On the use of Spatial Ecological Modelling as a tool for improving the assessment of geographic range size of threatened species

SEM and Species geographic range size

Fabio Attorre\textsuperscript{a,*}, Michele De Sanctis\textsuperscript{a}, Alessio Farcomeni\textsuperscript{b}, Alfredo Guillet\textsuperscript{c}, Edoardo Scepi\textsuperscript{a}, Marcello Vitale\textsuperscript{a}, Francesca Pella\textsuperscript{d}, Mauro Fasola\textsuperscript{d}

\textsuperscript{a}Dipartimento di Biologia Ambientale, Sapienza Università di Roma, P. le A. Moro 5 - 00185, Roma, Italy
\textsuperscript{b}Dipartimento di Salute Pubblica e Malattie Infettive, Sapienza Università di Roma, P. le A. Moro 5 - 00185, Roma, Italy
\textsuperscript{c}Unità Tecnica, DGCS, Ministero Italiano degli Affari Esteri, Via S. Contarini 25, 00194 Roma, Italy
\textsuperscript{d}Dipartimento Scienze Terra Ambiente, Università degli Studi di Pavia, Via Ferrata 9, 27100 Pavia, Italy

*Correspondence: Fabio Attorre, Environmental Biology Department Sapienza University of Rome, Italy E-mail: fabio.attorre@uniroma1.it
Abstract
We analysed endemic threatened tree and reptile species of Socotra Island (Yemen), characterised by different ecological requirements and spatial distribution, in order to evaluate the usefulness of spatial ecological modelling in the estimation of species extent of occurrence (EOO) and area of occupancy (AOO). Point occurrences for the entire species range were used to model their spatial distribution by Random Forest (RF) and Generalised Linear Model (GLM). For each species the suitability area (SA) was obtained by applying the 0% omission error criterion on the probability map, and compared or integrated with EOO and AOO area obtained by topological methods such as the minimum convex polygon (MCP), α-hull and 2x2 km grid.

RF showed a lower prediction error than GLM. Higher accuracy was achieved for species with higher number of occurrences and narrower ecological niche. SA was always greater than AOO measured with the 2x2 km grid method. SA was greater than EOO, measured by both MCP and α-hull methods, for species with localised distribution, while it was smaller for widely distributed species. EOO-α-hull area was equal or smaller than that calculated by MCP depending on the spatial distribution of species. AOO measured considering the SA within the EOO-MCP was greater than that measured using the standard 2x2 km grid. Conversely, AOO calculated considering the suitable area within the EOO-α-hull showed variable results, being smaller or greater than the 2x2 km grid AOO depending on the ecological niche and spatial distribution of species. According to our results, SEM does not provide an effective alternative to topological methods for the estimate of EOO and AOO. However, it may be considered a useful tool to estimate AOO within the boundaries of EOO measured by the α-hull method, because it reduces some potential sources of inconsistency and bias.

Keywords: Area of occupancy, Extent of occurrence, Random Forest, Red List Criteria, Socotra, Suitability maps
Introduction

Species conservation status can be assessed according to current international criteria (IUCN 2001, 2010). Among these criteria, population size and trends are difficult laborious to gain, while more informative biological data on ecological requirements, population structure, dispersal mechanism are unknown for many threatened species. For this reasons, most assessments are based mainly on the size of the distributional ranges. This is the case of trees, which account for the majority of plant species assessed in the IUCN Red List of Threatened Species (http://www.iucnredlist.org/) and have been evaluated mainly according to the B Criterion on geographic range size (Lughadha et al. 2005; Newton & Oldfield 2008). The same applies for most animal taxa as well for which distributional range is the only quantitative information available and on which therefore the vulnerability assessment is usually based.

The use of the B Criterion is based on two parameters: extent of occurrence (EOO) and area of occupancy (AOO). EOO is defined as the area contained within the shortest continuous boundary drawn to encompass all sites of occurrence of a taxon. The minimum convex polygon (MCP), the smallest polygon in which no internal angle exceeds 180° and which contains all species’ sites, is the standard method for estimating EOO (IUCN 2001, 2010). Burgman and Fox (2003) discussed the bias of the MCP method when dealing with disjunction and outlying occurrences and suggested that the α-hull method should be preferred for estimating the extent of and trends in species’ ranges. The α-hull should provide a more detailed representation of the space contained by the occurrences, allowing outlying occurrences or poorly represented regions to be removed and so reducing some bias intrinsic in the MCP method (Burgman & Fox 2003).

AOO is the part of EOO occupied by a taxon and, as recommended by IUCN (2010), it is generally calculated by summing up the grid squares occupied by a species, the resolution being dependent on the biological features of the taxon under analysis. For this reason, AOO is strongly dependent on the spatial resolution used (Gaston & Fuller 2009) and the choice of a scale can be a source of inconsistency and bias. In order to improve AOO assessments other topological methods have been proposed (see Hernández & Navarro 2007), even if it has been suggested that in many cases a grid size of 2 x 2 km is an appropriate scale (IUCN 2010).

IUCN (2010) indicated that not only actually known sites but also sites inferred from presence of known appropriate habitat can be used to measure EOO and AOO. Based on this indication, several attempts have been made to also include also potential areas in the estimation of species range size. For instance, Good et al. (2006) used a very simple approach for assessing the extinction risk of plant species of Coleeae tribe (Bignoniaceae) in Madagascar by calculating the AOO, as the
remaining area of primary forest in 2000, within the EOO measured using the MCP method. More recently spatial ecological modelling (SEM) of species was proposed as a tool to support the assessment of EOO and AOO. In particular, based on occurrences obtained from herbarium data, Sérgio et al. (2007) used SEM (based on well-known methods such as GARP, MAXENT and ENFA) as an alternative to topological methods (MCP and \( \alpha \)-hull), to estimate, for Portugal, the EOO of four bryophyte species, with a spatial resolution of 1 km\(^2\). Jiménez-Alfaro et al. (2012), using the Maximum Entropy algorithm and very fine resolution data (15 x 15m), evaluated the AOO of one, rare and endangered vascular plant species (\textit{Empetrum nigrum}) in an area of the Cantabrian range (Spain) and compare it with the results of the grid method. Both studies, although providing useful insights in the application of SEM for the assessment of species range size, were of limited use for attempting generalizations on this topic for two main reasons. First, they were focused on only one of the two parameters (EOO or AOO). Second, the available occurrences allowed the analysis of only a part of the species ecological niche, being limited in one case on an administrative boundary (Portugal) and on the other one on an arbitrary delimited area of the Cantabrian range.

In order to further explore the issue of using SEM as a tool to perform or improve the assessment of EOO and AOO, we used as a case study several endemics of Socotra Island (Yemen): seven tree species and six reptile species with different vulnerability status. This choice turned to be particularly suitable for several reasons: i) the analysed species are characterised by different ecological requirements and spatial distribution; ii) recent and exhaustive surveys of all the known populations of these species are available; iii) the assessment of species endemic to the island corresponds to a global assessment (Gärdenfors et al. 2001).

**Methodology**

*Study area and species*

Socotra is part of a continental archipelago comprising four islands and a few islets in the Indian Ocean between 12°06’-12°42’N and 52°03’- 54°32’E; Abd al Kuri, the westernmost island, lies about 80 km from Cape Guardafui in Somalia, and Socotra about 380 km south of the Arabian Peninsula (Fig. 1). The island with an area of about 3600 km\(^2\) is characterised by an undulating plateau ranging from 300 to 900 m composed of a thin stratum of Cretaceous and Tertiary limestone that overlies an igneous and metamorphic basement (Beydoun & Bichan 1970). On the coastal plains and in the inner depressions, Quaternary and recent deposits of marine, fluvial and continental origin overlie the older substrata.
Local climate is influenced by large-scale weather phenomena, particularly the seasonally reversing monsoon: the NE (winter) monsoon, which lasts from November to March, and the SW (Indian or summer) monsoon, which blows from May to September. Mean annual rainfall and temperature are 216 mm and 28.9°C respectively, but the former varies across the island with exposure and elevation: on the coastal plains annual precipitation may be as little as zero, while in the mountains it can reach exceptional levels of 1000 mm, of which preliminary measurements suggest that fog-derived moisture may constitute over two-thirds of the total (Scholte & De Geest 2010).

The long isolation of Socotra archipelago has contributed to the evolution of unique flora and fauna, with high ratio of endemics (Wranick 2003; Miller & Morris 2004, De Sanctis et al. 2012). The traditional, sustainable land practices have allowed such a “hot spot” of biological diversity to be conserved. The Socotra Island flora hosts 308 endemic flowering plant species, out of an estimated total of 825, a 37.3% level of endemism (Miller & Morris 2004). Among them, several tree species are now threatened with extinction according to the assessment carried out by Miller & Morris (2004) based on IUCN classification system and criteria (IUCN 2001, 2010). Among animals, endemism in the Socotra archipelago is also particularly high for certain taxa, such as the terrestrial reptiles that include 28 endemic species out of 31 (Razzetti et al. 2011, and the new species recently described by Sindaco et al. 2012).

We selected for this study a few endemic tree and reptile species. Among the trees, we analysed some of the most important threatened species: *Dracaena cinnabari, Dendrosicyos socotrana* and five *Boswellia* species: *B. ameero, B. dioscorides, B. elongata, B. socotrana* and *B. popoviana*. *Dracaena cinnabari* is arguably the main flagship species of Socotra. It is one of the six arboreal species (dragonblood tree) of the genus (Marrero et al. 1998). These species are considered remnants of the Mio-Pliocene Laurasian subtropical forests, which are now almost extinct as a result of climate change in the late Pliocene leading to the desertification of North Africa (Quetzel 1978). *Dracaena cinnabari* still provides a reddish resin used in traditional medicine, cosmetics and artisanal pottery (Miller & Morris 2004). *Dendrosicyos socotrana* is one of the few species in the cucumber family (*Cucurbitaceae*) with a tree growth form and it has a noticeable swollen, succulent trunk (Miller & Morris 2004). *Boswellia* spp. represent one of the most interesting examples of adaptive radiation on Socotra (Turelli et al. 2001) and this process has led to the evolution of seven endemic species, giving the relatively small island the highest concentration of frankincense tree species in the world (Thulin & Al-Gifri 1998). *Boswellia* can be classified according to growth form (Miller & Morris 2004): the so-called cliff-rooted species (*B. bullata, B. dioscorides, B. nana, B.
popoviana); and the ground-rooted species (B. ameero, B. elongata, B. socotrana). Since early historical times, frankincense has been collected mainly from two ground species, B. elongata and B. socotrana, and then used for house cleansing and in traditional and veterinary medicine, or exported as a natural product.

Among the reptiles, we selected from different Families the following six endemic species, with varying IUCN vulnerability status. 1) the Socotra Chameleon (Chamaeleo monachus, Chamaeleonidae), classified as “Near Threatened” due to scattered distribution and alteration of the preferred habitat by overgrazing. 2) A lizard (Mesalina balfouri Lacertidae), “Least Concern” because widespread, with a large population size but unknown trend. 3) A skink (Hakaria simonyi, Scincidae) “Near Threatened” due to small area of occupancy, estimated at only 144 km², and to reduction of habitat by overgrazing. 4) The skink Socotra Mabuya (Trachylepis socotrana Scincidae), “Least Concern” because widespread but unknown population size and trend. 5) The snake Günther's Racer (Ditypophis vivax, Pseudoxyrhophiidae), “Least Concern” because widespread but unknown population size and trend. 6) The snake Socotran Racer (Hemerophis socotrae, Colubridae), “Near Threatened” due to habitat fragmentation and increasing road kills.

The main threats to all these species of plants and reptiles are believed to be habitat degradation by overgrazing and overexploitation, landscape fragmentation by new roads, increasing pressure by tourism, increasing drought caused by climate change, invasive species spread and several other impending environmental changes (Attorre et al. 2007a, 2011a; Scholte & De Geest 2010, Van Damme & Banfield 2011, Senan et al. 2012).

Data set

Point occurrence data for plants were obtained by sampling all known localities where the species had been recorded during previous surveys (Thulin & Al-Gifri 1998; Mies et al. 2000; Miller & Morris 2004), and where the occurrence of the species was suggested by the local staff of the Environmental Protection Authority of Socotra. The actual presence was confirmed during field surveys conducted by our team from 2006 to 2009. For the reptiles, we used a data set that included all the recent (after 1950) bibliographic records that could be precisely georeferenced, plus the records collected during five herpetological surveys of the entire island, from May 2007 to 5 April 2010: These surveys were based on 215 diurnal and nocturnal transects, according to a systematic, time-constrained sampling protocol, and conducted by teams ranging from two to seven trained herpetologists. Further transects were specifically aimed to record the presence of nocturnal species that are particularly difficult to detect, like chamaleos (Razzetti et al. 2011).
The number of georeferenced occurrences of each species is given in table 1.

Environmental variables potentially influencing the current spatial distribution of these species were selected in order to map their potential distribution. Climatic, topographical and geological data, in grid format with a high spatial resolution of 100 m, were used as independent variables. As elevation data we used the 3 arc-seconds spatial resolution ‘hole-filled’ SRTM Digital Elevation Model (DEM, Jarvis et al. 2008). Secondary data such as slope and aspect were derived from the DEM using the Spatial Analyst module of ESRI ArcGIS 9.3 software. Climatic maps were obtained by interpolating precipitation and temperature data recorded in 10 manual meteorological stations and calculating the average data for the 2000-2008 period. As interpolating method Universal Kriging with a trend function defined on the basis of a set of covariates (DEM, slope, aspect, distance from the coast), was used. This method produces reliable climatic and bioclimatic raster maps on a regional scale when complex topographical effects are present (Attorre et al. 2007b). Besides mean annual temperature and annual precipitation, a moisture index (Mi) was used based on: Mi = P/PET, where P is the mean annual precipitation and PET is the potential evapotranspiration. A simplified geological map (scale 1:350,000, Beydoun & Bichan 1970) including granitic, limestone, alluvial and sand substrata, was used as a surrogate of a pedological map, which still does not exist for the whole area.

For reptiles we used also a land cover map produced by classifying high-resolution satellite images that includes: coastal vegetation, grasslands, shrublands and woodland (Malatesta et al. submitted).

Data analysis

EOO was calculated using the MCP and the α-hull methods and, as suggested by IUCN (2010) a α value of 2 was chosen. For each species EOO was calculated using the ade-4 and deldir packages in R (http://cran.r-project.org). AOO was estimated using two methods. First we superimposed a grid square of 2 x 2 km resolution on species distribution maps, as recommended by IUCN (2010). Second, by applying an approach similar to that proposed by Boitani et al. (2008) for large-sized African mammals, we inferred potential AOO as estimated by SEM within the boundaries of EOO as measured with topological methods (MCP and α-hull).

As SEM methods we used two well established models: Random Forest (RF) and Generalised Linear Models (GLM). RF has widely proved to be efficient in predicting the spatial distribution of tree species (for more technical details and comparative analyses see also Benito Garzón et al. 2006; Prasad et al. 2006; Cutler et al., 2007; Williams et al. 2009; Scarnati et al. 2009; Attorre et al.
2011b; Barbet-Massin et al. 2012) and GLM is also widely used in species distribution modelling because of its strong statistical foundation and ability to realistically model ecological relationships (Elith et al. 2006, Austin 2007, Guisan et al. 2006, 2007, Thuiller et al. 2009, Williams et al. 2009). The GLM is parsimonious and easily interpretable, even if in ecological modelling it may impose too strong constraints on the model structure, leading to lack of fit. On the other hand, RF may be too flexible and overfit in certain cases. In our approach we used a classical logistic regression model, where a log-odds link for the probability of occurrence is coupled with the assumption that the outcome is distributed like a Binomial random variable with one trial. In order to select the model for prediction, we performed a stepwise selection minimizing the Akaike Information Criterion (Akaike 1974). The stepwise procedure was forward, and it was implemented using the software R (as in all other statistical analyses).

Since the use of presence-only data can bias the analysis and lead to overoptimistic predictions of the potential distribution, both RF and GLM require records of absences. In the case of rare species, absence data are difficult to obtain: a given location may be classified in the ‘absence’ set both when for historical reasons the species is absent, even though the habitat is suitable, or when it is not detected, albeit being present, and when the habitat is truly unsuitable; only the latter is relevant for predictions. When no true absence data are available, one approach is to generate ‘pseudo-absences’ and to use them in the model as absence data for the species. As many pseudo-absences as there are presences are sampled randomly without replacement along a case-control scheme (Rothman 1986). The choice of the method for generating pseudo-absences may influence the performance of the models (Chefaoui & Lobo 2008). In this study a simple random sampling was chosen to allow for pseudo-absences to be regarded as a random sample from the background population and the probabilistic properties of this random sample to be known (see, for instance, Ward et al. 2009).

The distribution maps based on probability were transformed into suitability maps, applying a threshold to each species so as to obtain a 0% omission error, meaning that all the occurrences are correctly predicted (Engler et al. 2004). Even if several methods are available to establish probability thresholds (Willson et al. 2005), we decided to use this criterion to make our results comparable with those of Sérgio et al. (2007) and Jiménez-Alfaro et al. (2012) and because it is the most similar to the basic IUCN principle for estimating species range size, that is to include all the locations for a given species (IUCN 2010).

Similarly to Sérgio et al. (2007) and Jiménez-Alfaro et al. (2012), the goodness of fit of the model for each species was evaluated through the out-of-bag prediction error. For each species we
repeatedly split the data in a test and a training set several times. Each model was fit on the training set and the fitted model was used to predict the response on the test set, thus evaluating the prediction error. The average prediction error obtained after randomly splitting the data 1000 times was used for model evaluation. The size of the test set was allowed to change from species to species, given the wide differences in available total sample size. In particular, let $n_j$ denote the sample size for the $j$-th species, we fixed the size of the test set as max $(0.1n_j, 2)$.

**Results**

**Spatial ecological analysis**

According to the evaluation procedure, prediction error for plant species was smaller than that for reptiles for both RF and GLM (Table 1). RF performed generally better than GLM for both plant species and reptiles. Moreover RF tends to produce suitability maps with a smaller predicted area than GLM (Appendix S1 for some examples). Only in two cases, GLM outperformed RF: *B. popoviana* (plant) and *Dityphophis vivax* (reptile), which are, within their taxonomical group, the species with the lowest number of occurrences.

Once assessed the lower prediction error of RF over GLM, we used RF outputs to produce suitability maps by applying a probability threshold according to the 0% omission error criterion (Fig. 2 a-o). For plants, two ground-rooted *Boswellia* species, *B. elongata* and *B. socotrana*, which are still used for incense production, showed the largest potential distribution (Fig. 2 c, e). Conversely, the third ground *Boswellia* species, *B. ameero*, being limited to the upper Haggheir mountain (Fig. 2 a) and the two cliff-rooted *Boswellia* species, *B. dioscorides* and *B. popoviana*, growing as they do in localised and restricted environments such as very steep escarpments, showed the smallest predicted areas, (Fig. 2 b, d). *Dendrosicyos socotrana* and *Dracaena cinnabari* were characterised by intermediate and almost equal potential distribution areas. The former is more widespread in the lower alluvial valleys, while the latter characterises the limestone plateau of the hilly areas up to the upper granitic Haggheir mountain (Fig. 3 c), sharing a large potential area with *B. elongata*.

The suitability maps for the reptiles, (Fig. 2 h-o) portrayed for all the species wide potential distributions across altitudinal ranges, that matched the actual occurrence from sea level to 900 m a.s.l. (*Dityphophis vivax, Hemerophis socotrae*), to 1000 m (*Chameleo monachus, Mesalina balfouri*), and almost up to the maximum elevations (1200 m for *Trachylepis socotrana* and 1400 m for *Hakaria simonyi*), but with specific preferences for certain areas.
Species range size analysis

For plant species the suitable area (SA) was greater than the EOO and AOO measured with the topological methods (Table 2), the only exception being *Dendrosicyos socotrana*, for which the EOO calculated with both the MCP and α-hull methods was noticeably greater than the potential area. On the other hand, for the reptiles EOO was always larger than the SA (Table 2).

As expected, EOO area, as measured by the α-hull method, was smaller or equal than that calculated by MCP (Table 2). Moreover, for *B. ameero* and *B. popoviana*, EOO measured with the α-hull method turned out to be smaller than the AOO areas measured with the 2x2 km grid method.

Potential AOO considering only the SA within the EOO-MCP was greater than the topological AOO measured using the standard 2x2 km grid method, the only exception being *B. popoviana* which, however, showed quite similar values. Conversely, potential AOO calculated considering the SA within the EOO-α-hull was smaller than the topological AOO for *B. ameero*, *B. dioscorides* and *B. popoviana*, while it was greater for the other plant species (*B. elongata*, *B. socotrana*, *Dendrosicyos socotrana* and *Dracaena cinnabari*), and for all the reptiles (Table 2).

Discussion

As already demonstrated by several comparative studies (Benito Garzón et al. 2006; Prasad et al. 2006; Scarnati et al. 2009; Williams et al. 2009; Attorre et al. 2011), lower prediction error (Table 1) confirmed RF to be particularly suitable for the analysis of the spatial ecological niche of species, because of its characteristics such as bootstrap-resampling, tree averaging and randomization of predictors. Moreover, the capacity of RF for producing suitability maps with a smaller predicted area than GLM (Appendix 1) is a feature to be preferred particularly when dealing with rare and threatened species in order to optimize costs for reintroductions or new field campaigns (Engler et al. 2004; Guisan et al. 2006).

In our study, reptiles showed greater prediction errors than plants, for both RF and GLM (Table 1) despite their larger data set of occurrences and the use of additional land cover variables and this could be explained considering a lower detectability (and consequent lower accuracy of the distribution data) of reptiles when compared to plants.

In accordance with Guisan et al. (2007), the greater prediction error was obtained for species with a smaller number of occurrences, e.g. *B. popoviana* and *Ditypophis vivax*, or for species with several occurrences, but with a wide ecological niche, such as, for instance, *B. socotrana*, which can grow both in very flat area of alluvial valleys and in moderately steep limestone escarpments, and *Trachylepis socotrana* that can be found in most habitats throughout the island (Razzetti et al 2010).
Conversely, the best results were obtained for species with a narrow ecological niche, such as *B. ameero* that grows only on the Haggher mountain on granitic substratum, *Dracaena cinnabari* that has the greatest number of occurrences and grows only above 400 m of altitude on limestone and granitic substrata, and *Hakaria simonyi* that has a peculiar habitat preference for vegetation cover with sparse stones (Razzetti et al 2010).

The application of the probability threshold, based on the 0% omission error criterion, allowed the production of maps of the SA for each species (Fig. 2 a-o). SA was greater than the EOO area calculated using the topological methods, MCP and α-hull, for almost all the plant species (Fig 2 a-e, g, Table 2). This result corresponds with that of Sérgio et al. (2007), who used SEM in addition to MCP and α-hull to estimate the potential EOO of several bryophyte species in Portugal. Conversely, the same result seems to question their expectations that “EOO based on an ecological approach will be low when modelling a taxon with a narrow ecological distribution”. In fact, in our case study, SA was greater than the EOO area measured with both the MCP and α-hull methods, for the plant species which were characterized by a more localised distribution and narrow ecological niche than the reptiles. Conversely, the latter and *Dendrosycios socotrana*, with a wide distribution throughout the entire island, showed a SA smaller than the EOO calculated by both the MCP and α-hull methods (Table 2). These results seem to suggest that SEM cannot be used directly for estimating EOO because species distribution patterns within the boundaries of the study area strongly affect SEM outputs. In other words species with a wide distribution throughout the area tend to have SA smaller than EOO, the opposite for species with localised distribution.

Concern also arises when applying SEM for assessing directly AOO. In fact contrary to Jiménez-Alfaro et al. (2012), who analysed one plant species (*Empetrum nigrum*) in an arbitrary defined area of the Cantabrian range, in our case study, SA was always larger than the AOO area produced by the traditional 2x2 km grid method (Table 2) even when high resolution data were used.

As far as the topological methods used for assessing EOO, our results are also coherent with the conclusions of Burgam and Fox (2003) who suggested the use of the α-hull method instead of the MCP one for assessing EOO when dealing with irregularly distributed species. This is the case, for instance, of *Dendrosycios socotrana, Chamaeleo monachus* and *Ditypophis vivax*, for which the α-hull method is able to avoid the overestimation of their range caused by disjunctions (Fig. 2, Table 2). However, even the α-hull method has some limitations. In fact, for *B. ameero* and *B. popoviana* the suggested value of 2 for parameter α determined an EOO that was smaller than the AOO calculated using the traditional 2x2 km grid method. Since this is inconsistent with the
definition of AOO, which is the area occupied by a taxon within its EOO; two options are available: to modify the α or to reduce the grid resolution. However, being both somewhat arbitrary and independent from the ecological and biological features of the species under analysis, they do not seem entirely satisfactory and are a potential source of analytical inconsistency and bias.

This problem can be better addressed by estimating AOO using SEM predictions within the boundaries of EOO as measured with α-hull, which, as noticed previously is able to better address discontinuities in species distributions. In this way, model-based AOO estimates resulted always smaller than EOO (Table 2). In this respect, our approach can be considered an improvement on that proposed by Good et al. (2006) and Biotani et al. (2008). As in their study, to them potential AOO is constrained within the boundary of EOO so to avoid the problem of overestimation (Hernández and Navarro 2007). But, differently to them, potential AOO is based on a robust model for the analysis of the relationship between species occurrences and environmental variables making it is also less dependent on the number of occurrences, which strongly affect the grid method (see Jiménez-Alfaro et al. 2012).

However, our study case did not produce univocal results when comparing model-based AOO within the boundaries of α-hull EOO with AOO measured with the 2x2km grid (Table 2). Whereas model-based AOO estimates were smaller for species with localised distribution and narrow ecological niche such as *B. ameero*, *B. dioscorides* and *B. popoviana*, they were greater for species with wider distribution and ecological niche such as *B. elongata*, *B. socotrana*, *Dendrosicyos socotrana* and all the reptile species. The variability of these results questions Jiménez-Alfaro et al. (2012), who, on the basis of the analysis of only one and rare species (*Empetrum nigrum*) and in an arbitrary defined area of the Cantabrian Range, concluded that: “estimates of AOO based on fine-resolution distribution models are more robust tools for risk assessment than traditional systems, allowing a better understanding of species ranges at habitat level”.

In conclusion, our results seem to indicate that, contrary to Sérgio et al. (2007) and Jiménez-Alfaro et al. (2012), SEM cannot be used as alternative to traditional topological methods for estimating EOO and AOO. In turn, the most promising application of SEM seems to be the estimation of potential AOO within the boundaries of EOO as measured with the α-hull method. However, before a generalised use of this approach could be definitely suggested, the effect of variable scale resolutions on the assessment of species range size needs to be evaluated (Guisan et al. 2007) since similar problems of scale-area relationships affecting the grid method can be expected (Hartley & Kunin 2003).
However, as the IUCN has identified the grid size of 2 km as a practical appropriate scale for the assessment, it can be suggested that the appropriate scale resolution for the model-based assessment varies from case to case and has to be chosen according to species and study area (Peterson 2006).

In this regard the availability of worldwide environmental data (e.g. topographical, climatic and land cover ones) will allow relevant comparative analyses, which could be also used to eventually review the IUCN thresholds for categorization.

Besides contributing to a more reliable extinction risk assessment, SEM can support in other ways the general process of elaborating strategies for the conservation of threatened species. For instance, by identifying potential sites of occurrence not covered by existing records, it would be possible to optimise the design of new field surveys to compensate for this lack of information before a re-assessment (Guisan et al. 2006, Williams et al., 2009). Moreover according to different climatic scenarios, SEM can be used to produce future potential species distribution maps, functional to the extinction risk assessment (Attorre et al. 2007a).

Acknowledgements

Special thanks go to the personnel of the Environmental Protection Authority of Socotra for their logistic support and help during the field work and in particular to Ahmed Saed, Ahmed Issa, Mohammed Nageeb and Fahmi Bahashwan. The herpetological surveys were accomplished by Edoardo Razzetti, Ugo Ziliani, Roberto Sindaco, Cristina Grieco, Fabio Pupin, and the data analysed by Roberto Sacchi. A special thanks to Denis Walker for the revision of the text. This work was supported by the Socotra Conservation and Development Program funded by GEF-UNDP, DGCS, the Italian Development Cooperation and the Government of Yemen.

References


dragonblood survive the next period of climate change? Current and future potential distribution

conservation strategies for endemic tree species when faced with time and data constraints:
*Boswellia* spp. on Socotra (Yemen). *Biodiversity and Conservation*, 20(7), 1483-1499.

species distribution models: how, where and how many? *Methods in Ecology and Evolution,*
Doi: 10.1111/j.2041-210X.2011.00172.x

Benito Garzòn, M., Blazek, R., Neteler, M., Sánchez de Dios, R., Sainz Ollero, H. & Furlanello, C.

Journal of the Geological Society of London*, 125, 413-446.

Boitani, L., Sinibaldi, I., Corsi, F., De Biase, A., d’Inzillo Carranza, I., Ravagli, M., Reggiani, G.,
based on habitat suitability models. *Biodiversity and Conservation*, 17, 605-621.


Burgman, M.A. & Fox, J.C. (2003). Bias in species range estimates from minimum convex
polygons: implications for conservation and options for improved planning. *Animal


Classification and distribution patterns of plant communities on Socotra Island, Yemen. *Applied

Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R.J.,
Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L., Loiselle, B.A., Manion, G.,
Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S.,
Richardson, K., Schachetti Pereira, R., Schapire, R.E., Soberón, J., Williams, S.E., Wisz, M. &


Table 1
RF and GLM prediction error (mean and 95% confidence interval) for plant and reptile species. The number of pseudo-absences was equal to the number of absences. The estimates are based on 1000 random splits in training and test set.

<table>
<thead>
<tr>
<th></th>
<th>N° occurrences</th>
<th>RF</th>
<th>GLM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Boswellia ameero</em></td>
<td>15</td>
<td>0.08 (0.08-0.11)</td>
<td>0.18 (0.16-0.20)</td>
</tr>
<tr>
<td><em>Boswellia dioscorides</em></td>
<td>17</td>
<td>0.15 (0.09-0.18)</td>
<td>0.17 (0.15-0.19)</td>
</tr>
<tr>
<td><em>Boswellia elongata</em></td>
<td>32</td>
<td>0.17 (0.16-0.20)</td>
<td>0.22 (0.19-0.25)</td>
</tr>
<tr>
<td><em>Boswellia popoviana</em></td>
<td>9</td>
<td>0.23 (0.22-0.39)</td>
<td>0.20 (0.18-0.22)</td>
</tr>
<tr>
<td><em>Boswellia socotrana</em></td>
<td>29</td>
<td>0.19 (0.19-0.21)</td>
<td>0.31 (0.28-0.34)</td>
</tr>
<tr>
<td><em>Dendrosicyos socotrana</em></td>
<td>61</td>
<td>0.12 (0.11-0.16)</td>
<td>0.20 (0.18-0.22)</td>
</tr>
<tr>
<td><em>Dracaena cinnabari</em></td>
<td>78</td>
<td>0.03 (0.02-0.05)</td>
<td>0.18 (0.16-0.20)</td>
</tr>
<tr>
<td>Reptiles</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Chamaeleo monachus</em></td>
<td>40</td>
<td>0.32 (0.30-0.36)</td>
<td>0.34 (0.31-0.37)</td>
</tr>
<tr>
<td><em>Ditypophis vivax</em></td>
<td>30</td>
<td>0.44 (0.41-0.47)</td>
<td>0.41 (0.38-0.45)</td>
</tr>
<tr>
<td><em>Hakaria simonyi</em></td>
<td>47</td>
<td>0.23 (0.20-0.26)</td>
<td>0.25 (0.23-0.28)</td>
</tr>
<tr>
<td><em>Hemerophis socotrae</em></td>
<td>38</td>
<td>0.32 (0.29-0.35)</td>
<td>0.41 (0.38-0.44)</td>
</tr>
<tr>
<td><em>Mesalina balfouri</em></td>
<td>185</td>
<td>0.31 (0.28-0.34)</td>
<td>0.36 (0.33-0.39)</td>
</tr>
<tr>
<td><em>Trachylepis socotrana</em></td>
<td>202</td>
<td>0.35 (0.32-0.38)</td>
<td>0.36 (0.33-0.39)</td>
</tr>
</tbody>
</table>
**Table 2**

Results of topological and spatial ecological analysis. EOO = Extent of Occurrence, AOO = Area of Occupancy, MCP = Minimum convex polygon. All areas are in km$^2$.

<table>
<thead>
<tr>
<th>Species</th>
<th>Probability threshold</th>
<th>Suitable Area (SA)</th>
<th>EOO MCP</th>
<th>EOO α-hull</th>
<th>AOO 2km grid</th>
<th>AOO MCP-SA</th>
<th>AOO α-hull-SA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Plants</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Boswellia ameero</em></td>
<td>0.80</td>
<td>374.2</td>
<td>62.8</td>
<td>7.6</td>
<td>20</td>
<td>46.4</td>
<td>4.7</td>
</tr>
<tr>
<td><em>Boswellia dioscorides</em></td>
<td>0.65</td>
<td>879.2</td>
<td>504.2</td>
<td>59.4</td>
<td>40</td>
<td>236.6</td>
<td>23.2</td>
</tr>
<tr>
<td><em>Boswellia elongata</em></td>
<td>0.76</td>
<td>1738.6</td>
<td>597.1</td>
<td>597.1</td>
<td>68</td>
<td>387.0</td>
<td>387.0</td>
</tr>
<tr>
<td><em>Boswellia popoviana</em></td>
<td>0.63</td>
<td>555.6</td>
<td>50.4</td>
<td>9.9</td>
<td>24</td>
<td>15.6</td>
<td>3.3</td>
</tr>
<tr>
<td><em>Boswellia socotrana</em></td>
<td>0.66</td>
<td>1473.2</td>
<td>881.1</td>
<td>284.2</td>
<td>72</td>
<td>421.4</td>
<td>150.2</td>
</tr>
<tr>
<td><em>Dendrosicyos socotrana</em></td>
<td>0.75</td>
<td>855.2</td>
<td>2038.0</td>
<td>874.9</td>
<td>112</td>
<td>488.5</td>
<td>326.1</td>
</tr>
<tr>
<td><em>Dracaena cinnabari</em></td>
<td>0.76</td>
<td>801.2</td>
<td>476.0</td>
<td>192.6</td>
<td>112</td>
<td>228.5</td>
<td>118.9</td>
</tr>
<tr>
<td><strong>Reptiles</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Chamaeleo monachus</em></td>
<td>0.62</td>
<td>578.6</td>
<td>2694.8</td>
<td>1782.5</td>
<td>148</td>
<td>417.9</td>
<td>326.0</td>
</tr>
<tr>
<td><em>Ditypophis vivax</em></td>
<td>0.60</td>
<td>847.6</td>
<td>2537.4</td>
<td>1578.5</td>
<td>112</td>
<td>717.4</td>
<td>542.7</td>
</tr>
<tr>
<td><em>Hakaria simonyi</em></td>
<td>0.56</td>
<td>518.9</td>
<td>2296.4</td>
<td>1016.6</td>
<td>132</td>
<td>385.7</td>
<td>286.1</td>
</tr>
<tr>
<td><em>Hemerophis socotrae</em></td>
<td>0.55</td>
<td>1105.9</td>
<td>3264.1</td>
<td>3264.1</td>
<td>124</td>
<td>968.1</td>
<td>968.1</td>
</tr>
<tr>
<td><em>Mesalina balfouri</em></td>
<td>0.52</td>
<td>1075.3</td>
<td>3433.1</td>
<td>3433.1</td>
<td>484</td>
<td>996.7</td>
<td>996.7</td>
</tr>
<tr>
<td><em>Trachylepis socotrana</em></td>
<td>0.53</td>
<td>1000.2</td>
<td>3436.4</td>
<td>3436.4</td>
<td>536</td>
<td>916.2</td>
<td>916.2</td>
</tr>
</tbody>
</table>
Figure captions

Fig. 1 - The study area, Socotra Archipelago.

Fig. 2 A-E - Current potential area (light green), occurrences (black triangles) and Extent of Occurrence calculated with the Minimum Convex Polygon method (black connected lines) for *Boswellia ameero* (A), *B. dioscorides* (B), *B. elongata* (C), *B. popoviana* (D) and *B. socotrana* (E).

Fig. 2 F-L - Current potential area (light green), occurrences (black triangles) and Extent of Occurrence calculated with the Minimum Convex Polygon method (black connected lines) for *Dendrosicyos socotrana* (F), *Dracaena cinnabari* (G), *Chamaeleo monachus* (H), *Dityphosis vivax* (I) and *Hakaria simonyi* (L).

Fig. 2 M-O - Current potential area (light green), occurrences (black triangles) and Extent of Occurrence calculated with the Minimum Convex Polygon method (black connected lines) for *Hemerophis socotrae* (M), *Mesalina balfouri* (N) and *Trachylepis socotrana* (O).

Appendix S1 – Suitable areas for *Chamaeleo monachus* A – Random Forest, B – Generalised Linear Model and *Dendrosicyos socotrana* C – Random Forest, D – Generalised Linear Model.
Figure 2
Click here to download high resolution image