

# Modelling terrestrial reptile species richness, distributions and habitat suitability in Saudi Arabia

Abdulaziz S. Alatawi<sup>a,b,\*</sup>, Francis Gilbert<sup>a</sup>, Tom Reader<sup>a</sup>

<sup>a</sup> School of Life Sciences, University of Nottingham, University Park, Nottingham, NG7 2RD, UK

<sup>b</sup> Biology Department, University of Tabuk, Tabuk City, Saudi Arabia

## ARTICLE INFO

### Keywords:

Habitat suitability  
Maxent  
Saudi Arabia  
Species distribution modelling  
Species richness  
Terrestrial reptiles

## ABSTRACT

Species distribution modelling is a powerful tool that can give us ecological insights about species distributions, and potential effects of environmental factors, in poorly known habitats. For the first time the distribution of terrestrial reptiles in Saudi Arabia was modelled, and environmental factors that affect their current distribution and richness investigated. Reptiles are a major vertebrate group in Saudi Arabia and protecting them should be a priority for conservation in such an arid environment. Temperature was the most important of eleven predictors. Maximum species richness of reptiles was predicted in the central plateau, north-western borders, and in coastal areas of Saudi Arabia. Overall, the predicted and the observed patterns of species richness followed a similar pattern. Our analysis revealed that large scattered parts of Saudi Arabia are considered under-sampled in terms of sampling efforts of terrestrial reptile species. Our results represent the most comprehensive description of terrestrial reptile diversity distributions and habitat suitability in Saudi Arabia to date.

## 1. Introduction

How and why species are distributed in space and time is considered a fundamental scientific question, especially in ecology (Rushton et al., 2004; Guisan and Thuiller, 2005; Brotons, 2014). Humans have observed the unique relationship that connects living organisms with their surrounding environment (Guisan and Thuiller, 2005; Elith and Leathwick, 2009), and for more than a century scientists have tried to come up with theoretical or experimental explanations for spatial and temporal patterns of presence/absence and abundance (Elton, 1927; MacArthur and Wilson, 1967). Today, many important scientific applications depend on understanding this relationship, such as natural resource management, predicting the impacts of climate change and invasive species, and conservation planning more generally (Guisan and Zimmermann, 2000; Guisan and Thuiller, 2005; Peterson et al., 2011). However, despite considerable progress in methods for understanding spatial and temporal patterns, our knowledge of contemporary levels of biodiversity and the distribution of species richness in many parts of the world remains poor (Mora et al., 2011; Ficetola et al., 2013).

Species richness is a key measure of biodiversity and its study can illuminate important ecological phenomena, such as the species-area relationship, and rules for community assembly (Brown et al., 2007). Species richness is connected with the quality and quantity of available habitat, and habitat degradation can adversely affect species richness

(Yi et al., 2016). The distribution of species richness is influenced by many biotic and abiotic factors; understanding and measuring these factors and how they influence the distribution of species is critical for proper conservation planning (Brown et al., 2007). However, it is not an easy task to model the spatial distribution of the richness of species, especially across large-scale areas (Pineda and Lobo, 2009). The gap in current knowledge about the distribution of species richness and the suitability of habitats could be fixed with enough biological data obtained from surveys, and accurate evidence of spatial patterns in environmental variables; these can be combined to produce predictive maps of species richness (Lobo et al., 2002; Graham and Hijmans, 2006). These descriptive maps of species richness are considered a basic tool in conservation decision-making (Lobo et al., 2002; Graham and Hijmans, 2006).

Techniques for modelling species distributions have developed rapidly in the last two decades (Guisan and Zimmermann, 2000; Guisan and Thuiller, 2005; Elith and Leathwick, 2009). Correlative species distribution models (SDMs) (also called ecological niche modelling, more details in Sillero, 2011) usually work by combining known species occurrence records with spatial layers of digital environmental data, and using statistical techniques to estimate geographical distribution and habitat suitability for species over an area of interest (Guisan and Zimmermann, 2000; Guisan and Thuiller, 2005; Sillero, 2011). The main objective is to generate new predicted distribution maps that

\* Corresponding author. School of Life Sciences, University of Nottingham, University Park, Nottingham, NG7 2RD, UK.

E-mail addresses: [mbxaa24@nottingham.ac.uk](mailto:mbxaa24@nottingham.ac.uk), [abalatawi@ut.edu.sa](mailto:abalatawi@ut.edu.sa) (A.S. Alatawi).

capture and explain the species-environment association (Guisan and Zimmermann, 2000; Guisan and Thuiller, 2005) which, if presence-only models are used, effectively represents habitat suitability for the target species (see Sillero, 2011). These maps show how habitat suitability varies through space and time, by identifying locations in the study area that have environmental conditions similar to those where the species has been found (Pearson, 2008; Sillero, 2011). However, despite considerable progress in the development of these techniques, predicting species distributions is still a complex and challenging process (Merow et al., 2014).

Many advanced modelling techniques/tools and quantitative methods have been developed and applied in distribution modelling, including various forms of regression analysis (generalized linear models [GLMs], generalized additive models [GAMs], and machine-learning methods Maximum Entropy [Maxent], and artificial neural networks [ANN]) (Peterson et al., 2011). All of these methods can be used to model potential species richness, but each uses a different process to generate the final distribution map. Some modelling techniques require presence-absence data (e.g., GLM) while others require only presence data as they can generate background points during modelling (e.g., Maxent) (Phillips et al., 2006). Choosing which method to apply depends on the available data and the purpose of the study (Elith and Leathwick, 2009).

Reptiles are widely distributed around the world and are considered important components of local ecosystems and biodiversity (AbuZinada et al., 2004; Carranza et al., 2018). Due to their ectothermy, reptiles are sensitive to the thermal characteristics of their environment (Wilms et al., 2011; Carranza et al., 2018), and they demonstrate various special adaptations as a result (Metallinou et al., 2015). Because their response to temperature variation is relatively well understood, they can be used as indicators to assess the impact of global warming on the environment (Wilms et al., 2011). Today, many reptiles are seriously threatened by many factors, such as habitat loss and conversion, invasive species and collection for the pet trade (Cox et al., 2012), all of which can negatively affect their spatial distribution.

The spatial distribution of reptiles is affected by various climate and topographical factors. Different studies have reported a variety of factors contributing to the explanation of reptile distributions, including precipitation (Fattahi et al., 2014; Sanchooli, 2017), temperature (Sillero and Carretero, 2013; Javed et al., 2017), altitude (El-Gabbas et al., 2016) and vegetation cover (Fattahi et al., 2014). However, among all these factors, temperature appears to dominate (directly and/or indirectly), which is not surprising since it is well known in affecting daily activities and reptile biology (Huey, 1982; Wilms et al., 2011). The extinction risk of most reptiles has not been evaluated properly yet (Carranza et al., 2018).

Reptiles are a major group of vertebrates in Saudi Arabia, with a list of 103 species having been recorded (AbuZinada et al., 2004). Reptiles are the second most diverse terrestrial vertebrate group in Saudi Arabia after birds (432 spp.), with 25% more species than the mammals (79 spp.) (AbuZinada et al., 2004). General knowledge about reptile ecology in Saudi Arabia is poor, and much work is required to understand fully their distribution in space and time (AbuZinada et al., 2004). Saudi Arabia occupies most of the Arabian Peninsula, encompassing a wide range of topographic and climatic conditions (AbuZinada et al., 2004; Vincent, 2008; Gosling et al., 2011). The richness and occurrence of reptiles in Saudi Arabia are expected to vary from one location to another in response to this variation. Unfortunately, published data and museum records for Saudi reptiles are scarce. The few surveys of Saudi reptiles conducted thus far have typically been biased towards easily accessible areas, near major cities, and have focussed on certain taxa of interest to particular researchers. Such bias may mean that the community is under-sampled, and species new to the science may have been entirely missed. Therefore, mapping and modelling of reptile distributions, and studies of the factors that may affect them, are urgently required before relevant conservation measures can be enacted.

Thus far, there have only been two attempts (to our knowledge) to estimate the distribution of reptile diversity in Saudi Arabia. Cox et al. (2012) performed a regional assessment for IUCN Red-listing of the majority of the reptiles of the Arabian Peninsula. By overlaying simple range-distribution polygons for each species, they predicted that high species-richness locations in Saudi Arabia would include the long chain of south-western mountains, the coast of the Eastern Province and a few locations in northern areas. Locations expected to have the lowest species richness were in the Empty Quarter (the Rub' al Khali), and central southern areas. A more detailed species distribution model was developed by Wilms et al. (2011) for just one species in Saudi Arabia - *Uromastyx aegyptia microlepis* - as part of their studies on its thermal biology and activity patterns.

Our objective in this paper is to model the distribution of reptiles in general, using presence-only data, and map their potential richness, and their habitat suitability in Saudi Arabia using Maxent. Maxent is capable of capturing complex non-linear response curves using a set of "features" (e.g. linear, quadratic, product, threshold; see Merow et al., 2013; Phillips et al., 2017). The distribution predictions made by Maxent are known to perform well, its regularization procedure obviating over-fitting (Merow et al., 2013). We built models of the spatial distribution of individual reptile species as a function of environmental variables in order to understand the magnitude of the impact of these variables on the species concerned. We then combined these models to produce two maps to describe predicted patterns of diversity across Saudi Arabia. One map shows the summed probability of species occurrence, and the other shows the expected species richness assuming a threshold rule to convert continuous habitat suitability values into a binary prediction of suitable/not suitable. The predictions that we produce represents the most comprehensive attempt to understand the current distribution of terrestrial reptiles in Saudi Arabia to date.

## 2. Methods and materials

### 2.1. Study area and presence records

Saudi Arabia is a large [2,149,690 km<sup>2</sup>] hyper-arid country (Fig. 1) (Vincent, 2008) with a list of 103 reptiles (at least 85 of them being terrestrial lizards and snakes) that have been recorded, comprising 60 lizards, 34 snakes, and 9 turtles (AbuZinada et al., 2004). For this study, we obtained records of occurrences of all terrestrial reptiles (lizards and snakes only) from museums, and relevant literature. An extensive georeferencing of locations was carried out. Hence, results are presented for a total of 62 reptile species (41 lizards and 21 snakes) with 3943 point records. The currently accepted names of the reptiles were checked using The Reptiles Database (The Reptiles Data-Base, 2017), and Catalogue of Life (Catalogue of Life, 2018). Overall, the geographical coverage of the sample points is satisfactory (Fig. 1). However, as is expected with museum data and non-systematic surveys (Newbold, 2010), the data show an unavoidable bias toward cities, road networks, and easily accessible areas. To minimise the impact of bias in our model predictions, we employed a target-group bias file (Phillips et al., 2009). This uses the sum of all the records of the taxon of interest as an estimate of sampling effort, smoothed using a Gaussian kernel estimation function (using the SDMtoolbox in ArcGIS: Brown et al., 2017). The bias file is used in Maxent to promote the selection of more background points from biased locations. It is a common approach to correct sampling bias used in distribution modelling and is known to improve model accuracy (Phillips et al., 2009).

### 2.2. Environmental variables

The environmental variables used in this study were downloaded from WorldClim Version 2.0 (19 bio-layers and 12 indices of solar radiation), and Version 1.4 (altitude) ([www.worldclim.org](http://www.worldclim.org)) (Hijmans et al., 2005; Fick and Hijmans, 2017). Because of the massive extent of

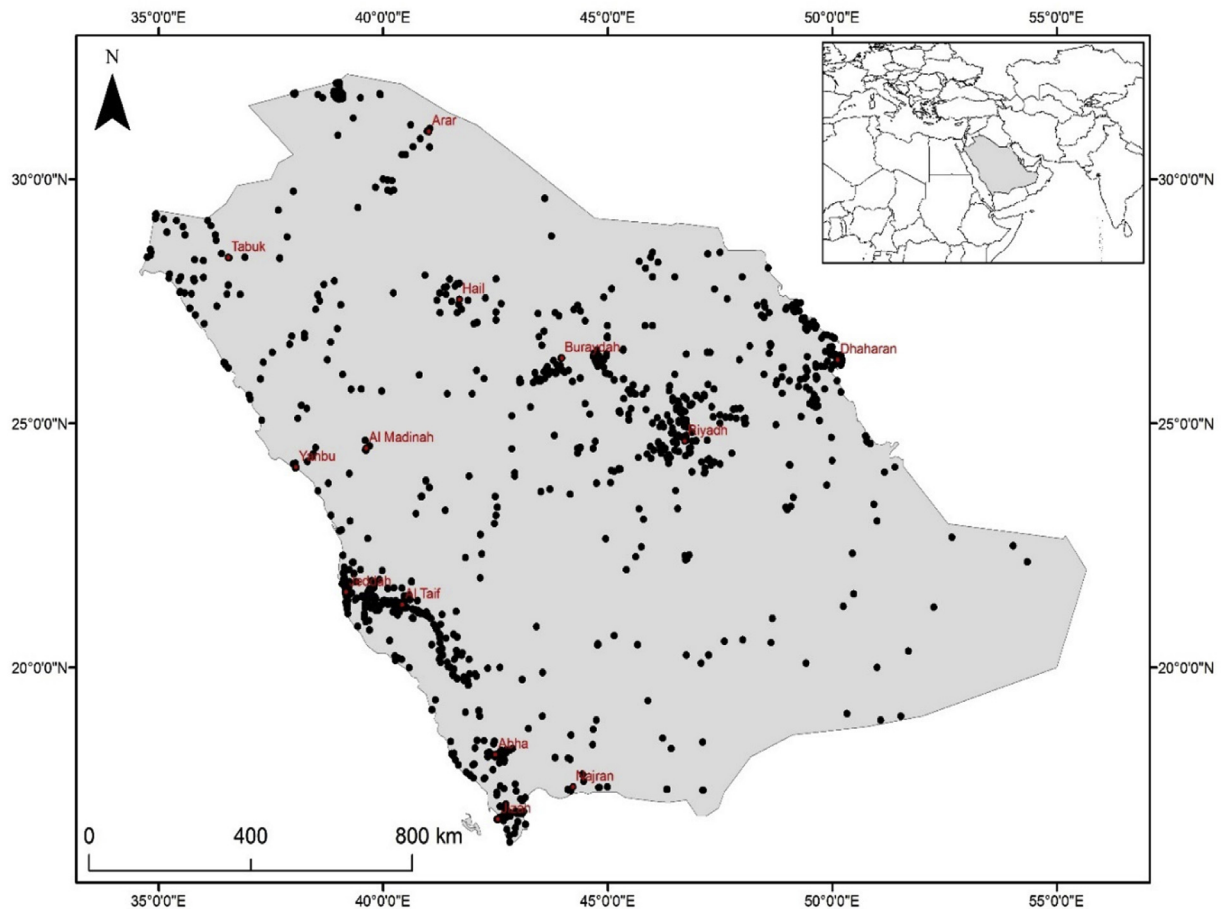


Fig. 1. The locations in Saudi Arabia of the observations of 62 terrestrial reptile species ( $n = 3943$  observations in total) that were used to build distribution models.

Table 1

Bioclimatic variables with a Variance Inflation Factor (VIF) of less than 10 which were used to build the reptile distribution models using Maxent. The last column shows the averaged percent contribution of each environmental variable to the final distribution models of all the species.

Variable	Meaning	% contribution in the models
Alt	Altitude	4.8
Bio_02	Mean Diurnal Range (mean monthly (max temp - min temp))	32.2
Bio_03	Isothermality (BIO2/BIO7) (* 100)	14.5
Bio_08	Mean Temperature of Wettest Quarter	3.2
Bio_09	Mean Temperature of Driest Quarter	1.6
Bio_14	Precipitation of Driest Month	4.8
Bio_15	Precipitation Seasonality (Coefficient of Variation)	12.9
Bio_19	Precipitation of Coldest Quarter	12.9
Srad_04	Solar radiation in April	3.2
Srad_06	Solar radiation in June	9.7
Srad_09	Solar radiation in September	0

the study area and relatively low number of records, environmental variables were downloaded at 30-sec resolution ( $\sim 1 \text{ km}^2$ ). Stepwise selection of environmental variables was performed by calculating the Variance Inflation Factor (VIF) using the library *usdm* in R (Naimi et al., 2014; R Development Core Team, 2018), in order to reduce multicollinearity among variables. Among the 32 original variables, only 11 with a VIF less than 10 were retained for modelling (Table 1). Vegetation cover was also ignored because it is very sparse in such an environment, and has been shown to be a poor predictor of reptile distributions in similar studies (see El Gabbas et al., 2016 for Egypt). All

the environmental variables were prepared by clipping them to the size of the study area and converting them to *ascii* files using ArcGIS software (Version 10.3.1).

### 2.3. Species richness distribution modelling and evaluation

We created distribution models using Maxent Version 3.4.1, downloaded from [https://biodiversityinformatics.amnh.org/open\\_source/maxent/](https://biodiversityinformatics.amnh.org/open_source/maxent/) (Phillips et al., 2006). Maxent identifies areas that have similar values on environmental variables to those at the locations of recorded occurrences. By default, Maxent uses 10000 background points chosen randomly from the study area. It is the most popular software to model species distribution with presence-only data (Merow et al., 2013). We chose Maxent to run the distribution modelling because: i) it is very robust, producing reliable models from presence-only data; ii) it works with low numbers of records; and iii) it appears to be relatively insensitive to sample bias (see Baldwin, 2009). Maxent has been shown to perform very well in comparisons with other modelling approaches (Baldwin, 2009; Merow et al., 2013).

We chose features and settings in Maxent that maximized the area under curve (AUC), on the basis of extensive testing. We therefore used: product and hinge features; ten subsampled replicates; 1000 iterations; and the new default *cloglog* output (for more details about *cloglog*, see Phillips et al., 2017). The replicates were produced by partitioning the data into 90% training and 10% testing subsets; and we used the '10% training presence' as the threshold rule to convert continuous habitat suitability into dichotomous suitable/unsuitable values (see Liu et al., 2005). The permutation importance table was used to determine the importance of each variable and the degree it contributed to the model. We also considered other approaches to test the model, including

partitioning data into 70% training and 30% testing subsets, and cross-validation with ten replications (see supplementary materials) to investigate the final patterns of habitat suitability distribution.

#### 2.4. Model evaluation

To evaluate the accuracy of each model, we used the threshold-independent Area Under the Curve (“AUC”) calculated from the “Receiver Operating Curve” (ROC) (Fielding and Bell, 1997; Pearce and Ferrier, 2000). The ROC curve plots sensitivity against 1-specificity for all possible values of the threshold habitat suitability above which the habitat is assumed to be suitable (Fielding and Bell, 1997). The AUC can be interpreted as the probability that a randomly chosen presence site will be more highly ranked than a randomly chosen absence one (Pearce and Ferrier, 2000; Merow et al., 2013).

The ten subsampled replicate maps for each species were reduced to a single average map by taking the arithmetic mean, or by creating a consensus binary map (see below). Then we combined the species distribution predictions for each modelled species, creating two maps summarising the distribution of terrestrial reptile diversity in Saudi Arabia: i) a probability map which shows the sum of the predicted habitat suitabilities for all species in each grid cell; and ii) a binary map which shows the “predicted potential species richness” based on the conversion of the continuous surface of habitat suitability for each species into a dichotomous surface (suitable = 1 and not suitable = 0). Both maps were created in ArcGIS (10.3.1) using the Raster Calculator and the Reclassify Tool. To create the probability map, all the averaged *ascii* maps for each species were added together to produce a final map of potential species richness. To create a consensus binary map from the 10 replicate maps for each species, the average of ten *ascii* thresholds for each species were added together, and a species was considered to be present in a pixel if more than five replicates of the model predicted its presence. Then, all the species binary maps were added together to create a final consensus species-richness map.

#### 2.5. Identifying areas of low sampling effort

We measured the degree of sampling in areas that probably have been under-sampled, and areas with low/moderate to high sampling effort in relation to the predicted habitat suitabilities, by calculating the difference between predicted and observed species richness at each grid square cell. To do so, 50 × 50-km grid cells were created across the study area and the observed species richness calculated for each grid cell. Next, each individual thresholded species map (i.e., the consensus binary map at a resolution of ~ 1 km) was checked and if any suitable cell was detected. Then 50 × 50 cell was marked as positive. The species maps at the large cell size were then added together to generate a new thresholded species richness map that matches the observed richness resolution.

### 3. Results

#### 3.1. Model performance and variables contribution

The distribution models provided good predictions for most reptile species, with AUCs ranging from 0.68 to 0.99 (Fig. 2). For 22 reptile species, models had excellent performance, with mean AUC > 0.90. The number of records used to create the best performing models varied considerably in size (10–79), and there was a significant negative relationship between the number of observations and mean AUC (Pearson's correlation:  $n = 62$ ,  $r = -0.326$ ,  $P = 0.01$ ). The species with highest mean AUC was *Scincus hemprichii* (0.998), and the species with lowest was *Stenodactylus doriae* (0.679).

The contribution of each environmental predictor to models for all 62 reptile species modelled is shown in Fig. 3. Temperature variables were the most frequently retained as significant predictors of the spatial

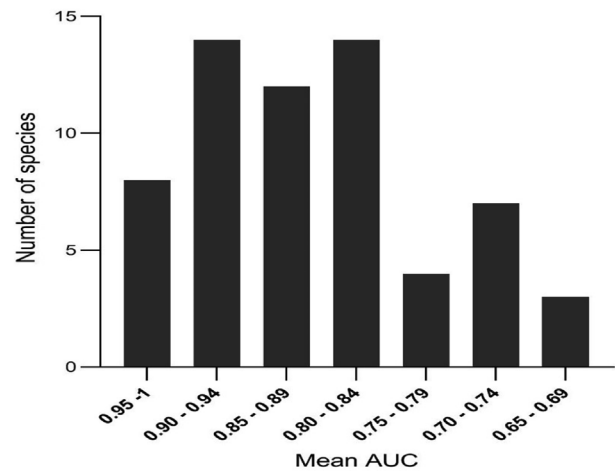


Fig. 2. The frequencies of mean AUC values for reptile distributions models in Saudi Arabia.

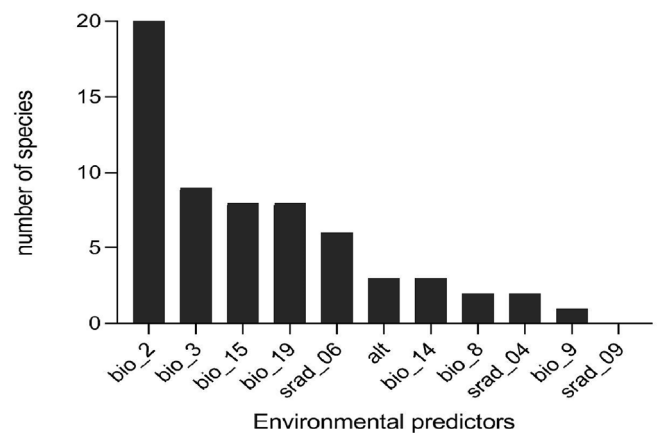
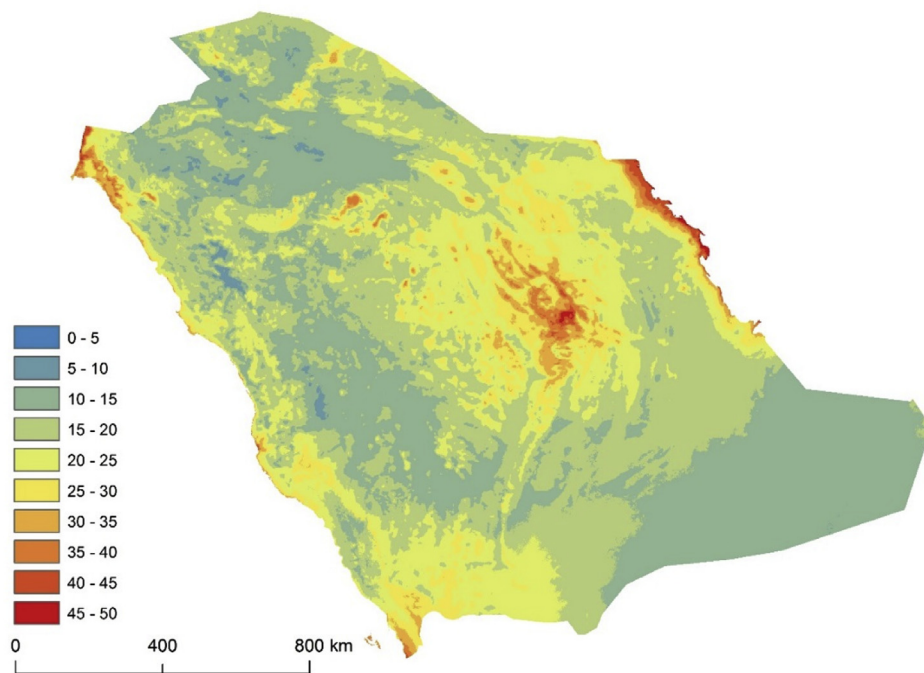


Fig. 3. The frequencies with which each environmental predictor was the most important contributor to the models of 62 Saudi reptile species. The variable contributions were assessed using permutation importance in the MaxEnt output. Abbreviations are explained in Table 1.

distribution of terrestrial Saudi reptiles. Overall, bio2 (mean diurnal temperature range) was most frequently the highest contributor to the models (20 species). The frequency of other predictors contributing to the models was moderate (see Fig. 3 & Table 1). Altitude had a low contribution to the models, while solar radiation in September never contributed significantly (Fig. 3 & Table 1). Full details about variable contributions and relevant information are provided in the supplementary material.

#### 3.2. Current patterns of species richness

The current range of spatial distribution of potential species richness and habitat suitability (Figs. 4 and 5) shows that large areas of Saudi Arabia are expected to have moderate to high species richness of reptiles. Overall, the two maps show a broadly similar range pattern. The central plateau of Saudi Arabia has high predicted species richness. This area covers from south of Riyadh Province up to Al-Qaseem Province. There is moderate to high predicted species richness in the area between the central plateau and the Eastern Province (eastern coast and surrounding areas). There are some fragmented locations in other parts of Saudi Arabia that have high predicted species richness, including the extreme north-west, particularly in Tabuk Province. Finally, the distribution models also predict that western areas surrounding Jeddah city, the long Asir mountain chain, and the south (Jizan and Najran)

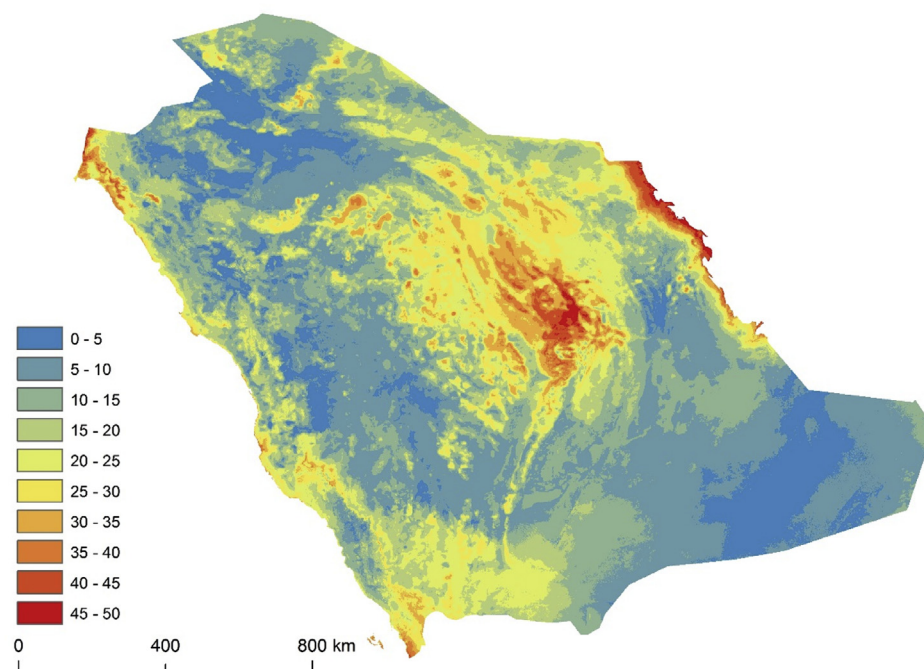


**Fig. 4.** Predicted terrestrial reptile species richness for Saudi Arabia, based on summed habitat suitabilities from individual distribution models for 62 species. The map was generated by summing all averaged *ascii* maps for each species, and then colours rescaled to match the range of Fig. 5. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

have moderate species richness. The predicted pattern declines from the northern edge of the Empty Quarter to the border with Oman, and is low between the central region and the western coast, apart from a few scattered locations. The northern border of Al-Jawf Province along the Jordanian border also has low predicted diversity except in a few fragmented locations. Interestingly, the other two partitioning approaches generated similar potential species richness and habitat suitability distributional patterns (see Figs. S1 and S2).

For the individual species, Maxent predicted distributions very well despite variable sample sizes. Several groups of reptile species had similar distribution patterns (see Supplementary Material Fig. S3). For example, Bosk's Fringe-fingered Lizard (*Acanthodactylus boskianus*)

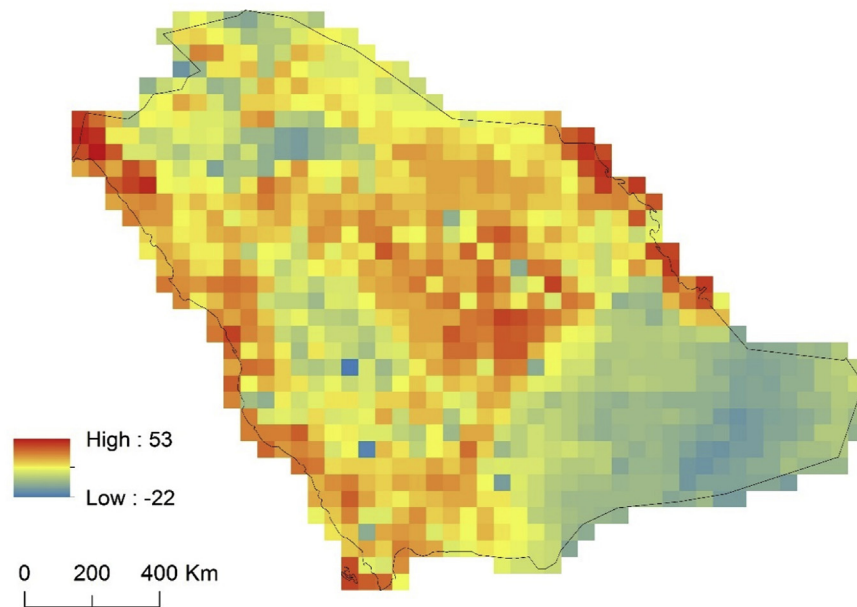
(n = 222 records), and Arabian horned viper (*Cerastes gasperettii*) (n = 155) were predicted to have high habitat suitability in the central plateau, along the coastlines, and in fragmented locations in the north and south. Other species such as Sandfish skink (*Scincus scincus*) (n = 62) and Small-spotted Lizard (*Mesalina guttulata*) (n = 42) had medium to high habitat suitability across most of the country. Even though in general the Empty Quarter had low habitat suitability, for some species it was predicted to be very suitable; for example, Arabian sand skink (*Scincus mitranus*) (n = 172) and Arabian sand boa (*Eryx jayakari*) (n = 96).



**Fig. 5.** Predicted terrestrial reptile species richness for Saudi Arabia, based on summed binary predictions of suitable habitats from individual distribution models for 62 species. The map was generated by averaging the threshold maps for each reptile species, and then summing the number of species predicted to be present in each pixel.



**Fig. 6.** Map of Saudi Arabia (50 × 50 km grid cells) representing the difference between the predicted thresholded species richness and the actual observations from the records. Negative values mean that the observed species richness was higher than the predicted. High values mean that the predicted number of species richness at these cells was larger than the observed.



### 3.3. Sampling effort across Saudi Arabia

Overall, the predicted and the observed species richness broadly followed similar patterns. The map of their differences (Fig. 6) shows that a large portion of the extreme north-west, the west coast to south-western areas, the majority of the central areas, and the east coast and its surroundings appear to have been greatly under-sampled because the predicted species richness recorded species richness was noticeably higher than the observed value. In contrast, areas corresponding to small differences between the predicted and the observed species richness were identified as above the central areas toward the borders with Kuwait, Iraq, and Jordan, with scattered locations in the mid-west, and the whole of the south-west.

## 4. Discussion

Our distribution models appear to predict the distribution of reptiles in Saudi Arabia well, since 35% of species had excellent performance (AUC > 0.90) and only 4% were poor (AUC < 0.70). Diurnal temperature range was the most effective environmental predictor in explaining the current distribution of reptiles. The models suggest that terrestrial reptile habitat suitability and their potential richness is highest in the eastern part of the central plateau of Saudi Arabia, in the mountain chains of the southern border, and around the coasts, especially in north-western and eastern areas. The degree of under-sampling was identified in relatively large parts of Saudi Arabia from large differences between the observed and the predicted species richness.

Knowing which environmental variables contribute most to distribution models for reptiles in Saudi Arabia is important in such an extreme environment, where diurnal temperature ranges are large, and drought is prevalent most of the year. Ecologically, it is unsurprising that diurnal temperature range contributed most to the models, given the ectothermy of reptiles (Huey, 1982; Wilms et al., 2011) and the nature of the environment (AbuZinada et al., 2004; Vincent, 2008). Temperature is expected to be a limiting factor in reptile distribution because of their behavioural and physiological characteristics (Huey, 1982; Wilms et al., 2011). Our maps predict that reptile species richness in Saudi Arabia is expected to be high in areas with moderate diurnal temperature ranges, but lower in areas which experience very high temperatures.

Other studies of reptiles have found similar results to ours in

different parts of the world, even where typical environmental characteristics are very different from those in Saudi Arabia. For example, Javed et al. (2017) found that mean diurnal temperature range (bio2) was one of best predictors of the distribution of the Indian Golden Gecko (*Calodactylodes aureus*) in India. In Eastern Asia and Taiwan, Ananjeva et al. (2015) found that temperature variables were amongst the most effective predictors of the distribution of Square-headed Cat Snake (*Boiga kraepelini*). In Europe, Sillero and Carretero (2013) found that temperature was an important predictor variable in their model of the distribution Carbonell's wall lizard (*Podarcis carbonelli*) on the Iberian Peninsula.

The predicted current distribution of reptiles generated by Maxent suggests that large areas of Saudi Arabia have habitats that are reasonably suitable for reptiles, even though they have not been surveyed. However, the predicted areas with highest potential species richness tend to be concentrated around locations with many known species occurrences. Sillero and Carretero (2013) reported similar results because their model predicted larger suitable areas over the Iberian Peninsula than currently known for *P. carbonelli*, but relatively concentrated around observed occurrences. In our model, predicted reptile species richness is highest in the centre and around the coast of Saudi Arabia. A similar coastal pattern was reported in North Africa by Kaliontzopoulou et al. (2008), and around the coastline of the Arabian Peninsula by Cox et al. (2012). Along these coastal areas in Saudi Arabia there are major cities (e.g., Jeddah, Jizan, Al-Khoabr, Al-Jubil), near which there are a lot of occurrence records.

The predicted pattern of high reptile diversity and habitat suitability around the coast of Saudi Arabia is in-line with the findings of El-Gabbas et al. (2016), who modelled reptile species richness and habitat suitability in Egypt. The authors found that some coastal areas of Egypt had high predicted species richness, especially around the Sinai Gulf and Mediterranean Sea. Our model predicted that some habitats in Saudi Arabia would be suitable for at least 50 out of 62 species, which is similar to the findings of El-Gabbas et al. (2016), who found that some locations in Egypt would be suitable for at least 52 out of 75 modelled species.

As well as predicting high potential species richness in coastal areas, our models predicted high potential species richness in the central areas of the Riyadh and Al-Qassem areas. The areas with high predicted potential species richness are often near centres of human population. This is a pattern seen in other studies of terrestrial reptiles in arid

environments (e.g. in Oman by Carranza et al., 2018), and could an artefact of sampling bias. However, we have attempted to control for sampling bias by using a bias file. Alternatively, the association between centres of human habitation and predicted reptile species richness may reflect the fact that human activity is associated with more food-rich or otherwise suitable environments, either because humans choose to live in areas which are benign or productive, or because humans influence reptile distributions positively by acting as a source of suitable habitat or resources (e.g. farmland, which can be exploited by reptiles) (Tytar et al., 2015; Carranza et al., 2018). A study by Tytar et al. (2015) modelled reptile distributions in the Western Podillya (Ukraine), and found that the human “footprint” contributed significantly (positively) as a predictor in the distribution model. Carranza et al. (2018) in Oman found some areas with high predicted species richness around the capital Muscat; having sampled thoroughly, they concluded that reptiles are not badly affected by the presence of human settlements and may actually benefit. Finally, the combination of low to moderate altitude and suitable temperatures most of the year along the coasts (AbuZinada et al., 2004; Vincent, 2008), which are relatively densely populated by humans, may provide good conditions for terrestrial reptiles in Saudi Arabia, including both generalist and specialist species.

The analysis of under-sampling suggests that even areas that are easily accessible and not isolated have received low sampling effort in terms of terrestrial reptile detection, which unfortunately represents large parts of the Saudi Arabian landscape (Fig. 6). This is a very important result that can help direct future sampling effort to provide better systematic conservation planning. A study by Sillero et al. (2009) used a similar approach to identify potential under-sampled areas in the Iberian Peninsula (see also Sillero et al., 2014).

We found that altitude in Saudi Arabia was not an important predictor of reptile distributions, which contradicts the findings of a similar study in Egypt (El-Gabbas et al., 2016). However, climate-related variables are generally known to influence reptiles more than topographical variables (Guisan and Hofer, 2003). According to Chettri et al. (2010), species richness of reptiles typically gradually declines with elevation (no species were recorded above 3000 m in their study), but the distribution of reptiles in Saudi Arabia may be restricted by factors other than climate and topography, such as interspecific competition (Chettri et al., 2010). Locations at high altitude in Saudi Arabia are not well sampled, and this makes it difficult to speculate about the reason why altitude is not an important predictor in this study. This underlines the need for more ground-truthing surveys in less sampled areas.

Since most of Saudi Arabia lacks systemic surveys for reptiles, and most of the available records were gathered from opportunistic observations (e.g., museum records), some species may have been completely overlooked, including perhaps species that are new to science. Our model predictions cover large areas that have not been surveyed at all (to our knowledge), emphasising the need for more ground-truthing surveys to test predictions and reveal the true extent of Saudi reptile diversity. For example, we have found that the Empty Quarter is not very suitable for terrestrial reptiles, in line with Cox et al. (2012), who reported similar patterns about the Empty Quarter. However, more sampling and ground-truthing is required to establish if diversity is really that low.

We have used a target-group bias file in Maxent to minimise the effect of the usual biases in museum data (Phillips et al., 2009). Biased observations are known to have an effect on the predictive quality of the model (Phillips et al., 2009; Fourcade et al., 2014). Corrections for any bias depend on the type and intensity of bias (Fourcade et al., 2014) in addition to species characteristics (Hernandez et al., 2006), and the best methods are still being actively researched (Phillips et al., 2009; Fourcade et al., 2014).

The resources that we used to collect information about Saudi reptile distributions did not contain records of species absences. Using

presence-only data to calibrate distribution models has some known drawbacks (see Zaniewski et al., 2002; Brotons et al., 2004) which may limit model performance (Brotons et al., 2004). Importantly, presence-only methods probably over-estimate species occurrence, because locations predicted to be suitable may not in fact be occupied, as a result of limited species dispersal. As a result, using presence-absence data is strongly recommended whenever available (Brotons et al., 2004). However, presence-only records often the only available information about species occurrences, and these are still informative about the true underlying distribution (Zaniewski et al., 2002). Despite using presence-only data, Maxent has been shown to perform well, generating predictive models even with biased data and small sample sizes (see Hernandez et al., 2006; Pearson et al., 2007; Wisz et al., 2008). In the absence of systematic surveys of the reptile fauna across the country, the Saudi data that we collected represent valuable information, and presence-only distribution modelling provides the current best option for describing and understanding patterns of species richness. Because of the tendency for this approach to over-estimate actual species occurrence, we must treat our predictions as likely upper-estimates of actual species richness. Only extensive systematic field sampling will prove whether these estimates are correct.

It is important to note that large portions of Saudi Arabia lack surveys, and the total number of observations upon which our model predictions are based is rather small. This is well known to limit the performance of models (Hernandez et al., 2006). The majority of our records were obtained from museums and the literature, and show signs of bias towards easily accessible locations and species that are easy to identify and catch. Although we have attempted to minimise the effect of sampling bias, it is unlikely that our predictions are unaffected by these issues.

Even though there is a general lack of detailed information about reptiles in Saudi Arabia (AbuZinada et al., 2004), they do represent a good case study because: i) they are very diverse and represent a large portion of vertebrate diversity; ii) they include endemic, vulnerable and unusual species; and iii) they are relatively well known and easy to sample, compared with many other taxa. It is important to encourage local environmental agencies and Saudi herpetologists to fill the knowledge gaps by establishing more reliable and more accurate datasets that are updated regularly. In particular, there should be systematic surveys of the different areas of the country, including locations have not been surveyed before. We consider our modelling exercise as the first step towards a more detailed understanding of the spatial distribution of reptile diversity, and biodiversity more generally, in this understudied part of the world.

## 5. Conclusion

Because efforts to study species distributions and patterns of species richness are unevenly distributed in many parts of the world, some locations are expected to have more species richness and uniqueness than currently reported (e.g., the south-western Arabian Peninsula; Ficitola et al., 2013). Reptiles are the major vertebrate taxon in the desert ecosystem in Saudi Arabia. Ensuring their conservation and long-term persistence should be a priority. In our case, even with unevenly distributed occurrence data, we were able to build informative spatial models to describe the likely current distribution of reptiles in Saudi Arabia. Distribution modelling studies such ours can provide important insights into patterns of diversity, in Saudi Arabia and elsewhere. We hope that this study will stimulate detailed future studies, including ground-truthing, to fill the gap in our knowledge about local terrestrial reptile distributions. Species distribution modelling has proven useful in many different subject areas, facilitating fieldwork and conservation planning, and helping us to understand the likely consequences of climate change and the threat of invasive species (Guisan and Thuiller, 2005; Pearson, 2008; Peterson et al., 2011). Distribution modelling has potential to contribute to these goals in Saudi Arabia, and may be vital

in future studies such as those testing the value of networks of protected areas or seeking to ameliorate the impact of global environmental change.

#### CRedit authorship contribution statement

**Abdulaziz S. Alatawi:** Data curation, Formal analysis, Writing - original draft. **Francis Gilbert:** Supervision, Formal analysis, Writing - original draft. **Tom Reader:** Supervision, Formal analysis, Writing - original draft.

#### Declaration of competing interest

None.

#### Acknowledgements

We thank Dr. Emad Kaky for advice during the analysis process. We thank the University of Tabuk, Saudi Arabia and the University of Nottingham, United Kingdom for funding this work.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jaridenv.2020.104153>.

#### References

- AbuZinada, A., Robinson, E.R., Nader, I., Al-Wetaid, Y., 2004. First Saudi Arabian National Report on the Convention on Biological Diversity. The National Commission for Wildlife Conservation and Development, Riyadh.
- Ananjeva, N.B., Golynsky, E.E., Lin, S.-M., Orlov, N.L., Tseng, H.-Y., 2015. Modeling habitat suitability to predict the potential distribution of the Kelung Cat Snake *Boiga kraepelini* Steineger, 1902. *Russ. J. Herpetol.* 22, 197–205.
- Baldwin, R.A., 2009. Use of maximum entropy modeling in wildlife research. *Entropy* 11, 854–866. <https://doi.org/10.3390/e11040854>.
- Brotans, L., 2014. Species distribution models and impact factor growth in environmental journals: methodological fashion or the attraction of global change science. *PLoS ONE* 9, e111996. <https://doi.org/10.1371/journal.pone.0111996>.
- Brotans, L., Thuiller, W., Araujo, M.B., Hirzel, A.H., 2004. Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography* 27, 437–448. <https://doi.org/10.1111/j.0906-7590.2004.03764.x>.
- Brown, J.L., Bennett, J.R., French, C.M., 2017. SDMtoolbox 2.0: the next generation Python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *PeerJ* 5, e4095. <https://doi.org/10.7717/peerj.4095>.
- Brown, R.L., Jacobs, L.A., Peet, R.K., 2007. Species richness: small scale. John Wiley & Sons Ltd eLS 1–8.
- Carranza, S., Xipell, M., Tarroso, P., Gardner, A., Arnold, E.N., Robinson, M.D., et al., 2018. Diversity, distribution and conservation of the terrestrial reptiles of Oman (Sauropsida, Squamata). *PLoS ONE* 13, e0190389. <https://doi.org/10.1371/journal.pone.0190389>.
- Catalogue of Life, 2018. Catalogue of Life. <http://www.catalogueoflife.org/>, Accessed date: 15 March 2018.
- Cox, N.A., Mallon, D., Bowles, P., Els, J., Tognelli, M.F., 2012. The Conservation Status and Distribution of Reptiles of the Arabian Peninsula. IUCN, and Sharjah, UAE: Environment and Protected Areas Authority, Cambridge, UK and Gland, Switzerland.
- Chettri, B., Bhupathy, S., Acharya, B.K., 2010. Distribution pattern of reptiles along an eastern Himalayan elevation gradient, India. *Acta Oecol.* 36, 16–22. <https://doi.org/10.1016/j.actao.2009.09.004>.
- El-Gabbas, A., Baha El Din, S., Zalat, S., Gilbert, F., 2016. Conserving Egypt's reptiles under climate change. *J. Arid Environ.* 127, 211–221. <https://doi.org/10.1016/j.jaridenv.2015.12.007>.
- Elith, J., Leathwick, J.R., 2009. Species distribution models: ecological explanation and prediction across space and time. *Annu. Rev. Ecol. Syst.* 40, 677–697. <https://doi.org/10.1146/annurev.ecolsys.110308.120159>.
- Elton, C., 1927. *Animal Ecology*. Macmillan Company, New York.
- Fattahi, R., Ficetola, G.F., Rastegar-Pouyani, N., Avci, A., Kumlutaş, Y., Ilgaz, Ç., Yousefkhani, S.S.H., 2014. Modelling the potential distribution of the Bridled skink, *Trachylepis vittata* (Olivier, 1804), in the Middle East. *Zool. Middle East* 60, 208–216. <https://doi.org/10.1080/09397140.2014.944428>.
- Ficetola, G.F., Bonardi, A., Sindaco, R., Padoa-Schioppa, E., 2013. Estimating patterns of reptile biodiversity in remote regions. *J. Biogeogr.* 40, 1202–1211. <https://doi.org/10.1111/jbi.12060>.
- Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* 37, 4302–4315. <https://doi.org/10.1002/joc.5086>.
- Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ. Conserv.* 24, 38–49. <https://doi.org/10.1017/S0376892997000088>.
- Fourcade, Y., Engler, J.O., Rödder, D., Secondi, J., 2014. Mapping species distributions with MAXENT using a geographically biased sample of presence data: a performance assessment of methods for correcting sampling bias. *PLoS ONE* 9, e97122. <https://doi.org/10.1371/journal.pone.0097122>.
- Gosling, S.N., Dunn, R., Carrol, F., Christidis, N., Fullwood, J., Gusmao, D.D., Golding, N., Good, L., Hall, T., Kendon, L., Kennedy, J., et al., 2011. *Climate: Observations, Projections and Impacts: Saudi Arabia*. The Met Office, UK.
- Graham, C.H., Hijmans, R.J., 2006. A comparison of methods for mapping species ranges and species richness. *Global Ecol. Biogeogr.* 15, 578–587. <https://doi.org/10.1111/j.1466-8238.2006.00257.x>.
- Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. *Ecol. Model.* 135, 147–186. [https://doi.org/10.1016/s0304-3800\(00\)00354-9](https://doi.org/10.1016/s0304-3800(00)00354-9).
- Guisan, A., Hofer, U., 2003. Predicting reptile distributions at the mesoscale: relation to climate and topography. *J. Biogeogr.* 30, 1233–1243. <https://doi.org/10.1046/j.1365-2699.2003.00914.x>.
- Guisan, A., Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. *Ecol. Lett.* 8, 993–1009. <https://doi.org/10.1111/j.1461-0248.2005.00792.x>.
- Hernandez, P.A., Graham, C.H., Master, L.L., Albert, D.L., 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography* 29, 773–785. <https://doi.org/10.1111/j.0906-7590.2006.04700.x>.
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* 25, 1965–1978. <https://doi.org/10.1002/joc.1276>.
- Huey, R.B., 1982. Temperature, physiology, and the ecology of reptiles. In: Gans, C., Pough, F.H. (Eds.), *Biology of the Reptilia, Physiology*. Academic Press, London, pp. 25–91.
- Javed, S.M., Raj, M., Kumar, S., 2017. Predicting potential habitat suitability for an endemic gecko *Calodactylodes aureus* and its conservation implications in India. *Trop. Ecol.* 58, 271–282.
- Kalioztopoulou, A., Brito, J.C., Carretero, M.A., Larbes, S., Harris, D.J., 2008. Modelling the partially unknown distribution of wall lizards (Podarcis) in North Africa: ecological affinities, potential areas of occurrence, and methodological constraints. *Can. J. Zool.* 86, 992–1001. <https://doi.org/10.1139/Z08-078>.
- Liu, C., Berry, P.M., Dawson, T.P., Pearson, R.G., 2005. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* 28, 385–393. <https://doi.org/10.1111/j.0906-7590.2005.03957.x>.
- Lobo, J.M., Lumaret, J.P., Jay-Robert, P., 2002. Modelling the species richness distribution of French dung beetles (Coleoptera, Scarabaeidae) and delimiting the predictive capacity of different groups of explanatory variables. *Global Ecol. Biogeogr.* 11, 265–277. <https://doi.org/10.1046/j.1466-822X.2002.00291.x>.
- MacArthur, R.H., Wilson, E.O., 1967. *Theory of Island Biogeography*. Princeton University Press.
- Merow, C., Smith, M.J., Edwards, T.C., Guisan, A., McMahon, S.M., Normand, S., Thuiller, W., Wiust, R.O., Zimmermann, N.E., Elith, J., 2014. What do we gain from simplicity versus complexity in species distribution models? *Ecography* 37, 1267–1281. <https://doi.org/10.1111/ecog.00845>.
- Merow, C., Smith, M.J., Silander, J.A., 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography* 36, 1058–1069. <https://doi.org/10.1111/j.1600-0587.2013.07872.x>.
- Metallinou, M., Červenka, J., Crochet, P.A., Kratochvíl, L., Wilms, T., Geniez, P., Shobrak, M.Y., Brito, J.C., Carranza, S., 2015. Species on the rocks: systematics and biogeography of the rock-dwelling *Pyrodactylus* geckos (Squamata: Phyllodactylidae) in north Africa and Arabia. *Mol. Phylogenet. Evol.* 85, 208–220. <https://doi.org/10.1016/j.ympev.2015.02.010>.
- Mora, C., Tittensor, D.P., Adl, S., Simpson, A.G., Worm, B., 2011. How many species are there on Earth and in the ocean? *PLoS Biol.* 9, e1001127.
- Naimi, B., Hamm, N.A.S., Groen, T.A., Skidmore, A.K., Toxopeus, A.G., 2014. Where is positional uncertainty a problem for species distribution modelling? *Ecography* 37, 191–203. <https://doi.org/10.1111/j.1600-0587.2013.00205.x>.
- Newbold, T., 2010. Applications and limitations of museum data for conservation and ecology, with particular attention to species distribution models. *Prog. Phys. Geogr.* 34, 3–22. <https://doi.org/10.1177/0309133309355630>.
- Pearce, J., Ferrier, S., 2000. Evaluating the predictive performance of habitat models developed using logistic regression. *Ecol. Model.* 133, 225–245. [https://doi.org/10.1016/S0304-3800\(00\)00322-7](https://doi.org/10.1016/S0304-3800(00)00322-7).
- Pearson, R.G., 2008. Species' distribution modeling for conservation educators and practitioners. Synthesis. American Museum of Natural History Available at: <http://ncep.amnh.org>.
- Pearson, R.G., Raxworthy, C.J., Nakamura, M., Peterson, A.T., 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *J. Biogeogr.* 34, 102–117.
- Peterson, A.T., Soberón, J., Pearson, R.G., Anderson, R.P., Martínez-Meyer, E., Nakamura, M., Araújo, M.B., 2011. *Ecological Niches and Geographic Distributions*. Princeton University Press.
- Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. *Ecol. Model.* 190, 231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>.
- Phillips, S.J., Anderson, R.P., Dudík, M., Schapire, R.E., Blair, M.E., 2017. Opening the black box: an open-source release of Maxent. *Ecography* 40, 887–893. <https://doi.org/10.1111/ecog.03049>.
- Phillips, S.J., Dudík, M., Elith, J., Graham, C.H., Lehmann, A., Leathwick, J., Ferrier, S., 2009. Sample selection bias and presence-only distribution models: implications for



- background and pseudo-absence data. *Ecol. Appl.* 19, 181–197.
- Pineda, E., Lobo, J.M., 2009. Assessing the accuracy of species distribution models to predict amphibian species richness patterns. *J. Anim. Ecol.* 78, 182–190. <https://doi.org/10.1111/j.1365-2656.2008.01471.x>.
- R Development Core Team, 2018. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.r-project.org/>.
- Rushton, S.P., Ormerod, S.J., Kerby, G., 2004. New paradigms for modelling species distributions? *J. Appl. Ecol.* 41, 193–200.
- Sanchooli, N., 2017. Habitat suitability and potential distribution of *Laudakia nupta* (De Filippi, 1843) (Sauria: Agamidae) in Iran. *Russ. J. Ecol.* 48, 275–279. <https://doi.org/10.1016/j.chemosphere.2017.07.130>.
- 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.
- Sillero, N., Brito, J.C., Skidmore, A.K., Toxopeus, A.G., 2009. Biogeographical patterns derived from remote sensing variables: the amphibians and reptiles of the Iberian Peninsula. *Amphibia-Reptilia* 30, 185–206. <https://doi.org/10.1163/156853809788201207>.
- Sillero, N., Campos, J., Bonardi, A., Corti, C., Creemers, R., Crochet, P., et al., 2014. Updated distribution and biogeography of amphibians and reptiles of Europe. *Amphibia-Reptilia* 35, 1–31. <https://doi.org/10.1163/15685381-00002935>.
- Sillero, N., Carretero, M.A., 2013. Modelling the past and future distribution of contracting species. The Iberian lizard *Podarcis carbonelli* (Squamata: Lacertidae) as a case study. *Zoologischer Anzeiger - A Journal of Comparative Zoology* 252, 289–298. <https://doi.org/10.1016/j.jcz.2012.08.004>.
- The Reptiles Data-Base, 2017, 2017. Reptiles Data-Base. <http://www.reptile-database.org/>, Accessed date: 20 October 2017.
- Tytar, V., Sobolenko, L., Nekrasova, O., Mezhzherin, S., 2015. Using ecological niche modeling for biodiversity conservation guidance in the western Podillya (Ukraine): reptiles. *Vestn. Zool.* 49, 551–558. <https://doi.org/10.1515/vzoo-2015-0065>.
- Vincent, P., 2008. Saudi Arabia: an Environmental Overview. Taylor & Francis Group, UK.
- Wilms, T.M., Wagner, P., Shobrak, M., Rödder, D., Böhme, W., 2011. Living on the edge? - on the thermobiology and activity pattern of the large herbivorous desert lizard *Uromastix aegyptia microlepis* Blanford, 1875 at Mahazat as-Sayd Protected Area, Saudi Arabia. *J. Arid Environ.* 75, 636–647. <https://doi.org/10.1016/j.jaridenv.2011.02.003>.
- Wisn, M.S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H., Guisan, A., NCEAS Predicting Species Distributions Working Group, 2008. Effects of sample size on the performance of species distribution models. *Divers. Distrib.* 14, 763–773.
- Yi, Y.-J., Cheng, X., Yang, Z.-F., Zhang, S.-H., 2016. Maxent modeling for predicting the potential distribution of endangered medicinal plant (*H. riparia* Lour) in Yunnan, China. *Ecol. Eng.* 92, 260–269. <https://doi.org/10.1016/j.ecoleng.2016.04.010>.
- Zaniewski, A.E., Lehmann, A., Overton, J.M., 2002. Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. *Ecol. Model.* 157, 261–280. [https://doi.org/10.1016/S0304-3800\(02\)00199-0](https://doi.org/10.1016/S0304-3800(02)00199-0).