organismes et écosystèmes aquatiques (BOREA UMR 7208), Muséum national d'Histoire naturelle, Sorbonne Univ., Univ. Pierre et Marie Curie, Univ. de Caen Normandie, Univ. des Antilles, CNRS, IRD, Paris, France.

Ecography
41: 1194–1208, 2018
doi: 10.1111/ecog.03464

Subject Editor: Tim Newbold
Editor-in-Chief: Miguel Araújo
Accepted 30 September 2017

Over the last two decades there has been an exponential increase in the use of correlative species distribution models (SDMs) to address a variety of topics in ecology, biogeography, evolution, and conservation biology. Conversely, the use of these statistical methods to study the potential distribution of subterranean organisms has lagged behind, relative to their above-ground (epigean) counterparts. The reason for this is possibly related to a number of peculiarities of subterranean systems, which pose important limits, but also opportunities, for these correlative models. The aim of this forum is to explore the caveats that need to be made when generalizing these statistical techniques to caves and other subterranean habitats. We focus on the typical bias in spatial datasets of cave-dwelling species, and provide advice for selecting the model calibration area. In parallel, we discuss the potential use of different large scale surface variables to represent the subterranean condition. A more widespread adoption of these statistical techniques in subterranean biology is highly attractive and has great potential in broadening our knowledge on a variety of ecological topics, especially in the fields of climate change and biodiversity conservation. Their use would especially benefit the study of the biogeographic patterns of subterranean fauna and the impact of past and future climate change on subterranean ecosystems.

Introduction

Species distribution modelling (SDM) refers to the practice of using mathematical algorithms to establish a relationship between occurrence data and environmental variables, with the aim of representing the ecological niche of a certain species or projecting a probability surface into a geographical space to represent its potential range of distribution (Franklin 2009, Peterson et al. 2011). Although these spatial models have often been referred to with different names (climatic envelope models, ecological niche models, species niche models, etc.), for simplicity we will use the acronym SDM hereafter; but see Sillero (2011) and Peterson and Soberón (2012) for semantic discussions. In the last few years there have been many developments in this field

NORDIC SOCIETY OIKOS

www.ecography.org

and numerous algorithms are currently available (Elith et al. 2006, 2011, Blonder et al. 2014, Qiao et al. 2015, 2016). These tools have given excellent opportunities to address a variety of questions in ecology, biogeography, evolution and conservation biology (Guisan and Thuiller 2005, Elith and Leathwick 2009, Booth et al. 2014). One of the most intriguing applications of these methods is the possibility of transferring current distribution models into new geographical spaces and time frames, with important applications in invasion biology and global change biology, respectively (Peterson and Vieglais 2001, Pearson and Dawson 2003, Peterson 2003, Hijmans and Graham 2006, Dormann 2007, Jiménez-Valverde et al. 2011).

For all these reasons, interest in SDM has grown exponentially in the last three decades, with hundreds of scientific papers published every year in peer-reviewed journals (Lobo et al. 2010: p. 104, f. 1). In contrast, the total number of published works applying these statistical techniques to the study of the distribution of subterranean organisms is orders of magnitude lower. As far as we are aware, in the last ten years there have been less than twenty contributions specifically employing SDM techniques to the study of subterranean species distribution (Table 1).

Subterranean habitats are among the most widespread terrestrial environments on our planet (a glossary of the jargon related to subterranean biology adopted in this text is given in Table 2). For example, over 45 000 caves have been reported from the United States (Culver 1999 to Christman and Culver 2001) and 20 000–25 000 caves are estimated solely for the Dinarides, the most important karst region in Europe (Zagmajster et al. 2008, 2010). However, Curl (1958, 1966, 1999) pointed out that the majority of caves have no entrances, meaning that the number of empty spaces beneath the earth's crust must be even higher. Moreover, troglobionts do not exclusively inhabit caves, but also dwell in smaller subterranean voids the size of which is not commensurable to the human scale (Culver and Pipan 2009, 2014, Giachino and Vailati 2010, Mammola et al. 2016a). Altogether, subterranean habitats contain a peculiar and spectacularly diverse fauna, with different taxonomic orders and even classes occurring in caves only a few kilometres apart (Culver et al. 2013). Culver and Holsinger (1992) estimated that there should be from 50 000 to 100 000 obligate cave-dwelling species worldwide, most of which have not yet been described, especially in the tropics (Trajano et al. 2016).

Considering that in the last two decades, the study of subterranean biology has seen many significant advances, especially from molecular and ecological standpoints (Juan et al. 2010, Moldovan 2013, Culver 2016, Pipan and Culver 2016), it is interesting to discuss why the application of correlative SDM techniques to subterranean fauna has lagged behind to this extent. There are different possible explanations for this paucity of studies. Foremost, it is probably counterintuitive to model the subterranean conditions using the typical external variables adopted in correlative SDMs. In many cases, the transfer of the variables, whether topographical or climatic, from the surface to subterranean habitats is not straightforward. Secondly, for most subterranean taxa there is a general lack of spatial datasets of species distribution, which are mandatory for SDM studies (Barbet-Massin et al. 2012), although recently some extensive occurrence databases have been developed, which should facilitate the development of SDMs for subterranean taxa (Zagmajster et al. 2008, 2014, Deharveng et al. 2009, Glanville et al. 2016). Thirdly, many subterranean species are extremely limited spatially, either due to inadequate sampling or to a high level endemism, and thus the minimum number of data points to calibrate a correlative model properly, may not be reached (Christman et al. 2016, Proosdij 2016). Furthermore, while a number of recent advances indicate that the study of macroecological patterns of subterranean species is a fast-moving field of research (Christman and Culver 2001, Culver et al. 2003, 2006, 2013, Zagmajster et al. 2008, 2010, 2014, Dole-Olivier et al. 2009, Gibert et al. 2009, Malard et al. 2009, Stoch and Galassi 2010, Niemiller and Zigler 2013, Eme et al. 2015, 2017, Bregović and Zagmajster 2016, Christman et al. 2016, Pellegrini et al. 2016), it may also be that SDMs have not yet reached popularity in the subterranean literature.

Regardless of the underlying reasons for this knowledge gap, it is arguable that an increase in the utilization of correlative SDMs would improve our understanding of the biogeography of cave-dwelling organisms, their niche dynamics and of the effects of climate change on subterranean ecosystems, among other topics. The aim of this forum is to explore the main caveats of SDM techniques when they are applied to caves and other subterranean habitats, with the overall goal of stimulating studies investigating the ecological niche and potential distribution of subterranean organisms.

Modelling the distribution of subterranean species

In order to fit a SDM it is necessary to collect occurrence data of the species of interest, and to relate these occurrences to a set of environmental variables (so-called predictors, covariates or inputs) within a calibration area (Soberón and Peterson 2005, Muñoz et al. 2009) (Fig. 1). However, the peculiarities of underground habitats and their organisms pose specific challenges for SDMs; see Schroeder (2008) for an example related to soils. In the following paragraphs, we discuss some important concepts regarding the selection of occurrence records, the environmental variables and the calibration area when applying models to subterranean habitats. We focus here on specificities of subterranean SDMs, and not on general guidance about the modelling procedures, model validation and the underlying statistical theory, which can be found in the literature. Many practical guidelines for

Table 1. Summary of the studies applying species distribution models (SDM) in subterranean ecosystems. We excluded studies on bats (Lisón et al. 2013, Soto-Centeno et al. 2015, Cooper-Bohannon et al. 2016, Herkt et al. 2016) and their parasites (Flory et al. 2012, Escobar et al. 2014) from this compilation, as bats are likely to use caves exclusively as shelters, and their distribution is likely to be driven by several external factors, such as the availability of feeding areas, flight paths, availability of roosts, etc.

Model organism(s)	Area	Variables	Modelling technique(s)	Aim(s)	Reference
Amphipod *Niphargus virei*	Europe (Jura Mountains, France)	Elevation, linear distance to the glacier	GLM	Influence of Pleistocene glaciations on current distribution	Foulquier et al. 2008
Amphipods (gen. *Niphargus*)	Europe (Istrian Peninsula, NW Balkans)	Bioclimatic variables	MaxEnt	Study of niche partitioning	Fišer et al. 2015
Amphipods (gen. *Niphargus*)	Southwestern England	Hydrogeological formations, water chemistry	GLM	Study of the factor driving the distribution and composition of stygobiont assemblages	Johns et al. 2014
Beetles (tribe Leptodirini)	Europe (Pyrenees)	Temperature variables	Bioclimatic envelope procedure based on grid cells	Study of climate change dynamics	Sánchez-Fernández et al. 2016
Cave salamanders (gen. *Eurycea* and *Plethodon*)	North America (northwestern Georgia)	Bioclimatic variables, NDVI indexes, soil parameters, vegetation	MaxEnt	Model potential distribution and thermal niche	Camp et al. 2014
Scorpion (*Pseudouroctonus reddelli*)	North America (Edwards Plateau karst region, Texas)	Bioclimatic variables	MaxEnt	Identify Pleistocene refugia	Bryson et al. 2014
Spiders (gen. *Pimoa*)	Europe (W-Alps, Appennine)	Bioclimatic variables, elevation	MaxEnt	Identify Pleistocene refugia. Study of niche partitioning	Mammola et al. 2015
Spiders (gen. *Meta*)	Great Britain	Elevation, geological substrate, precipitation, temperature	MaxEnt	Predict the present and future distribution of native and alien congeneric species	Mammola 2017
Spiders (gen. *Meta*)	Europe, north Africa	Bioclimatic variable, elevation	MaxEnt, n-dimensional hypervolumes	Study of niche partitioning and climate change dynamics	Mammola and Isaia 2017
Spiders (gen. *Troglohyphantes*)	Europe (NW Italian-Alps)	Bioclimatic variables, elevation, Pleistocene glaciers	MaxEnt, GLM, BRT	Study of climate change dynamics	Mammola et al. 2017
Troglobionts	North America (Illinois to Virginia and New York to Alabama)	Elevation, geological substrate, temperature and precipitation, vegetation, soil data.	GLM	Predict occurrence of cave-dwelling fauna	Christman et al. 2016
Troglobionts	Canary Island (Gran Canaria)	Vegetation, ecgological substrate, precipitation	MaxEnt	Predict occurrence of cave-dwelling fauna	Naranjo et al. 2014
True bug (gen. *Zelurus*)	South America	Bioclimatic variables, elevation	MaxEnt	Predict potential distribution	Ferreira et al. 2016
Obligate groundwater species (19 spp.)	Europe (Jura Mountains, France)	Elevation, ecological substrate, habitat variables, physical-chemical water parameters	GAM	Predict present and past potential distribution	Castellarini et al. 2007

BRT = boosted regression tree; ENFA = ecological niche factor analysis; GLM = generalized linear model; GAM = generalized additive model; MaxEnt = maximum entropy modelling.

applying SDMs are available (Muñoz et al. 2009, Elith et al. 2011, Merow et al. 2013), as well as statistical platforms which facilitate the computation of the different algorithms (Phillips et al. 2006, Thuiller et al. 2009, Warren et al. 2010, Qiao et al. 2012, Hijmans et al. 2014, Cola et al. 2016, Naimi and Araújo 2016).

Occurrence data

Sampling biases

Assembling reliable datasets of occurrences of obligate subterranean organisms is challenging for multiple reasons. Foremost, sampling in caves and other subterranean habitats

Table 2. Glossary.

Term	Definition adopted herein	Reference
Subterranean habitat(s)	All the aphotic undeground spaces and cavities harbouring species showing adaptation to the subterranean life (troglobiomorphic traits). These include human-accessible caves ('proper caves' sensu Curl 1964), but also the network of fissures and voids of smaller size, the shallow subterranean habitats (defined below) and various artificial subterranean habitats (mines, bunkers, cellars, etc.).	Culver and Pipan 2009, Romero 2009
Shallow subterranean habitats (SSHs)	The aphotic subterranean habitats close to the surface, harbouring species showing troglobiomorphic traits, such as epikarst, lava tubes, deep leaf litter, soil strata and the so-called Milieu Souterrain Superficiel.	Culver and Pipan 2014
Troglobiont	A terrestrial species strictly bound to a subterranean habitat.	Sket 2008
Stygobiont	An aquatic species strictly bound to a subterranean habitat.	
Troglophile	Essentially epigean species, but able to maintain a permanent subterranean population or temporarily to inhabit a subterranean habitat. Although Sket (2008) distinguished between sub- and eu-troglophiles, for simplicity we used the general term troglophile in the text.	
Trogloxene	An epigean species exploiting the subterranean habitats for occasional purposes.	

is particularly challenging because human access to these habitats is difficult, often demanding the use of specific caving techniques and equipment (Zagmajster et al. 2010). In addition, the low abundance of subterranean organisms often results in low detectability, requiring repeated visits/ sampling sessions to sample species (Schneider and Culver

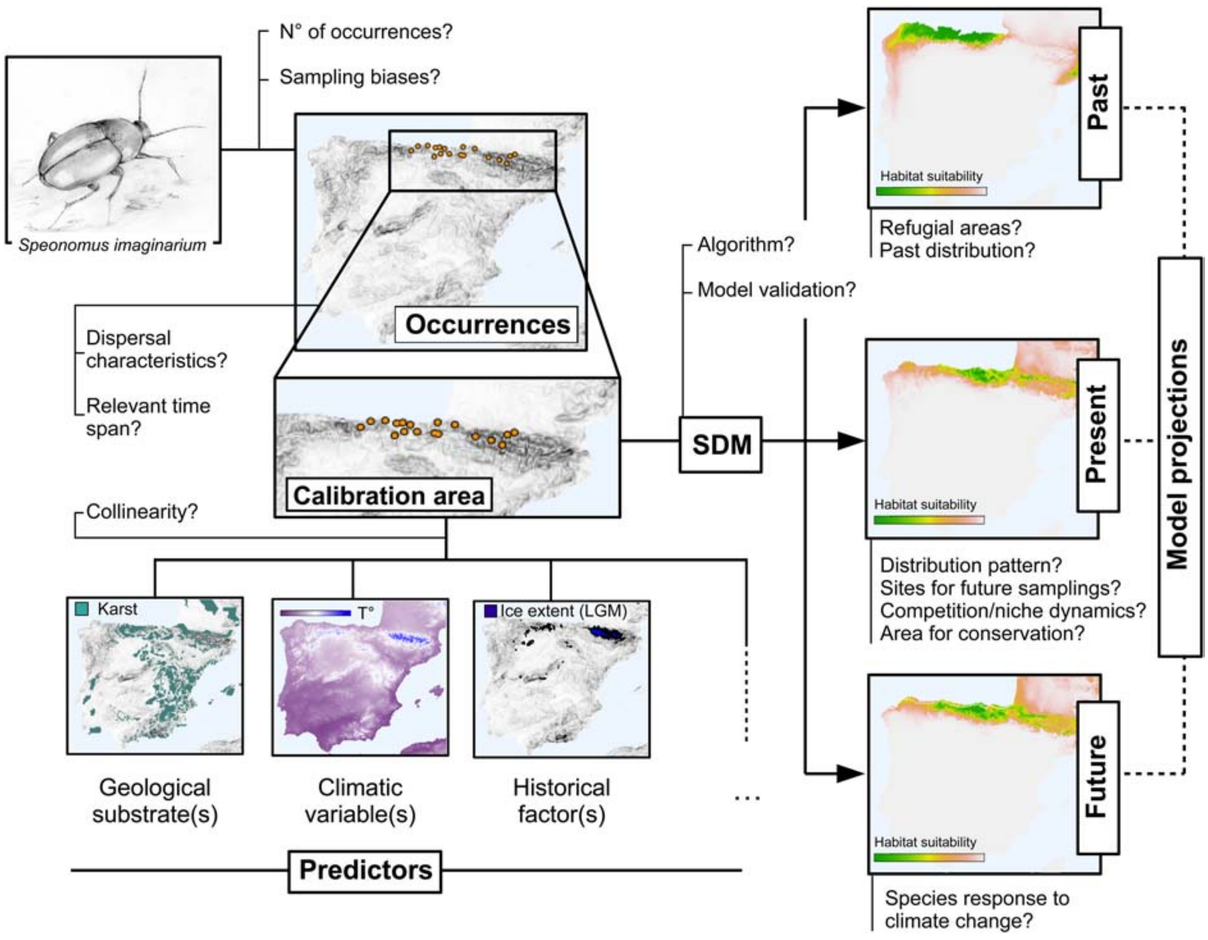

Figure 1. General workflow for species distribution modelling (SDM) of a subterranean species – an imaginary cave-dwelling beetle endemic to the Iberian Peninsula. In order to fit the model, it is necessary to collect occurrence data of the species (see section 'Occurrence data' for details), and to relate these occurrences to a set of explanatory variables (see section 'Environmental predictors' for details) within a calibration area (see relative section for details). Model, in turn, can be projected onto different geographic areas and temporal time-frames, in order to address various ecological and evolutionary questions (see section 'Applications of SDM in subterranean biology').

2004, Eberhard et al. 2009, Halse and Pearson 2014). Once samples have been acquired, a further problem lies in species identification afterwards, due to the lack of taxonomists. This problem is especially serious when considering taxa that are underrepresented within subterranean communities, such as minor orders of arthropods.

Secondly, a growing body of evidence reveals that a large number of subterranean organisms dwell in networks of small fissures (Giachino and Vailati 2010, Culver and Pipan 2014, Mammola et al. 2016a) or in caves with no entrances (Christman and Culver 2001). Since these habitats are not accessible to humans, except by indirect means, collection of species in these habitats requires the use of special sampling techniques – see Mammola et al. (2016a) for a review. Therefore, biological studies focusing on the subterranean fauna take place almost exclusively in caves.

The general consequences of these caveats are that a substantial part of subterranean biodiversity is unknown, absences are difficult to ascertain, and that distribution data are biased toward caves in the most investigated regions. These biases may result in a number of issues for SDMs, although not exclusive to the specific case of subterranean habitats. First, the likely unreliable nature of absences should lead to ponder the choice between presence-only and presence–absence modelling techniques (see Guisan and Zimmerman 2000, Guillera-Arroita et al. 2015, Jarnevich et al. 2015, Yackulic et al. 2015 which discuss in detail the assumptions, requirements and caveats of different modelling techniques). Second, given the aforementioned likelihood of imperfect detection in subterranean systems, the impact on SDM performance should be considered. It can be done by collecting data in a way that allows detectability to be modelled, which does not necessarily require a higher sampling effort (Lahoz-Monfort et al. 2014). Third, depending on sampling biases and on the modelling approach, SDMs may not be able to estimate occurrence probability, highlighting the necessity of a critical evaluation of the output given by a SDM to ensure that it can be applied to the defined research question (Guillera-Arroita 2017).

Furthermore, as discussed in the introduction, when dealing with species with high levels of endemism and small distribution ranges, even a complete sampling could yield an insufficient number of data points to estimate species niche with SDMs. In such cases, one possible solution would be to consider data at the genus or trophic level, i.e. a species complex instead of a single species (Johns et al. 2014, Mammola et al. 2017). It is clear that this approach will be possible only if species share similar ecological traits, e.g. a similar thermal physiology, and show a generally high niche overlap (Qiao et al. 2017).

Variation of sampling effort is also a common problem, especially in regions where the number of caves is large or because of the lack of scientists who are able to investigate underground habitats. For example, this issue is especially serious in regions where subterranean biology has only recently started to grow, such as the Neotropics (Trajano and Bichuette 2009, Silva and Ferreira 2016, Trajano et al. 2016). Incidentally, the distribution of most taxa is expect to be underestimated even in areas with historical and consolidated speleobiological traditions (Zagmajster et al. 2008, 2010) – although the use of modern surveillance tools, such as environmental DNA, is recently helping to overcome this impediment (Gorički et al. 2017, Năpăruș-Aljančič et al. 2017, Niemiller et al. 2017, Vörös et al. 2017).

For these reasons, occurrences of cave-dwelling species may be unevenly distributed in environmental space, containing gaps and areas of higher density (Christman and Culver 2001, Christman et al. 2005). Several methods to test for spatial auto-correlation in occurrence datasets and statistical methods to deal with different sources of sampling bias have been put forward in the SDM literature (Dormann et al. 2007, Phillips et al. 2009, Newbold 2010, Syfert et al. 2013, Olivera et al. 2014, Jarnevich et al. 2015). However, most auto-correlation correction measures will prevent correct extrapolation to new environments, either spatially and/or temporally. At the very least, we recommend aggregating occurrence data in grids at the resolution of environmental variables, rather than using raw point-locality occurrence data, to avoid inflation of the numbers of presences and absences as an effect of spatial sampling heterogeneity (Mammola 2017).

Mismatches between occurrence and species distribution

The issues previously discussed represent the primary sources of potential errors in spatial datasets of subterranean species that have to be taken into account prior to model fitting. However, there are other issues that can be associated with spatial datasets, which are often related to the way in which geographical coordinates are measured by cavers. GPS coordinates are almost always recorded at the cave entrance; a potential source of error is therefore introduced if the species under study actually dwells far from the entrance, so there may be a mismatch between the environmental values extracted from the occurrence point geo-referenced at the cave entrance, and the real habitat occupied by the species (Fig. 2).

To correct for this mismatch manually, the topographic position of the ground plan of the cave has to be derived, in order to establish the correct position occupied by the species within the subterranean environment, and to project it to the surface. Sophisticated tools that permit georeferencing of the internal geomorphology of karst are available in the specialized literature (Florea et al. 2002, Litwin 2008, Gallay et al. 2016). However, the decision to spend time and invest funds to obtain very precise GPS coordinates rather than approximations will largely depend on the scale of the phenomenon under investigation. This mismatch only relates to very big cave systems, and whenever the resolution of the environmental layers employed in the SDM is smaller than the cave length. It does not apply to vertical caves (pits and abysses) or in the case of species dwelling in the vicinity of the cave entrance. From a statistical point of view, the

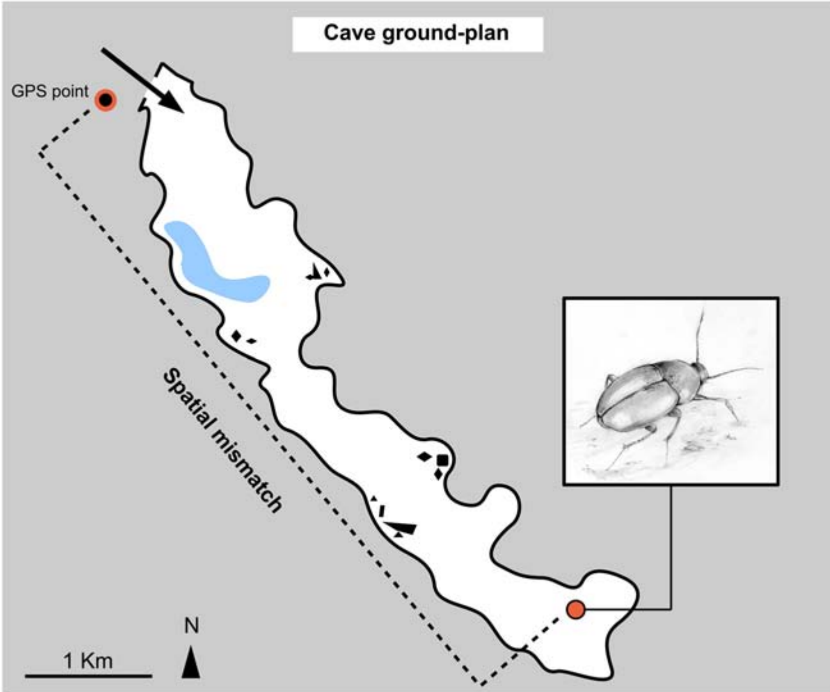

Figure 2. The ground plan of a hypothetical cave. A species of cave-dwelling beetle was collected at ca 5 km from the entrance, in the final chamber (orange dot). The collector associated the GPS coordinates corresponding to the cave entrance with this record (dot and arrow, respectively). If this occurrence is used for species distribution modelling (SDM), there will be a mismatch between the environmental values extracted at the entrance, and those in which the species actually dwells (dotted line). However, this mismatch will represent a potential source of error only if the spatial resolution of the environmental layer employed is < 5 km.

potential impact of this spatial mismatch on SDMs can be visualised and tested (Naimi 2013, Naimi et al. 2013).

Environmental predictors

Among the most important aspects that has hindered the development of spatial modelling in subterranean biology are i) the unavailability of useful environmental GIS variables and ii) the difficulties encountered when transferring external variables from the surface to the subsurface. However, the recent development of readily available large scale spatial variables can overcome the first impediment. Accordingly, the usability of different surface variables for understanding the patterns and processes in subterranean habitats has been recently discussed in the speleobiological literature (Stoch and Galassi 2010, Zagmajster et al. 2014, Eme et al. 2015, Mammola et al. 2015, 2017, Bregović and Zagmajster 2016, Christman et al. 2016, Mammola 2017).

The different responses obtained in studies published so far, suggest that the relative importance of different variables in determining distribution patterns is likely to vary depending on the scale of analysis (Field et al. 2009) – a nested hierarchy of environmental factors may determine species patterns at different scales (Stoch and Galassi 2010) – and the taxon under consideration. For example, while current climatic conditions and historical factors have proved to be important for explaining the potential distribution of different terrestrial arachnids (Bryson et al. 2014, Mammola et al. 2015, 2017, Mammola and Isaia 2017), with respect to aquatic crustaceans it has been shown that recent climate is a very poor predictor of subterranean distribution patterns (Zagmajster et al. 2014).

Accordingly, when applying SDM to subterranean systems it is crucial to evaluate the potential inclusion of a variety of environmental predictors, such as topographic, climatic, geological and historical variables. Moreover, in line with the recent literature (Murray et al. 2009, Brandt et al. 2017), the choice of variables should not only be based on statistical selection, but also include expert-based selection considering the specific biology of the taxon under study, e.g. terrestrial versus aquatic subterranean species. We hereafter discuss a list of potential predictors which can be used as proxies of the availability of suitable underground conditions, considering their potential transferability to the subterranean habitats.

Topographic factors and habitat heterogeneity

Habitat availability and habitat (pore) size are the first factors limiting the occurrence of a subterranean species (Christman and Culver 2001, Pipan and Culver 2017). Logically, a higher habitat and topographic heterogeneity should enhance a wider distribution in most subterranean taxonomic groups (Cornu et al. 2013, Eme et al. 2015, Christman et al. 2016). Accordingly, many variables that can be used as proxies of subterranean habitat availability have been put forward in the recent literature.

Different authors have suggested that elevation range is one of the simplest surrogates of topographic heterogeneity (Zagmajster et al. 2014, Eme et al. 2015, 2017, Bregović and Zagmajster 2016), and the inclusion of elevation data rasters is therefore common in most of the subterranean SDM studies published to date (Table 1). Different authors also have used variables related to the geological substrate(s) as a proxy for the availability of subterranean habitat. For instance, Christman et al. (2016) included the percent area of carbonate bedrock in 20 × 20 km grid cells as a measure of potential karst, whereas Cooper-Bohannon et al. (2016) used the distance from karst areas in their models. Both of these approaches are suitable for predicting the potential presence of suitable habitats in karst landscape because they reflect the potential presence of limestone caves. However, they would underestimate the availability of suitable habitats in non-karst substrates, e.g. tectonic caves, shallow subterranean habitats (SSHs; see Table 2 for the definition), or artificial subterranean habitats (e.g. mines, cellars, bunkers).

In order to account for a larger variety of potentially suitable habitats, more heterogeneous and detailed geomorphological maps have to be included as predictors, reflecting the variety of geological substrates that can potentially host subterranean habitats. In other words, the most appropriate habitat metric should be used depending on the subterranean habitat and taxon under consideration. For instance, Herkt et al. (2016) developed a specific variable representing the ruggedness of the terrain to reflect the presence of caves and crevices, whereas Christman et al. (2016) included a topographic position index as an indirect measure of landscape rugosity.

In some cases, environmental variables can have a threshold effect on species, i.e. their nature can lead to systematic species absence (e.g. aquatic species cannot thrive out of water). For example, species can be limited to karst landscapes, and thus the absence of karst will result in species absence; e.g. calciphilous invertebrates such as certain species of Gastropoda. If such information is available a priori, then these environmental variables do not need to be included as predictors in the model; they should rather be used as a binary filter to refine model outputs to suitable areas [an example of this two-step modelling process can be found in Gillard et al. (2017)].

Recently, more comprehensive representations of subterranean habitat availability have also been made available. Cornu et al. (2013) published a map of the distribution of groundwater habitats in Europe, which can be used for modelling aquatic subterranean species. The inclusion of predictors directly derived from this map has since become standard practice in studies addressing the distribution patterns of subterranean aquatic species in Europe (Zagmajster et al. 2014, Eme et al. 2015). We believe that further developments of similar maps, extending the geographic coverage and also considering the distribution of terrestrial subterranean habitats, would therefore represent stimulating endeavours, representing major advances in the understanding of distributional patterns of subterranean species.

Current climate

The majority of SDM studies of above-ground (epigean) ecosystems are based on the 19 'bioclimatic' variables, which constitute a set of variables derived from monthly means and extreme temperatures and rainfall regimes (Table 3); see Kriticos et al. (2012, 2014), Title and Bemmels (2017) and for further additions to the standard set of bioclimatic variables. Bioclimatic variables are available at different spatial resolutions, from coarse (10') to very fine (30"), for a current time scale (1970–2000; Fick and Hijmans 2017) and for several past and future scenarios (Hijmans et al. 2005, Kriticos et al. 2012).

One of the primary concerns about the use of bioclimatic variables to model subterranean species niches and distributions is that these variables may not represent subterranean conditions adequately (but see Table 3). This is probably because, in a first approximation, the link between external and internal (subterranean) climate is not intuitive. However, an ample body of literature documents how the temperature of the underground compartment generally reflects the climatic regime at the surface (Moore and Sullivan 1964, Smithson 1991, Badino 2004a, b, 2010). Despite the fact that underground temperature increases with depth due to the geothermal gradient, caves typically exhibit relatively constant temperatures. For example, the world's deepest cave (Krubera-Voronja) spans 2200 m vertically, with a temperature range of less than 5°C (Sendra and Reboleira 2012). The general explanation for this is found in the physics of the infiltrating meteoric fluids, which are regarded as the primary factors determining the general climatic conditions found underground (Badino 2004a, 2010, Brookfield et al. 2016), having a stronger effect than geothermal fluxes or air currents (Badino 2005, 2010). On a geological time scale, a certain rock layer is indeed forced into thermal equilibrium with the water flowing through it. Consequently, the temperature of the rock surrounding the cave equilibrates with the mean temperature of the infiltrating waters, which, in turn, is equal to the annual mean temperature outside (Badino 2010, Brookfield et al. 2016). This fact implies that it is possible to approximate the microclimate of the underground compartment on the basis of external climatic predictors (Mammola et al. 2017), especially the mean annual temperature (details in Table 3).

Past climate and other historical factors

Many studies suggest that, at least in temperate regions, there is a significant link between the current distribution of subterranean species and various paleogeographical events, such as glacial cycles and other climatic upheavals (reviewed by Culver and Pipan 2010). It has been suggested that these historical factors are especially important in explaining species patterns at broad scales (Stoch and Galassi 2010).

Accordingly, it is possible to incorporate past climatic reconstructions into SDM. For instance, the bioclimatic

Table 3. The bioclimatic variables (after Hijmans et al. 2005) and their relationships to subterranean conditions.

Variable	Description	Unit	Relation with the subterranean conditions
Bio01	Annual mean temperature	°C	Bio01 represents an almost ideal approximation of the thermal condition found in most deep subterranean habitats (Moore and Sullivan 1964, Smithson 1991, Badino 2004a, 2010). Elevation can also be used as a surrogate (Castellarini et al. 2007), but only for restricted geographical areas.
Bio02	Mean diurnal range	°C	Bio02–04 and Bio07 can represent proxies of the daily and seasonal divergence found in the vicinity of the cave entrance, where the external weather still has a daily and seasonal influence on the microclimate (Smithson 1991, Badino 2010). Also, they may represent more realistically the climatic conditions of very superficial caves and other shallow subterranean habitats (Pipan et al. 2011, Mammola et al. 2016a).
Bio03	Isothermality (Bio2/Bio7) (× 100)	°C	
Bio04	Temperature seasonality (standard deviation × 100)	°C	
Bio07	Temperature annual range (Bio05–Bio06)	°C	
Bio05	Max temperature of warmest month	°C	Given the constant cave temperatures and reduced temperature divergence, average values such as Bio01 are often prioritized when dealing with deep subterranean habitats. Bioclimatic variables related to maxima and minima may become useful for representing the conditions found in the outermost sectors of caves and in other shallow subterranean habitats. These variables may therefore be useful when the aim is to model the distribution of troglophile and trogloxene species (Novak et al. 2014).
Bio06	Min temperature of coldest month	°C	
Bio08	Mean temperature of wettest quarter	°C	
Bio09	Mean temperature of driest quarter	°C	
Bio10	Mean temperature of warmest quarter	°C	
Bio11	Mean temperature of coldest quarter	°C	
Bio12	Annual precipitation	mm	Infiltrating meteoric fluids are the primary factors determining the general climatic conditions found underground (see text for details). In addition, meteoric fluids represent an important route for organic materials to enter subterranean ecosystems (hydrochoric transportation). Therefore, the precipitation regime may represent a proxy for external energy inputs (Bregović and Zagmajster 2016). Fišer et al. (2015) also suggested that heavy precipitation may cause catastrophic ecological drift in the underground aquifer. Thus, bioclimatic variables related to precipitation can reflect the susceptibility of certain subterranean habitats to disturbance [see Ortuño et al. (2013) for an example].
Bio13	Precipitation of wettest month	mm	
Bio14	Precipitation of driest month	mm	
Bio15	Precipitation seasonality (coefficient of variation)	mm	
Bio16	Precipitation of wettest quarter	mm	
Bio17	Precipitation of driest quarter	mm	
Bio18	Precipitation of warmest quarter	mm	
Bio19	Precipitation of coldest quarter	mm	

variables introduced in the previous section are currently available for different past time-frames, such as the last interglacial (~120 000–140 000 yr ago), the last glacial maximum (~22 000 yr ago), and the middle Holocene (~6000 yr ago). So, by incorporating past climate into spatial models, Zagmajster et al. (2014) showed that past climatic oscillations were the most important determinants of range sizes of groundwater crustaceans, rather than current climatic seasonality and the availability and heterogeneity of habitats.

Different authors have also attempted to incorporate biogeographical factors into SDM. Castellarini et al. (2007) and Foulquier et al. (2008) included the linear distance of each occurrence point from the Würm glaciers as a predictor. Mammola et al. (2017) employed a raster layer representing the shape of the glaciers during the Last Glacial Maximum (LGM), in order to explain cave occupancy by subterranean spiders as a function of the glacial history.

Future climate

One of the timeliest applications of SDM is the study of global climate change. In fact, models fitted in present-day conditions can be transferred to novel time periods under different energy, land-use, and emissions projections (Moss et al. 2010, IPCC 2014, Riahi et al. 2017), thus forecasting potential variations in species ranges due to anthropogenic climate and land-use change (Beaumont et al. 2008).

As far as caves are concerned, it is worth noting that direct field observations (Domìnguez-Villar et al. 2015) and theoretical models (Badino 2004a, Covington and Perne 2015) indicate that caves have thermal inertia, which results in time-lags of up to several decades in the warming of deep subterranean systems. As an example, Domìnguez-Villar et al. (2015) estimated a time-lag of ca 20 yr for Postonja cave, Slovenia. In practical terms, this time-lag implies that for modelling future habitat modifications of deep subterranean species, the projection of SDM should be always accompanied by a time interval rather than a precise time-scale – see Mammola et al. (2017) for a discussion. Conversely, for species living in superficial sections of caves and in other SSHs, the standard projected emission time frames can be used (Badino 2004a).

It is also worth noting that, for epigean species, correlative distribution models prioritize the usage of extreme values, the maxima and minima, instead of average values, as the former represent more realistic physiological constraints for species. However, for subterranean species, average values may be more appropriate, because of more constant temperatures and reduced temperature excursions (details in Table 3). Interestingly, the use of average values may increase the reliability of future projections, given that most future climate projections used in the SDM literature come from statistically downscaled general circulation models, which represent average trends better than extremes (Trzaska and Schnarr 2014; but see Zimmermann et al. 2009).

Trophic resources

There is a direct connection between surface productivity and the availability of food below ground (Souza Silva et al. 2011, Venarsky et al. 2017), consisting mainly of organic products which have drifted underground from the soils and by other routes (Gers 1998, Poulson and Lavoie 2000; Table 3). Thus, it has been hypothesized that higher surface productivity parallels a richer subterranean community (Culver et al. 2006, Eme et al. 2015). Data on soil characteristics, e.g. water capacity, permeability, thickness, percent organic matter, and the above ground vegetation cover should therefore be considered among the potential predictors of subterranean species distribution.

Several variables reflecting surface productivity can also be employed. The most logical are raster layers representing plant cover at the surface or relative metrics such as the normalized difference vegetation index (NDVI) index (Camp et al. 2014) or the mean annual evapotranspiration (Eme et al. 2015, Bregović and Zagmajster 2016). However, when lacking these kinds of GIS data at high spatial resolutions, other proxies for surface productivity can be used, namely the precipitation regime or the mean annual temperature (Bregović and Zagmajster 2016; Table 3). Nevertheless, in such cases it may become difficult to disentangle between direct effects of climate on species and indirect effects through surface vegetation.

Biological factors

There are some cave-dwelling organisms that are likely to be limited by the availability of very specific biological factors. It is self-evident, for instance, that guanobionts (organisms feeding on guano) will be present in a certain cave, only in the presence of bat colonies; or that the presence of specific parasites and parasitoids will be paralleled by the presence of their hosts. Accordingly, it is likely that the distribution of these specific organisms will significantly overlap the distribution of the organisms they are associated with. A distribution map of the latter may therefore be included as a predictor of the niche/distribution of the former. However, strong support for the interaction between the organisms should be available a priori, to support the use of a distribution map as a predictor. Otherwise, there is a risk that the inclusion of species distribution as a predictor in the model of another species will not reflect a biotic interaction, but simply the absence of key environmental predictors (Guisan and Thuiller 2005, Araújo and Luoto 2007).

Similarly, inter-specific competition is also likely to play an important role in constraining the spatial distribution of different subterranean organisms. It is worth noting that with respect to subterranean ecosystems, competition dynamics between species has been elucidated almost exclusively at a local scale (Culver 1973, 1975, Poulson 1977, Novak et al. 2010, Fišer et al. 2012, Mammola and Isaia 2014, 2016, Mammola et al. 2016b, Resende and Bichuette 2016), whereas patterns of inter-specific competition at a regional scale remain largely unexplored (but see Mammola and Isaia 2017). Lacking detailed information about competitive exclusion dynamics, the decision to include potential biotic interactions in the models should be evaluated on a case-by-case basis.

Calibration area

In recent years, many authors have shown how the extent of the study area (calibration area), can significantly influence the results of SDM (Anderson and Raza 2010, Giovanelli et al. 2010, Owens et al. 2013). The calibration area, sometimes referred to as the **M** area (Soberón and Peterson 2005, Barve et al. 2011), represents a region that is thought to have been 'sampled' by the species throughout its relevant history. In other words, the calibration area is the geographic extent hypothesized to fall within the long-term dispersal and colonization potential for a certain species during its evolutionary history.

Despite its importance (Saupe et al. 2012, Merow et al. 2013), to date, no standard rationale has been put forward to define the extent of the study area for constructing SDM (Anderson and Raza 2010). Its extent largely depends on the natural history, dispersal characteristics and biogeographic history of the species in question, and the geographic distribution of suitable habitats in the landscape of interest. Barve et al. (2011) suggested that authors should evaluate these aspects when defining the calibration area, explicitly stating the rationale chosen for defining its extent in the Methods section of the paper. As far as subterranean species are concerned, some specific considerations apply, as detailed in the next two paragraphs.

Dispersal characteristics

The first step in defining the calibration area, is the evaluation of the dispersal potential of the species under study, using existing knowledge about species dispersal abilities, genetic structuring and natural barriers hindering dispersal. Species with limited dispersal potential should have a small calibration area (a subregional extent), whereas species with large dispersal potential should have a larger one (from a regional to continental extent). When knowledge about species dispersal potential is lacking, it can be approximated on the basis of habitat connectivity – e.g. identifying major geomorphological barriers to gene flow (Chiari et al. 2012, Rizzo et al. 2017).

As a general rule, the colonization potential of most subterranean species is extremely reduced, with cave-adapted species often having small geographical ranges and high levels of endemism (Barr and Holsinger 1985, Culver et al. 2000, Michel et al. 2009). During the process of adaptation to a subterranean medium, animals have indeed fine-tuned their physiological tolerance to the stable condition of their habitat, which hampers their dispersal ability via non-subterranean habitats. Accordingly, studies conducted so far have uncovered pronounced genetic structuring and low gene flow – if any – between cave populations (Caccone 1985, Bohonak 1999, Dixon and Zigler 2011, Mammola et al. 2015, Weckstein et al. 2016). For these reasons, a small calibration area can be approximated for most subterranean species.

However, it should be kept in mind that subterranean habitats are often connected through a network of small cracks and voids inaccessible to humans, which may represent an important dispersal route for invertebrates (Romero 2012, Culver and Pipan 2014). In this respect, subterranean organisms dwelling in freshwater habitats (stygobionts) should generally have larger dispersal potential (Lamoreux 2004, Buhay and Crandall 2005, Lefébure et al. 2006), as subterranean aquatic habitats display higher connectivity than terrestrial ones (Christman and Culver 2001). For instance, Cornu et al. (2013) estimated that in Europe, groundwater habitat patches have average areas of 315 ± 4257 km^2, which is considerably higher than the area of most terrestrial subterranean complexes documented to date. A larger calibration area can also be expected a priori for troglophile and trogloxene species showing higher dispersal ability via non-subterranean habitats, such as Rhaphidophoridae cave crickets (Ketmaier et al. 2013) and ballooning cave spiders (Smithers 2005, Mammola and Isaia 2014, 2017, Mammola 2017).

Finally, the potential dispersal of the species of interest may be limited by inter-specific competitive exclusion dynamics – usually with more competitive closely-related species. In such a case, the area occupied by competitor(s) could be excluded from the final shape of calibration area.

Relevant time span

When a species under investigation is a new arrival in a landscape, basically its calibration area can only be estimated based on dispersal characteristics. However, when a species has a longer evolutionary history in a landscape, the estimation of the calibration area is much more complex, as it will also depend on the environmental changes that have occurred throughout its evolutionary history (see details in Barve et al. 2011). If this concept is applied to troglobionts, the timing of subterranean colonization or the time since isolation from the immediate epigean sister-group should be estimated somehow. It is, however, often difficult to reconstruct the early stages of the evolution of ancient subterranean lineages (Strecker et al. 2003), and basically this information is lacking for most of the taxa described to date (Trajano 2007).

If the timing of cave formation in the landscape of interest is available, one could use this information as an approximation of the relevant time span. However, this practice is not without its own problems, as it would underestimate the relevant time span for a species that was present in the area prior to the cave formation (e.g. SSH-dwelling species) and conversely, would overestimate the time span for a species that has colonized the subterranean medium in recent times. Some subterranean species do indeed represent 'ancient lineages' that are 'young troglobites' (Zhang and Li 2013), meaning that their process of subterranean adaptation took place only after the cave formation.

Applications of SDM in subterranean biology

As with epigeal and marine ecosystems, the application of SDM techniques to subterranean habitats can significantly improve our general understanding of a variety of topics. Whilst it is not possible to make a unique and prescriptive list of the main applications of these tools, the following compilation is intended to highlight some possible lines of research.

Distribution patterns and macroecology

In the last two decades efforts have been made to unravel distribution patterns of subterranean organisms and to address a variety of macro-ecological questions (Christman and Culver 2001, Culver et al. 2003, 2006, 2013, Dole-Olivier et al. 2009, Malard et al. 2009, Cardoso 2012, Niemiller and Zigler 2013, Eme et al. 2014, 2017, Zagmajster et al. 2014, Bregović and Zagmajster 2016, Pellegrini et al. 2016). Regarding this, species distribution models represent additional tools available to subterranean biologists, suitable for predicting subterranean biodiversity and its variation over space and time. As explained above, these tools may serve to predict the distribution of subterranean fauna based on the features at the earth's surface (Naranjo et al. 2014, Christman et al. 2016). Also, these tools can be used to identify unexplored areas of high environmental suitability for improving the geographical gaps of species distributions (Raxworthy et al. 2003, Engler et al. 2004). Predicting subterranean species distribution has indeed the potential to fill knowledge gaps, for example by using inferred distribution maps to understand seemingly incomplete distribution data, and even suggesting suitable sites for further sampling. Furthermore, applying SDMs to subterranean species will open up new opportunities for biodiversity conservation [see Guisan et al. (2013) for a review].

Biotic interactions

In recent years, it has been shown that biotic interactions could affect the prediction of SDMs even at larger scales (Davis et al. 1998, Araújo and Luoto 2007). Animal communities in extreme environments such as caves, present intriguing test cases for the study of niche-based processes and competition dynamics (Fišer et al. 2012, Mammola et al. 2016b). In this respect, correlative models represent interesting tools, allowing the investigation of spatial segregation at the landscape level (Mammola and Isaia 2017) and the assessment of overlaps in species bioclimatic niches (Fišer et al. 2015). However, when dealing with this topic, caution should be exercised because habitat is itself highly clustered. As a result, spatial segregation may not be based on competition or niche separation, but the result of simple physical processes.

Future climate change

To date, the effects of anthropogenic climate change on biological communities in subterranean ecosystems have seldom been explored. This is mostly because there is a general lack of historical series of meteorological data in caves. As such, the study of climate change dynamics in subterranean habitats has primarily relied on theoretical models (Badino 2004a).

However, recent applications of SDM to cave ecosystems have shown that these techniques can help to overcome this impediment (Table 1). A substantial increase in the application of SDM to subterranean ecosystems has the potential to fill a major gap in our estimation of climate change impacts on biodiversity (Bellard et al. 2012). Given the contradictory results obtained in the rare SDM studies (cf. Sánchez-Fernández et al. 2016 and Mammola et al. 2017), the sensitivity of subterranean ecosystems to climate change is largely unknown, thereby stressing an urgent need for further studies.

Past climate change

The theory of relicts and refuges (Botosaneanu and Holsinger 1991) puts great emphasis on long-term climatic changes, such as glaciation cycles and other large-scale climatic upheavals, as the main factors that prompted the colonization of the subterranean habitat, at least in temperate regions. In this area, SDM can be used to reconstruct the past distribution of subterranean species, and identify areas that acted as refugia during these adverse climatic conditions, such as the Pleistocene glaciations. This is possible because current distribution models based on bioclimatic variables can be projected onto past climatic reconstructions. This approach is particularly useful in combination with other lines of evidence, including molecular data (Bryson et al. 2014, Mammola et al. 2015).

Conclusions

In the last few decades, the potential of correlative distribution models as useful tools in ecology and biogeography has risen in the scientific literature. However, as with any model-fitting exercise, there are many approximations and uncertainties associated with SDM, and thus the outputs of these models should always be interpreted with caution. Precautions apply to an even greater extent when dealing with subterranean ecosystems, because their modelling heavily depends on explanatory variables that are external to the system. As such, special efforts should be made to undertake the necessary methodological steps to minimize uncertainties. In parallel, a clear statement about the modelling assumptions should always be included in the Methods section of the paper (see also Culver et al. 2013). A progress would be to combine SDMs with other lines of evidence, such as molecular and physiological studies, to improve our capacity to understand and predict biogeographic patterns (Carstens and Richards 2007, Knowles et al. 2007, Peterson 2009, Jessica et al. 2014). There is little doubt that this integrated approach would greatly benefit the science of subterranean biology, prompting a fast and significant advance in knowledge, especially in the fields of climate change and biodiversity conservation.

Acknowledgements – We would like to dedicate this work in memory of the extraordinary physicist and speleologist Giovanni Badino, who recently passed away after a long illness. He offered useful consultancy regarding underground microclimate. We are grateful to Marco Isaia for insightful discussions on several topics addressed in this work. We thank Aldyth Nys for proof-reading our English. Three anonymous referee provided a number of valuable advices which significantly increased the quality of the manuscript.
Author contributions – SM conceived the study and led the writing. BL provided fundamental statistical arguments and contributed to the writing of the manuscript.

References

Anderson, R. P. and Raza, A. 2010. The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela. – J. Biogeogr. 37: 1378–1393.

Araújo, M. B. and Luoto, M. 2007. The importance of biotic interactions for modelling species distributions under climate change. – Global Ecol. Biogeogr. 16: 743–753.

Badino, G. 2004a. Cave temperature and global climatic change. – Int. J. Speleol. 33: 103–114.

Badino, G. 2004b. Clouds in caves. – SEKA 2: 1–8.

Badino, G. 2005. Underground drainage systems and geothermal flux. – Acta Carsol. 34: 277–316.

Badino, G. 2010. Underground meteorology. What's the weather underground? – Acta Carsol. 39: 427–448.

Barbet-Massin, M. et al. 2012. Selecting pseudo-absences for species distribution models: how, where and how many? – Methods Ecol. Evol. 3: 327–338.

Barr, T. C. and Holsinger, J. R. 1985. Speciation in cave faunas. – Annu. Rev. Ecol. Syst. 16: 313–317.

Barve, N. et al. 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. – Ecol. Model. 222: 1810–1819.

Beaumont, L. J. et al. 2008. Why is the choice of future climate scenarios for species distribution modelling important? – Ecol. Lett. 11: 1135–1146.

Bellard, C. et al. 2012. Impacts of climate change on the future of biodiversity. – Ecol. Lett. 15: 365–377.

Blonder, B. et al. 2014. The n-dimensional hyper-volume. – Global Ecol. Biogeogr. 23: 595–609.

Bohonak, A. J. 1999. Dispersal, gene flow and population structure. – Q. Rev. Biol. 74: 21–45.

Booth, T. H. et al. 2014. BIOCLIM: the first species distribution modelling package, its early applications and relevance to most current MAXENT studies. – Divers. Distrib. 20: 1–9.

Botosaneanu, L. and Holsinger, J. 1991. Some aspects concerning colonization of the subterranean realm – especially subterranean waters: a response to Rouch and Danielopol, 1987. – Stygologia 6: 11–39.

Brandt, L. A. et al. 2017. Comparison of climate envelope models developed using expert-selected variables versus statistical selection. – Ecol. Model. 345: 10–20.

Bregović, P. and Zagmajster, M. 2016. Understanding hotspots within a global hotspot – identifying the drivers of regional species richness patterns in terrestrial subterranean habitats. – Insect Conserv. Divers. 9: 268–281.

Brookfield, A. E. et al. 2016. Effects of changing meteoric precipitation patterns on groundwater temperature in karst environments. – Groundwater doi: 10.1111/gwat.12456

Bryson, Jr, R. W. et al. 2014. Caves as microrefugia: pleistocene phylogeography of the troglophilic North American scorpion *Pseudouroctonus reddelli*. – BMC Evol. Biol. 14: 9.
Buhay, J. E. and Crandall, K. A. 2005. Subterranean phylogeography of freshwater crayfishes shows extensive gene flow and surprisingly large population sizes. – Mol. Ecol. 14: 4259–4273.
Caccone, A. 1985. Gene flow in cave arthropods: a qualitative and quantitative approach. – Evolution 1223–1235.
Camp, D. A. et al. 2014. Role of temperature in determining relative abundance in cave twilight zones by two species of lungless salamander (family Plethodontidae). – Can. J. Zool. 92: 119–127.
Cardoso, P. 2012. Diversity and community assembly patterns of epigean vs. troglobiont spiders in the Iberian Peninsula. – Int. J. Speleol. 41: 83–94.
Carstens, B. C. and Richards, C. L. 2007. Integrating coalescent and ecological niche modeling in comparative phylogeography. – Evolution 61: 1439–1454.
Castellarini, F. et al. 2007. Modelling the distribution of stygobionts in the Jura Mountains (eastern France). Implications for the protection of ground waters. – Divers. Distrib. 13: 213–224.
Chiari, Y. et al. 2012. Phylogeography of sardinian cave salamanders (genus *Hydromantes*) is mainly determined by geomorphology. – PLoS One 7: e32332.
Christman, M. C. and Culver, D. C. 2001. The relationship between cave biodiversity and available habitat. – J. Biogeogr. 28: 367–380.
Christman, M. C. et al. 2005. Patterns of endemism of the eastern North American cave fauna. – J. Biogeogr. 32: 1441–1452.
Christman, M. C. et al. 2016. Predicting the occurrence of cave-inhabiting fauna based on features of the earth surface environment. – PLoS One 11: e0160408.
Cola, V. D. et al. 2016. ecospat: an R package to support spatial analyses and modeling of species niches and distributions. – Ecography doi: 10.1111/ecog.02671
Cooper-Bohannon, R. et al. 2016. Predicting bat distributions and diversity hotspots in southern Africa. – Hystrix 27 doi: 10.4404/hystrix-27.1-11722
Cornu, J. F. et al. 2013. The distribution of groundwater habitats in Europe. – Hydrogeol. J. 21: 949–960.
Covington, M. D. and Perne, M. 2015. Consider a cylindrical cave: a physicist's view of cave and karst science. – Acta Carsol. 44: 363–380.
Culver, D. C. 1973. Competition in spatially heterogeneous systems: an analysis of simple cave communities. – Ecology 54: 102–110.
Culver, D. C. 1975. Interaction between competition and predation in cave stream communities. – Int. J. Speleol. 7: 229–245.
Culver, D. C. 1999. Ecosystem and species diversity beneath our feet. Terrestrial ecoregions of North America. – In: Ricketts, T. H. et al. (eds), A conservation assessment. Island Press, pp. 56–59.
Culver, D. C. 2016. 23rd international conference on subterranean biology. – Subt. Biol. 19: 65–85.
Culver, D. C. and Holsinger, J. R. 1992. How many species of troglobites are there? – Nat. Speleol. Soc. Bull. 54: 79–80.
Culver, D. C. and Pipan, T. 2009. The biology of caves and other subterranean habitats. – Oxford Univ. Press.
Culver, D. C. and Pipan, T. 2010. Climate, abiotic factors, and the evolution of subterranean life. – Acta Carsol. 39: 39–577.
Culver, D. C. and Pipan, T. 2014. Shallow subterranean habitats: ecology, evolution, and conservation. – Oxford Univ. Press.
Culver D. C. et al. 2000. Obligate cave fauna of the 48 contiguous United States. – Conserv. Biol. 14: 386–401.
Culver, D. C. et al. 2003. The North American obligate cave fauna: regional patterns. – Biodivers. Conserv. 12: 441–468.
Culver, D. C. et al. 2006. The mid-latitude biodiversity ridge in terrestrial cave fauna. – Ecography 29: 120–128.
Culver, D. C. et al. 2013. Paving the way for standardized and comparable subterranean biodiversity studies. – Subt. Biol. 10: 1–43.
Curl, R. L. 1958. A statistical theory of cave entrance evolution. – Natl Speleol. Soc. Bull. 20: 9–22.
Curl, R. L. 1964. On the definition of a cave. – Natl Speleol. Soc. Bull. 26: 1–6.
Curl, R. L. 1966. Caves as a measure of karst. – J. Geol. 74: 798–830.
Curl, R. L. 1999. Entranceless and fractal caves revisited. – Karst Model. 5: 183–185.
Davis, A. J. et al. 1998. Making mistakes when predicting shifts in species range in response to global warming. – Nature 391: 783–786.
Deharveng, L. et al. 2009. Groundwater biodiversity in Europe. – Freshwater Biol. 54: 709–726.
Dixon, G. B. and Zigler, K. S 2011. Cave-obligate biodiversity on the Campus of Sewanee: the University of the South, Franklin County, Tennessee. – Southeast. Nat. 10: 251–266.
Dole-Olivier, M. J. et al. 2009. Relationships between environmental gradients and groundwater biodiversity at a regional scale. – Freshwater Biol. 54: 797–813
Domìnguez-Villar, D. et al. 2015. Is global warming affecting cave temperatures? Experimental and model data from a paradigmatic case study. – Clim. Dyn. 45: 569–581.
Dormann, C. F. 2007. Promising the future? Global change projections of species distributions. – Basic Appl. Ecol. 8: 387–397.
Dormann, C. F. et al. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. – Ecography 30: 609–628
Eberhard, S. M. et al. 2009. Exploring the relationship between sampling efficiency and short-range endemism for groundwater fauna in the Pilbara region, Western Australia. – Freshwater Biol. 54: 885–901.
Elith, J. and Leathwick, J. R. 2009. Species distribution models: ecological explanation and prediction across space and time. – Annu. Rev. Ecol. Evol. Syst. 40: 677–697.
Elith, J. et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. – Ecography 29: 129–151.
Elith, J. et al. 2011. A statistical explanation of MaxEnt for ecologists. – Divers. Distrib. 17: 43–57.
Eme, D. et al. 2015. Multi-causality and spatial non-stationarity in the determinants of groundwater crustacean diversity in Europe. – Ecography 37: 1–10.
Eme, D. et al. 2017. Do cryptic species matter in macroecology? Sequencing European groundwater crustaceans yields smaller ranges but does not challenge biodiversity determinants. – Ecography doi: 10.1111/ecog.02683
Engler, R. et al. 2004. An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. – J. Appl. Ecol. 41: 263–274.
Escobar, L. E. et al. 2014. Potential for spread of the white-nose fungus (*Pseudogymnoascus destructans*) in the Americas: use of Maxent and NicheA to assure strict model transference. – Geospatial Health 9: 221–229.

Ferreira, M. I. et al. 2016. The genus *Zelurus* Hahn, 1826, in Brazilian caves: description of new species and comments on the potential distribution of the genus in South America. – Zootaxa 4170: 250–270.
Fick, S. E. and Hijmans, R. J. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. – Int. J. Climatol. doi: 10.1002/joc.5086
Field, R. et al. 2009. Spatial species-richness gradients across scales: a meta-analysis. – J. Biogeogr. 36: 132–147.
Fišer, C. et al. 2012. Niche-based mechanisms operating within extreme habitats: a case study of subterranean amphipod communities. – Biol. Lett. 8: 578–581.
Fišer, Ž. et al. 2015. Morphologically cryptic amphipod species are "ecological clones" at regional but not at local scale: a case study of four *Niphargus* species. – PLoS One 10: e0134384.
Florea, L. J. et al. 2002. Karst GIS advances in Kentucky. – J. Cave Karst St. 64: 58–62.
Flory, A. R. et al. 2012. Environmental conditions associated with bat white-nose syndrome mortality in the north-eastern United States. – J. Appl. Ecol. 49: 680–689.
Foulquier, A. et al. 2008. The imprint of Quaternary glaciers on the present-day distribution of the obligate groundwater amphipod *Niphargus virei* (Niphargidae). – J. Biogeogr. 35: 552–564.
Franklin, J. 2009. Mapping species distributions: spatial inference and prediction. – Cambridge Univ. Press.
Gallay, M. et al. 2016. Geomorphometric analysis of cave ceiling channels mapped with 3-D terrestrial laser scanning. – Hydrol. Earth Syst. Sci. 20: 1827–1849.
Gers, C. 1998. Diversity of energy fluxes and interactions between arthropod communities: from soil to cave. – Acta. Oecol. 19: 205–213.
Giachino, P. M. and Vailati, D. 2010. The subterranean environment. Hypogean life, concepts and collecting techniques. – WBA Handbooks.
Gibert, J. et al. 2009. Assessing and conserving groundwater biodiversity: synthesis and perspectives. – Freshwater Biol. 54: 930–941.
Gillard, M. et al. 2017. Present and future distribution of three aquatic plants taxa across the world: decrease in native and increase in invasive ranges. – Biol. Invasions 19: 2159– 2170.
Giovanelli, J. G. et al. 2010. Modeling a spatially restricted distribution in the Neotropics: how the size of calibration area affects the performance of five presence-only methods. – Ecol. Model. 221: 215–224.
Glanville, K. et al. 2016. Biodiversity and biogeography of groundwater invertebrates in Queensland, Australia. – Subt. Biol. 17: 55–76.
Gorički, Š. et al. 2017. Environmental DNA in subterranean biology: range extension and taxonomic implications for *Proteus*. – Sci. Rep. 7: 45054.
Guillera-Arroita, G. 2017. Modelling of species distributions, range dynamics and communities under imperfect detection: advances, challenges and opportunities. – Ecography 40: 281–295.
Guillera-Arroita, G. et al. 2015. Is my species distribution model fit for purpose? Matching data and models to applications. – Global Ecol. Biogeogr. 24: 276–292.
Guisan, A. and Zimmerman, N. 2000. Predictive habitat distribution models in ecology. – Ecol. Model. 135: 147–186.
Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. – Ecol. Lett. 8: 993–1009.
Guisan, A. et al. 2013. Predicting species distributions for conservation decisions. – Ecol. Lett. 16: 1424–1435.
Halse, S. and Pearson, G. 2014. Troglofauna in the vadose zone: comparison of scraping and trapping results and sampling adequacy. – Subt. Biol. 13: 17–34.
Herkt, K. M. B. et al. 2016. A high-resolution model of bat diversity and endemism for continental Africa. – Ecol. Model. 320: 9–28.
Hijmans, R. J. and Graham, C. H. 2006. The ability of climate envelope models to predict the effect of climate change on species distributions. – Global Change Biol. 12: 2272–2281.
Hijmans, R. J. et al. 2005. Very high resolution interpolated climate surfaces for global land areas. – Int. J. Climatol. 25: 1965–1978.
Hijmans, R. J. et al. 2014. dismo: species distribution modeling. – R package ver. 1.0-5, <http://CRAN.R-project.org/package=dismo>.
IPCC 2014. Climate change 2014: synthesis report. Contribution of working groups I, II and III to the fifth assessment report of the intergovernmental panel on climate change. – IPCC.
Jarnevich, C. S. et al. 2015. Caveats for correlative species distribution modeling. – Ecol. Inform. 29: 6–15.
Jessica, L. et al. 2014. Integrating multiple lines of evidence into historical biogeography hypothesis testing: a *Bison bison* case study. – Proc. R. Soc. B 281: 20132782.
Jiménez-Valverde, A. et al. 2011. Use of niche models in invasive species risk assessments. – Biol. Invasions 13: 2785–2797.
Johns, T. et al. 2014. Regional-scale drivers of groundwater faunal distributions. – Freshwater Sci. 34: 316–328.
Juan, C. et al. 2010. Evolution in caves: Darwin's 'wrecks of ancient life' in the molecular era. – Mol. Ecol. 19: 3865–3880.
Ketmaier, V. et al. 2013. Molecular divergence and evolutionary relationships among Aemodogryllinae from southern China, Laos and Thailand (Orthoptera, Rhaphidophoridae). – Subt. Biol. 10: 1–25.
Knowles, L. L. et al. 2007. Coupling genetic and ecological-niche models to examine how past population distributions contribute to divergence. – Curr. Biol. 17: 940–946.
Kriticos, D. J. et al. 2012. CliMond: global high-resolution historical and future scenario climate surfaces for bioclimatic modelling. – Methods Ecol. Evol. 3: 53–64.
Kriticos, D. J. et al. 2014. Extending the suite of bioclim variables: a proposed registry system and case study using principal components analysis. – Methods Ecol. Evol. 5: 956–960.
Lahoz-Monfort, J. J. et al. 2014. Imperfect detection impacts the performance of species distribution models. – Global Ecol. Biogeogr. 23: 504–515.
Lamoreux, J. 2004. Stygobites are more wide-ranging than troglobites. – J. Cave Karst St. 66: 18–19.
Lefébure, T. et al. 2006. Phylogeography of a subterranean amphipod reveals cryptic diversity and dynamic evolution in extreme environments. – Mol. Ecol. 15: 1797–1806.
Lisón, F. et al. 2013. Effectiveness of the Natura 2000 Network for the conservation of cave-dwelling bats in a Mediterranean region. – Anim. Conserv. 16: 528–537.
Litwin, L. 2008. Atlas of karst area based on Web GIS technology. – Environ. Geol. 54: 1029–1036.
Lobo, J. M. et al. 2010. The uncertain nature of absences and their importance in species distribution modelling. – Ecography 33: 103–114.

Malard, F. et al. 2009. Diversity patterns of stygobiotic crustaceans across multiple spatial scales in Europe. – Freshwater Biol. 54: 756–776.

Mammola, S. 2017. Modelling the future spread of native and alien congeneric species in subterranean habitats – the case of *Meta* cave-dwelling spiders in Great Britain. – Int. J. Speleol. 46: 427–437.

Mammola, S. and Isaia, M. 2014. Niche differentiation in *Meta bourneti* and *M. menardi* (Araneae, Tetragnathidae) with notes on the life history. – Int. J. Speleol. 43: 343–353.

Mammola, S. and Isaia, M. 2016. The ecological niche of a specialized subterranean spider. – Invertebr. Biol. 135: 20–30.

Mammola, S. and Isaia, M. 2017. Rapid poleward distributional shifts in the European cave-dwelling *Meta* spiders under the influence of competition dynamics. – J. Biogeogr. doi: 10.1111/jbi.13087

Mammola, S. et al. 2015. Alpine endemic spiders shed light on the origin and evolution of subterranean species. – PeerJ 3: e1384.

Mammola, S. et al. 2016a. Ecology and sampling techniques of an understudied subterranean habitat: the Milieu Souterrain Superficiel (MSS). – Sci. Nat. 103: 88.

Mammola, S. et al. 2016b. Step back! Niche dynamics in cave-dwelling predators. – Acta Oecol. 75: 35–42.

Mammola, S. et al. 2017. Climate change may drive cave spiders to extinction. – Ecography doi: 10.1111/ecog.02902

Merow, C. et al. 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. – Ecography 36: 1058–1069.

Michel, G. et al. 2009. Reserve selection for conserving groundwater biodiversity. – Freshwater Biol. 54: 861–876.

Moldovan, O. 2013. One step forward for subterranean biology. – Subt. Biol. 11: 1–2.

Moore, G. W. and Sullivan, G. N. 1964. Out of phase seasonal temperature fluctuations in Cathedral Cave, Kentucky. – Geol. Soc. Am. 76: 313.

Moss, R. H. et al. 2010. The next generation of scenarios for climate change research and assessment. – Nature 463: 747–756.

Muñoz, M. E. S. et al. 2009. openModeller: a generic approach to species' potential distribution modelling. – GeoInformatica 15: 111–135.

Murray, J. V. et al. 2009. How useful is expert opinion for predicting the distribution of a species within and beyond the region of expertise? A case study using brush-tailed rock-wallabies *Petrogale penicillata*. – J. Appl. Ecol. 46: 842–851.

Naimi, B. 2013. usdm: uncertainty analysis for species distribution models. – R package ver. 1.1-12, <http://CRAN.R-project.org/package=usdm>.

Naimi, B. and Araújo, M. B. 2016. sdm: a reproducible and extensible R platform for species distribution modelling. – Ecography 39: 368–375.

Naimi, B. et al. 2013. Where is positional uncertainty a problem for species distribution modelling? – Ecography 36: 1–13.

Năpăruş-Aljančič, M. et al. 2017. GIS analysis to assess the groundwater habitat pollution of black proteus. – Nat. Sloveniae 9: 47–49.

Naranjo, M. et al. 2014. ¿Dónde buscar troglobiontes? Ensayo de una cartografía predictiva con MaxEnt en Gran Canaria (islas Canarias). – Arxius Miscel. Zool. 12: 83–92, in Spanish.

Newbold, T. 2010. Applicatons and limitatons of museum data for conservaton and ecology, with particular attention to species distribution models. – Progr. Phys. Geogr. 34: 3–22.

Niemiller, M. L. and Zigler, K. S. 2013. Patterns of cave biodiversity and endemism in the Appalachians and Interior Plateau of Tennessee, USA. – PLoS One 8: e64177.

Niemiller, M. L. et al. 2017. Evaluation of eDNA for groundwater invertebrate detection and monitoring: a case study with endangered *Stygobromus* (Amphipoda: Crangonyctidae). – Conserv. Genet. Resour. doi: 10.1007/s12686-017-0785-2

Novak T. et al. 2010. Niche partitioning in orbweaving spider *Meta menardi* and *Metellina merianae* (Tetragnathidae). – Acta Oecol. 36: 522–529.

Novak, T. et al. 2014. Cold tolerance in terrestrial invertebrates inhabiting subterranean habitats. – Int. J. Speleol. 43: 265–272.

Olivera, G. de et al. 2014. Evaluating, partitioning, and mapping the spatial autocorrelation component in ecological niche modeling: a new approach based on environmentally equidistant records. – Ecography 37: 637–647.

Ortuño, V. M. et al. 2013. The "Alluvial Mesovoid Shallow Substratum", a new subterranean habitat. – PLoS One 8: e76311.

Owens, H. L. et al. 2013. Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. – Ecol. Model. 263: 10–18.

Pearson, R. G. and Dawson, T. P. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? – Global Ecol. Biogeogr. 12: 361–371.

Pellegrini, T. G. et al. 2016. Linking spatial scale dependence of land-use descriptors and invertebrate cave community composition. – Subt. Biol. 18: 17–38.

Peterson, A. T. 2003. Predicting the geography of species' invasions via ecological niche modeling. – Q. Rev. Biol. 78: 419–433.

Peterson, A. T. 2009. Phylogeography is not enough: the need for multiple lines of evidence. – Front. Biogeogr. 1: 19–25.

Peterson, A. T. and Vieglais, D. A. 2001. Predicting species invasions using ecological niche modeling: new approaches from bioinformatics attack a pressing problem. – BioScience 51: 363–371.

Peterson, A. T. and Soberón, J. 2012. Species distribution modeling and ecological niche modeling: getting the concepts right. – Nat. Conserv. 10: 102–107.

Peterson, A. T. et al. 2011. Ecological niches and geographical distributions. – Princeton Univ. Press.

Phillips, S. J. et al. 2006. Maximum entropy modeling of species geographic distributions. – Ecol. Model. 190: 231–259.

Phillips, S. J. et al. 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. – Ecol. Appl. 19: 181–197.

Pipan, T. and Culver, D. C. 2016. Contributions to speleobiology appearing in Acta Carsologica. – Acta Carsol. 45: 13– 129.

Pipan, T. and Culver, D. C. 2017. The unity and diversity of the subterranean realm with respect to invertebrate body size. – J. Cave Karst St. 79: 1–9.

Pipan, T. et al. 2011. Temperature variation and the presence of troglobionts in terrestrial shallow subterranean habitats. – J. Nat. Hist. 45: 257–273.

Poulson, T. L. 1977. A tale of two spiders. – Cave Res. Found. Annu. Rep. pp. 245–248.

Poulson, T. L. and Lavoie, K. H. 2000. The trophic basis of subsurface ecosystems. – In: Wilkens, H. et al. (eds), Ecosystems of the World 30. Subterranean ecosystems. Elsvier, pp. 231–249.

Proosdij, A. S. et al. 2016. Minimum required number of specimen records to develop accurate species distribution models. – Ecography 39: 542–552.

Qiao, H. et al. 2012. mmweb – an online platform for employing multiple ecological niche modeling algorithms. – PLoS One 7: e43327.
Qiao, H. et al. 2015. No silver bullets in correlative ecological niche modelling: insights from testing among many potential algorithms for niche estimation. – Methods Ecol. Evol. 6: 1126–1136.
Qiao, H. et al. 2016. A cautionary note on the use of hypervolume kernel density estimators in ecological niche modelling. – Global Ecol. Biogeogr. doi: 10.1111/geb.12492
Qiao, H. et al. 2017. Using data from related species to overcome spatial sampling bias and associated limitations in ecological niche modeling. – Methods Ecol. Evol. doi: 10.1111/2041-210X.12832
Raxworthy, C. J. et al. 2003. Predicting distributions of known and unknown reptile species in Madagascar. – Nature 426: 837–841.
Resende, L. P. A. and Bichuette, M. E. 2016. Sharing the space: coexistence among terrestrial predators in Neotropical caves. – J. Nat. Hist. 50: 2107–2128.
Riahi, K. et al. 2017. The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: an overview. – Global Environ. Change 42: 153–168.
Rizzo, V. et al. 2017. Substratum karstificability, dispersal and genetic structure in a strictly subterranean beetle. – J. Biogeogr. doi: 10.1111/jbi.13074
Romero, A. 2009. Cave biology. – Cambridge Univ. Press.
Romero, A. 2012. Caves as biological space. – Polymath 2: 1–15.
Sánchez-Fernández, D. et al. 2016. Thermal niche estimators and the capability of poor dispersal species to cope with climate change. – Sci. Rep. 6: 23381.
Saupe, E. E. et al. 2012. Variation in niche and distribution model performance: the need for a priori assessment of key causal factors. – Ecol. Model. 237: 11–22.
Schneider, K. and Culver, D. C. 2004. Estimating subterranean species richness using intensive sampling and rarefaction curves in a high density cave region in West Virginia. – J. Cave Karst St. 66: 39–45.
Schroeder, B. 2008. Challenges of species distribution modeling belowground. – J. Plant Nutr. Soil Sci. 171: 325–337.
Sendra, A. and Reboleira, A. S. P. S. 2012. The world deepest subterranean community – Krubera-Voronja Cave (western Caucasus). – Int. J. Speleol. 41: 221–230.
Sillero, N. 2011. What does ecological modelling model? A proposed classification of ecological niche models based on their underlying methods. – Ecol. Model. 222: 1343–1346.
Silva, M. S. and Ferreira, R. L. 2016. The first two hotspots of subterranean biodiversity in South America. – Subt. Biol. 19: 1–21.
Sket, B. 2008. Can we agree on an ecological classification of subterranean animals? – J. Nat. Hist. 42: 1549–1563.
Smithers, P. 2005. The early life history and dispersal of the cave spider *Meta menardi* (Latreille, 1804) (Araneae: Tetragnathidae). – Bull. Br. Arachnol. Soc. 13: 213–216.
Smithson, A. 1991. Inter-relatonships between cave and outside air temperatures. – Theor. Appl. Climatol. 44: 65–73.
Soberón, J. and Peterson, A. T. 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. – Biodivers. Inform. 2: 1–10.
Soto-Centeno, J. A. et al. 2015. The importance of late Quaternary climate change and karst on distributions of Caribbean mormoopid bats. – Am. Mus. Novit. 3847: 1–32.
Souza Silva, M. et al. 2011. Trophic dynamics in a Neotropical limestone cave. – Subt. Biol. 9: 127–138.
Stoch, F. and Galassi, D. M. 2010. Stygobiotic crustacean species richness: a question of numbers, a matter of scale. – Hydrobiologia 653: 217–234.
Strecker, U. et al. 2003. Genetic divergence between cave and surface populations of *Astyanax* in Mexico (Characidae, Teleostei). – Mol. Ecol. 12: 699–710.
Syfert, M. M. et al. 2013. The effects of sampling bias and model complexity on the predictive performance of MaxEnt species distribution models. – PLoS One 8: e55158.
Thuiller, W. et al. 2009. BIOMOD – a platform for ensemble forecasting of species distributions. – Ecography 32: 369–373.
Title, P. O. and Bemmels, J. B. 2017. ENVIREM: an expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. – Ecography doi: 10.1111/ecog.02880
Trajano, E. 2007. The challenge of estimating the age of subterranean lineages: examples from Brazil. – Acta Carsol. 36: 191–198.
Trajano, E. and Bichuette, M. E. 2009. Diversity of Brazilian subterranean invertebrates, with a list of troglomorphic taxa. – Subt. Biol. 7: 1–16.
Trajano, E. et al. 2016. Spots of high diversity of troglobites in Brazil: the challenge of measuring subterranean diversity. – Biodivers. Conserv. 25: 1805–1828.
Trzaska, S. and Schnarr, E. 2014. A review of downscaling methods for climate change projections. – United States Agency for International Development by Tetra Tech ARD.
Venarsky, M. P. et al. 2017. Experimental detritus manipulations unite surface and cave stream ecosystems along a common energy gradient. – Ecosystems doi: 10.1007/s10021-017-0174-4
Vörös, J. et al. 2017. Surveying Europe's only cave-dwelling chordate species (*Proteus anguinus*) using environmental DNA. – PLoS One 12: e0170945.
Warren, D. L. et al. 2010. ENMTools: a toolbox for comparative studies of environmental niche models. – Ecography 33: 607–611.
Weckstein, J. D. et al. 2016. Comparative phylogeography of two codistributed subgenera of cave crickets (Orthoptera: Rhaphidophoridae: *Ceuthophilus* spp.). – J. Biogeogr. 43: 1450–1463.
Yackulic, C. B. et al. 2015. To predict the niche, model colonization and extinction. – Ecology 96: 16–23.
Zagmajster, M. et al. 2008. Species richness patterns of obligate subterranean beetles (Insecta: Coleoptera) in a global biodiversity hotspot-effect of scale and sampling intensity. – Divers. Distrib. 14: 95–105.
Zagmajster, M. et al. 2010. Evaluating the sampling bias in pattern of subterranean species richness: combining approaches. – Biodivers. Conserv. 19: 3035–3048.
Zagmajster, M. et al. 2014. Geographic variation in range size and beta diversity of groundwater crustaceans: insights from habitats with low thermal seasonality. – Global Ecol. Biogeogr. 23: 1135–1145.
Zhang, Y. and Li, S. 2013. Ancient lineage, young troglobites: recent colonization of caves by *Nesticella* spiders. – BMC Evol. Biol. 13: 183.
Zimmermann, N. E. et al. 2009. Climatic extremes improve predictions of spatial patterns of tree species. – Proc. Natl Acad. Sci. USA 106: 19723–19728.