The effects of paleoclimate on the distributions of some North West African Lizards.

Lee Ellis

A thesis submitted in partial fulfilment of the requirements of Liverpool John Moores University for the degree of Master of Philosophy

July 2018

Abstract

As awareness grows regarding impacts of global climate change, so does concern over the effects these changes have on a species habitat and distribution. Climate change is thought to have a major effect on the distribution of species, with the potential to cause isolated/fragmented populations, which could lead to genetic divergence. In this study species distribution modelling was applied to species occurrence data on northwest African lizards from Morocco, with corresponding environmental data. The aim was to identify how intraspecific divergence might be related to historical climatic events. Species distribution models (SDMs) were used to quantify a species niche and define the constraining factors that affect that niche. SDMs predict areas of suitable habitat under different climatic scenarios that replicate prehistoric climates, and used to examine if there is evidence to suggest historical divergence or historical splits in distributions that correspond to current patterns of geographical divergence within species. MaxEnt was used to develop the SDMs and define the species niche and variable constraints.

Previous studies have shown that the estimated divergence times of species discussed in this study range between 1–15 Ma. Environmental data dating back to these divergence times are unavailable or unreliable. Therefore, the Last Interglacial (LIG ~120,000 -140,000 years BP) and Last Glacial Maxima (LGM~ 21,000 years BP) datasets were used as a surrogate to earlier interglacial and glacial maximum climates, to analyse species distributions under earlier climatic scenarios which can then be inferred.

The models produced from this study portray geographical fragmentation/isolations of suitable habitat between currently recognised subspecies for all species studied. The results from this study give insights into potential events that could cause intraspecific divergence. Given that glacial patterns occur in a cyclic manner during the Earth's history, it is clear that they provide potential opportunities for disrupting a species habitat range and causing divergence due to oscillations between arid and humid environment.

Contents

Introduction

Species Distribution Models (SDMs)

Previous studies have shown climate change to pose a significant threat to biodiversity worldwide (Cahill et al., 2012; Thomas et al., 2004). Species can respond to climate change in a variety of ways, such as altering the timing of phenological events (Walther et al., 2002), changing their biotic interactions (Tylianakis et al., 2008) and shifting their distributions (Knowles et al., 2007). These changes in distribution could lead to range expansions for some species (Davey et al., 2012) and restrictions in others (Dirnböck et al., 2011). Predicting the impacts of climate induced changes is a rapidly increasing field for researchers (Bell et al., 2010), as these can be used to identify future risks to biodiversity to aid conservation projects (Guisan et al., 2013) and also help understand paleo events that could have caused species divergence (Knowles et al., 2007).

Species distribution modelling is used as a tool to help researchers answer theoretical questions in relation to a species success under hypothesised climate scenarios (Rossetto et al., 2012). Species distribution models (SDMs) are mathematical applications that combine two types of data; species occurrence records and environmental (predictor) variables (biotic and abiotic) (Austin et al., 2002). SDMs help estimate a species environmental limitations/ecological tolerances. They are fundamentally important when reintroducing species into the wild (Hunter et al., 2015), relocating a species to a more appropriate area, identifying potential areas for conservation (Gusian et al., 2013), selecting areas for growing crop plants or when predicting habitat for newly invaded regions (invasive species), (Elith et al., 2009; Johnson & Gillingham, 2008; Trabucco et al., 2010; Wilson et al., 2011). SDMs also allow the testing of ecological or biogeographical hypotheses of a species distribution, by determining factors such as upper and lower limits of constraining variables to see how the modelled distributions alter. SDMs can be used to estimate the influence of changes in climate (abiotic) on a predicted niche as well as non-climatic variables such as food production and competition (biotic). SDMs are able to extrapolate over past, present or future climate datasets, these predictions under different climatic scenarios can help us to understand how ancestral or future populations may distribute. These inferences can be achieved by using a predicted digitalised niche to project over these

different climatic/environmental scenarios (Phillips et al., 2006; Knowles et al., 2007; Bell et al., 2010).

SDMs typically aim to derive a species niche and the type of habitat required to support a species by identifying/quantifying variable importance. Many different mathematical methods and techniques can be used such a GLM and boosted regression (BRT) (Lehmann et al., 2002; Stryszowska et al., 2016; Benito et al., 2011). Essentially all methods aim to mathematically quantify the relationships between the observed species and the environmental variables they inhabit; this information is then portrayed in spatial form, normally a map. SDMs are able to manage large datasets and are able to use either presence, presence-absence, or abundance observations based on a random or stratified field-sampling technique, or observations obtained opportunistically (Graham et al. 2008). Observational data are geo-referenced using latitude and longitude coordinates for use in SDMs.

The availability and access to large electronic datasets are increasing (www.gbif.org), these sources stem from records in museums and field observations. The ability to measure the relationships in a quantitative manner allows data to circulate through many fields of expertise. A quantitative description of a species niche equips researchers with more tools to explore some of the questions at the forefront of ecology, evolution and conservation, which allows scientists to make inferences on whether or not a species can survive under certain hypothesised conditions or climatic scenarios.

Constraints on a species distribution

Environmental

Generally, not all species are evenly distributed over a spatial area, with potential to cluster in certain areas. Differences in areas of suitability indicates that a number of factors influence a species habitat (Phillips et al., 2006). The most common and easily available factors regularly used for modelling are abiotic factors. Abiotic factors are classed as non-living, physical or chemical factors that impose physiological constraints on a species existence (Soberon, 2005), and provide resources for existence in an area (Austin et al., 2002; Guisan et al., 2002). Abiotic factors such as precipitation and temperature are considered major influences on a species distribution, particularly for species in this study (Chamaille-james et al., 2005; Dickman et al., 1999; Ryan et al., 2016; Shine et al., 2002). A species free from interference with all other organisms, which has the use of all abiotic factors and resources to survive, is determined as occupying its *fundamental/potential* niche (Hutchinson, 1957).

Another group of factors are classed as biotic; biotic factors can be described as a living component that affects another organism, or shapes the ecosystem. Factors such as plants, animals, fungi and competition. Biotic factors can be classed as intra or interspecific (although interspecific factors are rarely used in SDMs, as they are hard to source and quantify). Intraspecific competition is competition within a species, whilst interspecific competition is the competition between two different species for a resource. Species that inhabit environments with limited resources means that populations have been found to be in equilibrium on a local scale, to minimize interspecific competition (Berendse, 1893). Distributions found to be at equilibrium are commonly found in forests, where competition for sunlight produces an even distribution of trees. Hutchinson (1957) considered biotic factors combined with abiotic factors to determine the *realised* niche of a species, a niche derived from the constraints of all variables.

Constraining factors can be categorised differently. Austin (2002) categorised environmental variables (resource) is either proximal or distal factors. These correspond to the position of the predictor variable in terms of how crucial its role is in regards to the species distribution. The most proximal variable is one that has the biggest effect on a species distribution whereas distal has a smaller effect. Austin (2002) also defined constraining factors as direct/indirect variables; direct variables are those that have a direct physiological effect on the species, for example temperature on lizards or territorial interactions. However, the classification depends on the species. For example water may be classified as a direct variable in plants, while temperature could be a direct variable for lizards. Indirect variables are those that do not have a direct physiological effect; for example, an indirect variable for plants could be the angle of slope, while competition could be an indirect variable for lizards. Indirect variables are considered easier to measure than proximal factors that tend to be more species specific. Although overlap between categories occurs, it is important to understand the subtle differences between proximal/distal, and indirect/direct. Indirect and distal are two types of predictor variables commonly used in most SDMs (Austin, 2002), as they are easier to source and quantify. It is thought that models based on proximal and direct gradients would produce the most mathematically robust models. However, it is known to be extremely difficult, and possibly impractical to use proximal variables for distribution modelling. The reason most models do not incorporate proximal variables is because they are hard to reliably quantify/measure and source, particularly due to their specificity (Araujo & Guisan, 2006; Austin & Van Niel, 2011). Therefore using variables such as proximal increases the potential for errors to be incorporated into a model.

Historical

Dispersal limitations and historical events are other factors that can constrain a species and restrict it from reaching equilibrium within its potential habitat, i.e. previous fires/floods or man-made (road)/natural structures (river) splitting a potential habitat (Condit et al., 2002). For this study we are particularly interested in events that occurred within NW Africa, Brown et al. (2002) suggested the rise of the Atlas Mountains during the Miocene caused the divergence of *Agama impalearis* through vicariance. This hypothesis has also been suggested for *T. weigmanni* (Sánchez & Escoriza, 2014). For species in northwest Africa, the Atlas Mountains act as a geographical barrier that could have possibly caused restrictions to a species potential distribution, leading to speciation through the restriction of gene flow, as suggested by Brown et al. (2002). The Rif Mountains of North Morocco could also act as a natural barrier and restrict gene flow. Mountainous regions can also create microhabitats that could differ significantly to environments at sea level; the steep incline means that within a relatively short distance the microhabitats can differ significantly from one another, changing environments from suitable to unsuitable in a short amount of space. These differences are often due to variations in rainfall, oxygen levels, soil quality, temperature and vegetation.

The end Messinian salinity crises and the consequent opening of the strait of Gibraltar (the collapse of the Gibraltar Arc) caused the re-flooding of the Mediterranean basin (Chumakov, 1973). This re-flooding led to the submergence of several major areas in northwest Africa, creating islands (in current inland areas) in the mountainous areas of NW Africa during shifts towards a humid climate approximately 5.5 Ma-3 Ma (Steininger & Rogl, 1984). These islands have been referred to as "fossil islands" that may have isolated faunas within NW Africa, which potentially led to speciation. Morocco's proximity to Europe and its contact during the Messinian stage of the late Miocene 5-6 Ma could also help explain the diversity of amphibians and reptiles (Hsü et al., 1973), due to events that have opened and closed the strait of Gilbraltar, hence, potential for vicariance speciation. A less important historical event in regards to dispersal limitations in NW Africa are tectonic movements, hence volcanic eruptions/earthquakes. Tectonic movements/volcanic eruptions have potential to cause local extinctions or destroy certain fundamental resources a species may be depend on throughout Morocco, Algeria and Tunisia for a considerable amount of time.

A species dispersal ability is another constraining factor and can be extremely useful when determining a species' *actual* distribution from its *potential* distribution. The discussed dispersal limitations and historical events have potential to drastically alter a species

distribution and should be considered when possible or when data for such events are available. A species may be absent from a site which is highly suitable based on climate and biota, however due to historical events and dispersal limitations a species is unable to inhabit all suitable areas, or becomes unable to reach suitable areas (Guisan & Zimmermann, 2000). Habitat fragmentation can lead to speciation, due to the interruption of gene flow between populations, which can lead to evolutionary independence. Evolutionary independence occurs when mutation, natural selection, gene flow and drift operate on populations separately. Fragmentation can either occur naturally, for example river or Mountain formations, or caused by human activity such as land conversion, a more direct effect that alters the environment much faster (Guisan & Zimmerman, 2000).

Species niche concept applied in SDMs

A frequent question in distribution modelling is; what is actually being modelled in a SDM, the *potential*/*fundamental* niche, the *actual*/*realised* niche or the probability of habitat use? (Elith et al., 2011). Studies can use the same environmental datasets, modelled under the same parameters but describe the output of the model as something different, *realised* instead of *probability of occurrence* for example (Guisan & Thuiller, 2005). Confusion is created between researchers and studies, but simply it depends on what variables have been used.

Occurrence records are samples of a species geographic distribution within the environmental "hypervolume", considered to include the effects of all factors that influence the species niche. Therefore, a current species niche is considered to be ingrained by both the biotic and abiotic elements as it is presently confined by both. Therefore, with the observed species population being ingrained by all the constraints of the environment more of a *realised* niche can be predicted through the SDM, rather than *potential* (Phillips et al., 2006). However, this theory is open to interpretation of how much biotic influence can be recorded in observation data. When the niche predicted by a statistical distribution model is mapped onto geographical space, it represents the potential distribution or habitat suitability (Araújo & Guisan, 2006).

In the majority of species distribution studies, choices of model variables are often limited to mainly abiotic factors, those that determine the *potential* niche. Other variables that influence distribution are rarely used as they are generally difficult to measure over spatial extents, too complex to interpret or simply unavailable. These are variable categories such as proximal/biotic resources (Thuiller et al., 2003; Soberon & Townsend Peterson, 2005). Using proximal/biotic

variables may improve model performance if incorporated, although difficult to incorporate all factors into a model without overfitting, therefore variable selection would be beneficial with a large amount of environmental (predictor) variables.

Interpretation of SDMs will depend on the context of the study in which it is used; the focus of a study may differ from the importance of *potential* habitat to the importance of specific variables for a species. Not all SDMs will predict a niche classified as *potential*, this depends on the methods and variables available. When biotic variables are unavailable only the *potential* niche can be calculated and results/predicted distributions can only be interpreted in terms of abiotic factors. Even though some biotic factors will have been intrinsically incorporated when calculating the niche, these biotic factors cannot be measured therefore the weight of influence on a species is unmeasurable. The reason for the calculated distribution being called "*potential*" is that the niche calculated is what a species with use of all resources could potentially cover, given the known abiotic constraints, without being influenced by all other organisms (biotic).

Niche evolution and conservation.

The idea of a species retaining its *fundamental* niche through time is called niche conservation, which suggests a species niche is constrained by the same ecological tolerances as its ancestors. If the ecological tolerances of an ancestral species are unknown then this assumption is necessary when extrapolating a species niche on to different climatic scenarios. It is argued that a species will always inhabit environments that share similar characteristics to those of close ancestors, and this may be supported by identifying morphological similarities between current species and its ancestor when possible. Therefore, to some extent a species niche is conserved through time, which can be used to test hypotheses on the causes of population divergence (Knowles et al., 2007). As a result, the question arises of how similar a species niche must be to another to be considered conserved (Wiens et al.,2010).

Niche conservation is often assumed in SDMs (Pearman et al., 2008), although it may not always be realistic, unfortunately it is difficult to make reliable inferences about extinct ancestors, even when fossil and other information are available. Speciation may occur due to one population conserving a species niche, while another isolated population evolves different adaptive traits, leading to divergence and eventually/potentially speciation. Different populations of the same species may adapt to different environmental conditions, in different parts of the inhabited range. If these adaptive traits differ between populations then over time divergence will occur,

especially if gene flow is low, as this will accelerate the process. Niche evolution is the alternative to niche conservation, and is the cause of geographical variation within the species niche (Peterson & Holt, 2003). It is where a species niche evolves over time brought on by different adaptive traits, adaptive traits that evolve in different populations brought on by differing selective pressures, leading to differentiated populations.

Niche comparative methods

Various methods have been used to compare the niche of a species and its respective subspecies. Niche comparative methods are used to determine the correlation between sister subspecies and the constraints imposed on their niche, these types of methods are useful to determine how similar or dissimilar the niches of sister subspecies are to one another. Two of the methods used to determine these relationships are; principle component analysis (PCA) and niche overlap analysis.

Principal component analysis

Principal component analysis (PCA) is a statistical procedure, which identifies a group of uncorrelated variables called principal components, from a potentially correlated dataset such as the bioclim dataset (Boulangeat et al., 2012). The purpose of a PCA is to help describe the maximum amount of variance within the least number of principal components (grouped variables). The first principal component has the largest possible variance, which accounts for as much of the variability in the data as possible. Each consecutive component in turn represents the highest residual variance possible. PCA is a useful tool for species distribution modelling as principal components can help clarify what can be a large uncorrelated array of variables and help define what the major constraining factors on a species are (Tamura & Tsujita, 2007). PCA can also be used to address collinearity between variables by grouping alike variables together. Principal component analysis was used in this study to identify any correlations/relationships between bioclim (environmental variables) and between sister subspecies which helps identify the differences in niches. More details of the PCA produced in the methods section of this study.

With the recognised drawbacks associated with a PCA, it is advised that outputs should be treated as an observational tool. Usually the PCA is a great and meaningful measure of central tendency but not always. One drawback of PCA is that it relies on linear assumptions; a PCA will try to find orthogonal relationships between data in order to find hidden linear correlations. However if the data are not linearly correlated then PCA will not be effective. In this case, the

Kernal PCA (KPCA) may be a better option. For nonlinear relationships, the only way PCA can represent that is by adding additional dimensions, this makes for a complicated, possibly uninterpretable, and frequently unstable solutions. As mentioned, a PCA tries to define orthogonal principal components within a dataset; however, this may not always be the best way to represent the data. A PCA fails to find the correct vectors due to a lack of orthogonal relationships, in which case an independent component analysis may be a better option. Consideration of how many PCs to discard following a PCA is also important. The PCA considers the low variance components in the data as noise and recommends discarding these components, even though sometimes those components may play a major role. Several techniques for deciding the most important components have been proposed (Cattell, R.B., 1966; Frontier, S., 1976).

Niche overlap analysis

Niche overlap analysis uses the Schoener's (1968) D value that is a measure derived from Hellinger distance called I to quantify the similarities in niches between two species or sister subspecies. The values are calculated by comparing the estimates of habitat suitability of a species (or in the case of this study subspecies). Calculated from each grid cell of the study area produced through Maxent, after normalizing each model so that all suitability scores within the geographic space sum to 1. This calculated score helps with niche comparison with a 0-1 scale, making it easy to see whether the niche of two different species or subspecies are similar to one another. The Schoener's D values range from 0 (niche models have no overlap), to 1 (niche models identical), with suggestions on how to interpret results; 0-0.2 = no or very limited overlap, 0.2-0.4= low overlap, 0.4-0.6=moderate overlap, 0.6-0.8= high overlap and 0.8-1= very high overlap (Rödder, D & Engler, 2011). Niche overlap was performed in this study to compare the estimated niches of sister subspecies.

Predictions through a SDM

The use of SDMs to predict the effects of climate change on a species distribution generally requires some assumptions. These are niche conservation, source habitat and a cyclic relationship between glacial cycles and climatic impact, these assumptions are discussed further in the methods section. A number of different statistical models and algorithms have been applied to SDMs. Certain statistical models are more specific to either fauna/flora or presence/presence-absence, depending on why the model was made, with numerous methods suggesting how the SDMs should be presented. Anderson (et al., 2006) reviews the different

statistical approaches commonly used in distribution modelling. These include multiple regression models, Environmental envelopes, Bayesian approach and machine learning methods.

One of the more popular techniques recently used for SDMs are regression models (GLMs and GAMs fall under regression models). These allow for a larger range (scale) of distributions to be predicted. This is due to the larger range of response curves available such as Gaussian, Poisson, Binomial and Gamma, unlike a simplistic model such as least squares regression (Elith & Graham, 2009). Environmental envelopes are also used for modelling, and have been mainly used in the past with large datasets for species or vegetation models (Shao et al., 1995). McKenzie et al (1992) developed the BIOCLIM model, a model that uses a species specific, fitted, multi-dimensional, minimal rectilinear environmental envelope. Originally, the model was developed to model plant species distributions in Australia, hence its practicality with large datasets. Soon after more techniques were developed based on the environmental envelope methodology used in BIOCLIM, these influenced approaches where DOMAIN (Carpenter et al., 1993) and HABITAT (Walker et al., 1991).

The Bayesian approach is another method popular in SDMs. It combines priori probabilities of observing a species, with a likelihood based approach on environmental predictor variables, for the area in which the species were observed. Machine learning tools have also been introduced to species distributions (MaxEnt) (Phillips et al., 2004), used with various methodologies that make modelling of large/global data sets much easier (DOMAIN/GARP).

Anderson et al. (2006) compared model performance between the main methods used in SDMs. The study produced and published 16 modelling methods, over 226 species from 6 regions to create a comprehensive set of model comparisons. The results from the study state that both AUC and *r* values indicate high predictive performance for MaxEnt.

Table 1. A comparative table comparing AUC scores and extra details were available of some of the modelling methods used for species distribution modelling derived from Anderson et al. (2006).

Algorithm	Mean AUC rank per species	Rank of mean AUC over all species	Description	Software	URL	Reference
BIOCLIM	10.85	16	Envelope model	DIVA-GIS	www.diva-gis.org	(Nix,1986), (Busby, 1991)
DOMAIN	8.7	9	Gower distance	DIVA-GIS	www.diva-gis.org	(Carpenter et al., 1993)
OM- GARP	8.92	11	Genetic algorithm	DesktopGarp	www.nhm.ku.edu/desktopgarp/index.html	(Stockwell &Peters, 1999)
DK-GARP	10.47	15	Rule sets from genetic algorithm	DesktopGarp	www.nhm.ku.edu/desktopgarp/index.html	(Stockwell &Peters, 1999)
GAM	8.26	7	Regression	GRASP	www.unine.ch.cscf/grasp	(Lehmann et al., 2002)
GLM	8.64	10	Regression	GRASP	www.unine.ch.cscf/grasp	(Lehmann et al., 2002)
MaxEnt	6.69	5	Maximum entropy	MAXENT	www.cs.princeton.edu/~schapire/maxent	(Phillips et al., 2006)
MaxEnt-T	6.42	3	Maximum entropy; with threshold features	MAXENT	www.cs.princeton.edu/~schapire/maxent	(Phillips et al., 2006)
BRUTO	8.79	12	Regression	R and Splus, mda package	https://rdrr.io/rforge/sdm/src/inst/methods/sdm/fda.R	
GDM	7.53	4	Generalised dissimilarity modelling	Specialized program, uses Arcview and Splus	www.arcgis.com	
GDM- SS	7.38	6	Uses community data	Specialized program, uses Arcview and Splus	www.arcgis.com	
LIVES	10.22	14	Regression splines	Specialized program		
MARS	8.92	8	Regression splines	R, mda package	https://cran.r-project.org/web/packages/mda/mda.pdf	(Mateo et al., 2010)
MARS - COMM	6.15	1	Regression splines	R, mda package	https://cran.r-project.org/web/packages/mda/mda.pdf	(Svenning et al., 2015)
MARS-INT	9.72	13	Regression splines; Interaction allowed	R, mda package	https://cran.r-project.org/web/packages/mda/mda.pdf	
BRT	6.2	2	Boosted regression trees	R, gbm package	https://cran.r-project.org/web/packages/mda/mda.pdf	(Elith et al., 2008)

Maximum Entropy

Maximum Entropy is a principle from mathematics that has been applied in a large spectrum of fields such as finance, astronomy and the medical industry (Anderson et al., 2006), and is the principle applied in the program MaxEnt. A maximum entropy state in mathematics is considered to be the closest to uniform and evenly spread, derived from the known constraints relevant to the subject (Phillips et al., 2004). The maximum entropy principle only bases predictions on information known about the subject/species such as environmental variables, occurrence records and biotic interactions (Jaynes, 1986), with no initial assumptions made that are not justified by the information gathered. It provides a suitable approach when trying to calculate a species niche due to the many unknown/ unmeasurable variables that constrain a species. A distribution with the largest remaining uncertainty, consistent with known constraints is said to have maximum entropy. To simplify, the distribution said to have maximum entropy, is the one that makes the least claim to being informed beyond the data gathered to construct the model framework. Therefore, no initial assumptions or bias are incorporated within a model.

MaxEnt

MaxEnt is the program that incorporates a general-purpose machine learning method and applies the maximum entropy methodology to environmental and observational datasets. MaxEnt has the ability to make inferences from incomplete data, capable of using presence only observations with predictor variables and able to handle a vast amount of data at one time. This means it has become widely used for SDMs in numerous fields (Phillips et al., 2004). An advantage of using presence-only observations is that the data is removed from any bias that could be created from unreliable absence records (Jimenez-Valverde, 2008). Further explanations on this sort of bias are explained later. The MaxEnt probability distribution follows a concise mathematical definition, therefore the results are easy to understand, interpret and amenable for analysis. MaxEnt is also a generative approach that uses the environmental data from across the study area rather than discriminative, which can be an inherent advantage when the amount of training data are limited (Phillips & Dudík, 2008; Phillips & Elith, 2013). Another advantage of Maxent is that the program is able to use both continuous and categorical (feature data) and can distinguish the interactions between them; meaning a broad range of variables can be applied within the model (abiotic/biotic). MaxEnt and its use of maximum entropy help remove bias from a sensitive model with options to create constructive analytical work, both of which are vital in producing work which others can interpret and progress from (Phillips et al., 2004). A detailed description of MaxEnt is available from Phillips et al., (2008).

For species distribution modelling in MaxEnt the geographical region in the study is considered the space which a distribution, and in essence a niche is defined. The projection layer can be larger or differ climatically, dependent on the purpose. The environmental layers are considered the available abiotic and biotic variables, and the occurrence localities are defined as sample points, from which the abiotic/biotic data are sampled from (Phillips et al., 2004). When producing SDMs in MaxEnt, the program produces a probability distribution over the pixels in the grid, starting from the uniform distribution and repeatedly improving using the maximum entropy method to fit the data. The probability it assigns to each pixel is typically very small, as values must sum to 1 over all the pixels in the grid.

A few drawbacks with MaxEnt have been mentioned in other studies that relate to the difficulties surrounding presence-only modelling and bias that can be introduced from using background data (Yackulic et al., 2013). However, it is acknowledged in other studies that bias can be incorporated within absence data as well (Phillips et al., 2009). Another possible drawback that affects the accuracy of presence only modelling relates to biases incorporated in the occurrence localities. Using remotely sensed datasets rather than field-based observations can help address location bias by providing a more complete dataset available for the study area and reducing sampling bias (Arnold et al. 2014). This should help remove some of the bias associated with presence only modelling. Phillips et al. (2006) also acknowledges that model validation tools within MaxEnt are limited and should be paired with external evaluation methods to validate their models such as null hypothesis etc. Chakraborty et al. (2011) also describes MaxEnt as more black-box like than better-known statistical analyse, which can lead to ill-informed interpretation of results based off misinformed models. Another limitation is the possibility of over-fitting, limiting the capacity of the model to generalise well to independent data. The 'regularization multiplier' parameter in MaxEnt aims to address this by limiting the complexity of the model and generating a less localized prediction (Phillips & Dudík, 2008).

Options in MaxEnt

In MaxEnt users are able to select options that treat the variables entered differently. These alternative options in MaxEnt are called features. The features available are; *Linear*; where continuous variables should be close to their observed values (their mean at occurrence localities). *Quadratic*; where variance of continuous variables should be close to observed values. *Product*; where the covariance of two continuous variables should be close to observed values. *Threshold*; where the proportion of the model that has values above a threshold for a continuous variable should be close to observed proportion. *Hinge*; linear features truncated at threshold and *Binary*; where the proportion of each category in a categorical feature should be close to the observed proportions. To

illustrate, if we use precipitation as a predictor, the linear feature class ensures that the mean value of precipitation where the species is predicted to occur approximately matches the mean value where it is observed to occur. A quadratic feature constrains the variance in rainfall where the species is predicted to occur to match observation. A product feature constrains the covariance of rainfall with other predictors and is equivalent to interaction terms in regression (when linear features are also included). Threshold features make a continuous predictor binary by generating a feature whose value is 0 below the threshold and 1 above. Hinge features are like threshold features, except that a linear function is used, instead of a step function. All features are rescaled to the interval [0,1] to make the coefficients comparable.

MaxEnt has three run types for users to choose from; Bootstrap, Subsample and Cross validation (k-fold). By default, the replication method used is *cross-validation*, where the occurrence data are randomly split into a number of equal-sized (k) subsets groups called "folds". Models are created by withholding each fold in turn, the remaining folds are then used for evaluation. Cross-validation has an advantage over using a single training/test split: it uses all of the data for validation, thus making better use of small data sets. *Subsample* is an alternative replicative run type. It is a renamed jackknife approach (avoiding confusion with jackknife for predictors). The subsample method repeatedly splits the presence points into random training and testing subsets. With this method, you can set the number of replicates and the percentage to be withheld from each replicated run. This method might be ideal for modellers who wish to specifically assign a percentage of data to withhold for testing and to control the amount of repetitions the model produces. This type of replication run is suited for modellers who have a larger occurrence dataset for their species of interest. Lastly *bootstrapping*; this model selects its training data by sampling with replacement from the presence points, with the number of samples equalling the total number of presence points. With bootstrapping, the number of presence points in each set equals the total number of presence points; therefore, the training datasets will have duplicate records. This method is also preferred by modellers with a small occurrence dataset as MaxEnt tests the model with occurrences that may have been used to train the model.

Model Validation

Model evaluation methods are used in SDMs as they evaluate the ability of a model to discriminate a presence from an absence, creating a challenge when using presence only data. Fortunately, methods such as *r*, KAPPA, AUC and sensitivity & specificity analysis can be applied to presence and background points/pseudo absences data. As the model distinguishes the available environment for the species,

it can determine the non-presence locations as absences. Below are some of the most common methods used:

Jackknife

One method used for model validation is a Jackknife test. Jackknifing involves exclusion of variables in turn; a model is then created with the remaining variables. Another model can then be created using each variable in isolation. This method allows one to determine which variables show high importance within the model.

AUC

Area under the curve (AUC) is another method commonly used for model validation of SDMs. The area calculated is that under the receiver operating characteristic (ROC) plot, originally a method developed in fields of medical diagnosis. The axis of the ROC plot comprises of true positive rate on the y-axis (Sensitivity) and false positive error rate on the x-axis (1- Specificity). Values are determined using each possible value of threshold probability. The values in which the threshold increases by can be set, and usually increases in steps of 0.01, i.e. 0.01-1. Once the AUC is calculated by summing the area under the ROC curve. The values help evaluate the model and its accuracy. AUC values can range from 0-1.0; 0.5 taken as random predictions means the model is performing no better than predicting randomly. AUC values that range 0.5-0.7 are considered poor in terms of model performance, 0.7-0.9 adequate and >0.9 being high (Manel et al., 2001). AUC is a method that has been used frequently to evaluate SDM predictive performance. The AUC alone does not give a complete assessment due to MaxEnts dependence on the unknown real species occurrence. Therefore, should be coupled with other evaluating methods such as the null hypothesis (Raes and Steege, 2007).

Pearson correlation coefficient

The Pearson correlation coefficient (r) is another method used for model validation. Also known as the biserial (bivariate) correlation, when one of the binary variables is used as a threshold independent measure of the predictive performance of SDMs *(*Anderson et al., 2006*)*. While the previously described AUC is a ranked based measure, the (r) coefficient takes into account the degree in which prediction varies linearly with the observation. As the Pearson correlation coefficient measures the linear association between the observations (test) and predictions, it has been suggested that it is sensitive to the effect of varying relative sampling intensity in the training data (Phillips et al., 2008). Variations of this method are used where concordance measures such as observation data are applied.

Due to the non-parametric nature of variables such as these the rank correlation (Spearman's) method is more plausible for evaluating SDMs (Pineda & Lobo, 2009).

Kappa

Kappa is another method used to measure predictive performance, though becoming less popular. The kappa value is derived by calculating the differences between observed and chance. Kappa was widely used in SDM research however more recently the Kappa value took criticism for being "inherently reliant" on prevalence (Lantz & Nebenzahl, 1996). It has been suggested that this sort of dependency introduces bias. Therefore, although Kappa has been used in species distributions it is becoming less popular due to its sensitivity to prevalence (Segurado & Araujo, 2004; Anderson et al., 2006). It has also been suggested to be insensitive to prevalence under certain conditions, hence the criticism this method has received. Neither (r) nor Kappa have been used in this study as evaluation methods. However, these are two commonly used principles used for SDM validation.

Sensitivity and Specificity analysis

Sensitivity and Specificity are considered statistical measurements that represent how well a model predicts its binary classifications (presence/absence), (Liu et al., 2013). This binary classification test is commonly used in the medical field. Sensitivity, also known as the true positive rate measures the proportion of actual positives that are correctly identified as such (the percentage of presence locations that where correctly identified as presences). Specificity, also known as the true negative rate measures the proportion of absences which where correctly identified as such (the percentage of absent locations that were correctly classified as absences). Both sensitivity and specificity are directly linked to the concept of type I and type II errors. A type I error (false positive rate) is a rejection of the null hypothesis. This is when a model predicts a presence in a location considered unsuitable. A type II error (false negative rate) is the failure to reject a false null hypothesis. Essentially this is when a model predicts an absence in a location that is considered suitable.

Confusion Matrix

Predicted (Model)		Data Set	
		Presence	Absence
	Presence	a	b
	Absence	c	d

Sensitivity = $\frac{a}{a+c}$ **Specificity =** $\frac{d}{b+d}$

False positive rate (α) = type I error = 1 − specificity

False negative rate (β) = type II error = 1 − sensitivity

A perfect performing model will measure 1 (on a scale of 0-1). This represents 100%, which means predicting all presences in the presence group. If we look at the sensitivity formula, 100% means a = a+c. When models specificity is valuated at 1, representing 100%, it shows that the model predicts absences where the data shows absences. Looking at the specificity formula, 100% means d = b+d.

Null hypothesis evaluation

A null hypothesis evaluation is a principle used by (Raes & Steege, 2007). The null model methodology is a way to measure whether the SDMs, produced with presence only data, differ significantly from what would be expected by chance. This is achieved by creating random observational datasets and seeing how they compare within a model. More details on this method are available in the methodology.

Datasets used in SDMs

Environmental datasets (predictor variables)

Environmental datasets are vital in distribution modelling; the types of data are selected on the ecological assumption that they have a direct or indirect correlation with the species, or are related with other variables that show correlations. In environmental datasets, the land is classified into multiple explicit groups such as annual temperature, annual precipitation and lowest temperature. This means one point on a map represents multiple variables. Data such as these are usually available

in two formats; vector and raster. They often vary in detail/resolution and scale depending on the source and time scale required. These datasets are amenable for their use in GIS programmes in a variety of ways, such as transformations.

Large environmental datasets such as global climate datasets are widely available; usually these datasets are constructed from station climatological averages. A recent form of a global climate dataset is WorldClim (www.**worldclim**.org*).* Hijmans et al., (2005) developed the WorldClim database of climate surfaces, with a 30 arc second spatial resolution. It compiled monthly averages of climates measured at weather stations, mostly for the 1950–2000 periods. The datasets consist of 19 variables listed in the method section of this study. The interpolated climate layers were compiled of databases by the Global Historical Climatology Network (GHCN), the FAO, the WMO, the International Centre for Tropical Agriculture (CIAT), R-HYdronet and a number of additional minor databases for Australia, New Zealand, the Nordic European Countries, Ecuador, Peru, Bolivia among others.

Past climatic datasets can also be sourced from WorldClim. The last glacial maximum (LGM) climatic database (~21,000 years BP) was produced by the Paleoclimate Modelling Intercomparison Project Phase II (PMIP$_2$). The last interglacial (LIG) climatic database (~120,000 -140,000 years BP) is an independent source, where the data was modelled for use in a project. Otto-Bliesner et al (2008) produced climate simulations for the LIG with a global, coupled, ocean-atmosphere-land-sea-ice general circulation model (NCAR) Community Climate System Model (CCSM).

Observational Data

Databases of animal specimens have been available for centuries preserved in the form of specimens in museums, providing a documented historical record of occurrence of species (Chapman, 1999). It is estimated that there are up to 3 billion records held in this form, each with an associated collection event, describing the time and place the specimen was found. Traditionally these records (adhoc) were used for taxonomic purposes. Digitalisation of such records has enabled observational data to be increasingly available due to improved computer processing, data digitalisation and storage capacities. Such georeferenced occurrence records are often derived from field surveys, designed to study a species distribution for mapping. The renewed interest in using such records in biogeographic studies such as distribution modelling has caused debate on data quality (see next chapter).

Another source of observational data is survey data, where species are observed in an area of interest and recorded within a given time period. A study produced by Valinia et al (2014) used survey data where all individuals (fish) were collected in the field and used to create species distribution models. Another study collected survey data of a North African Lizard to be used in predicting suitable habitat

(Kaliontzopoulou et al., 2008). Data quality is important to SDMs due to the implications poor quality data can have in regards to over/ under estimating populations (Araujo & Guisan, 2006). A common disadvantage of survey data is the variation in sample intensity, both spatially and temporally meaning species may be unobserved, hence creating false negatives. The types of the errors that produce poor quality observational data are discussed in the next chapter.

The introduction of digitised observation data and the accessibility of such data via the internet has allowed observational data to become more available. Digitalised data has contributed hugely in the increased amount of studies in conservation planning, reserve selection and climate change studies that incorporate SDMs. The majority of species observational data that are collected and used for SDMs can be classified as presence only data with each species observation are given in longitude-latitude coordinates.

It has been suggested that sampling a large geographical region could potentially increase the amount of fundamental niche that is represented. This is based on the assumption that in a larger study area spatial variation exists in community composition. This allows a greater number of biotic interactions to be incorporated that are difficult to record (Peterson & Holt, 2003). This idea stems from the thinking that presence records are imprinted by all the factors that constrain them (both abiotic, biotic and dispersal), therefore factors that correspond with absences (Manel et al., 2001). It is thought that if a species is absent from a potentially suitable location, due to local extinction for example, the imprint of that absence is still found in the distribution of the presence records. For example, if a population is affected due to over predation, an unknown biotic variable, the effect will be indirectly represented in the distribution, without absence records and without predator occurrence records. The pattern of presence will suggest the most suitable areas for a species regardless of absence data; therefore the surveying size of the study area can affect the impact of this theory (Manel et al., 2001). Enough area should be selected in the model to account for these missing absence records. If the area selected is too large in comparison to the habitat, then it may lead to the model producing false absences or false pseudo absences, affecting the models predictive performance. Idealistically the model should cover the fundamental niche, an area quite elusive without any previous studies on the particular species.

Many modelling techniques are purposely designed for presence only data. These incorporate absences by different methods. One method determines background points, randomly sampled from background data and uses them with the presence records. This method is found in some packages such as GARP, ENFA, MaxEnt and regression methods (Elith & Leathwick, 2009). Background data are not attempting to guess at absence locations, but rather to characterize environments in the study

region (Phillips et al. 2009). In this sense, background is the same, irrespective of where the species has been found. Background data establish the environmental domain of the study, whilst presence data should establish under which conditions a species is more likely to be present than on average. Background data are not attempting to guess at absence locations, but rather to characterize environments in the study region (Phillips *et al.* 2009). MaxEnt assumes that the species is equally likely to occur anywhere in the sample region, which assumes that every pixel has the same probability of being selected as background.

A closely related but different concept is "pseudo-absences", also used for generating the non-presence class for logistic models. This refers to the situation where it is unknown whether a species occurs there or not. In this case, researchers sometimes try to make informed decisions where absences might occur within the background data based on some prior distribution knowledge; they may sample the whole region except at presence locations, or they might sample at places unlikely to be suitable for the species.

Quality & accuracy with species observational data

Specimen data can be broadly categorised into three dimensions; space, time, and identity (Wieczorek et al., 2004), all of which can incorporate issues of data quality. Issues of data quality can be split generally into error and bias, although the two are connected. Bias refers to problems intrinsic with the data not encompassing the full environmental niche whereas error refers to a mistake in the data such as misidentification or human error in georeferencing. Error can result in bias incorporated if not identified as such.

Error in species observational data

Error refers to human errors; one example would be a taxanomic misidentification of a species, which could lead to an occurrence being identified at an incorrect location due to a misidentification of a specimen (Wieczorek et al., 2004). Another error in observational data is spatial error; this stems from georeferencing errors, which are imprecisions of an observation record or error in the original location of a record. Generally specimen data from museums have typically been recorded as textual descriptions with geographical coordinates often added after the collecting event. Adding the

coordinates after the collecting event can cause many errors; human error due to incorrect coordinates selected or copied (Wieczorek et al., 2004). Alternatively, the textual information of a species incorrectly imputed, with location descriptions being outdated. Generally, when field surveys obtain the observational records there is rarely any record of the procedures taken, the assumptions made or any uncertainties proposed when generating the coordinates (Chapman, 2004).

It should be noted not all data of low precision is deemed low quality, what are important is if the data are fit for the purpose for what they are being applied to. Chapman (1999) notes that errors in observational data are common and to be expected. However, the use of Global Positioning System (GPS) in field surveys has vastly increased and hoped to help aid in the reduction of errors incorporated in terms of geographical error. Although studies have shown to suggest errors with data collected with GPS (Hijman et al., 1999).

Bias in observational data

Sample Bias

Bias can be incorporated into a model through a species' observed population, whether that is a poor sampled area or a species that is hard to detect. This means bias can be introduced through presence and presence/absence records alike. Species observational data may incorporate bias due to collectors sampling in areas they would expect to find what they are looking for, sampling where conveniently accessible or when collected opportunistically. Collectors have often been found to sample along rivers, roads, areas of high diversity and areas easily accessible, resulting in a neglection of inaccessible areas and restricting the number of observed populations (Chapman, 2004; Hijman et al, 1999). This type of bias is considered to have the potential of a much stronger impact on presence-only models rather than presence-absence models (Phillips et al., 2009).

Issues around accuracy can also occur when datasets are particularly coarse, and the low resolutions affect accuracy of the geographic locations as they may deviate from the true observed area. The availability of Global Positioning System (GPS) has dramatically improved issues with accuracy, and should reduce geographical errors when assigning co-ordinates to environmental datasets. However Hijman et al., (1999) found even data associated with GPS can have errors.

There are arguments over absence records, as absence records are considered less reliable and can incorporate bias. The reason they are considered to incorporate more bias is that they are considered more difficult to realistically definitively say whether a species is absent in an area and difficult to

determine the correct geographical range of a species distribution. Generally, less time is spent identifying absence locations when compared with identifying presence locations.

Ecological bias

Generally it is assumed a species population is at equilibrium with its environment and all constraining factors. However bias could be incorporated by assuming an equilibrium population when in fact it is not, through invasion or disease for example. Therefore further predictions are based on wrong assumptions, introducing bias.

It is suggested presence-only records have potential to incorporate low bias if sampled correctly, as it is thought absences bear strong imprints of biotic interactions and dispersal constraints (Manel et al., 2001; Phillips et al., 2009) which are hard to quantify and therefore easier to incorporate bias from. To try reduce the bias incorporated when collating observation locations a systematic sampling approach could be used, with notes made about any areas deemed unsatisfactory and any assumptions made. If observational data are not reliable then the bias is ingrained into the original data itself, adding incorrect assumptions. Potential for ill-informed predictions to be made about suitable and unsuitable habitats.

If sample error/bias has not been accounted for or considered a factor, then the model might be a closer fit to that of the observational data rather than to the true distribution of the species (Phillips et al., 2009). Jimenez-Valverde (2008) suggests that presence only data have the potential to alleviate us from the problems that surround absence records. The study emphasises that absences can carry strong imprints on factors such as dispersal constraints and biotic influences. These factors are not easily measured or quantified, suggesting that no abiotic factors are free from biotic interactions.

Species distribution in Paleobiology

Past distributions of organisms and their abundances are key matters for not only paleobiology but also biology and geology. How an organism's spatial distribution changes through time is fundamental to understanding the evolution of biodiversity in geographic patterns (Lomolino, 2010), as well as the conditions that are required to preserve a species niche (Willis et al., 2010). A major problem when modelling a species past distribution is the lack of, or incomplete information available. A lack of data is also a problem when modelling current distributions, though it is more apparent when modelling past distributions. Climatic variables that predate the LIG (140,000 yrs BP) are hard to source and

usually coarse as they are essentially predictions themselves. Another problem when modelling a species past (paleo) distribution is the assumption that the niche constraints of the ancestral species are similar or identical to that of the current species. Generally, data are unavailable to enable researchers to quantify if/how much a species has evolved over millions of years, therefore it is generally an assumption most researchers use when modelling the paleo distribution of a species.

The use of SDM's to predict future or ancestral distributions

Predicting a species distribution dating further back than the LIG & LGM can be achieved by using LGM and LIG datasets. These datasets act as a surrogate to earlier GM and IG periods for which data are unreliable (Knowles et al., 2007). This is possible under the assumption that glacial cycles follow a cyclic pattern that corresponds to Earths circumnavigation around the Sun; this is known as the Milankovitch theory and is discussed later in the study (Hewitt, 2000; Kozak & Wiens, 2010). The climatic tolerances and quantifiable variables that impose constraints on a species can be measured using specific programs such as MaxEnt, using the present day climate data, a 50-year average (1950-2000). Once the constraints (biotic/abiotic) have been defined and quantified, a predicted digitalised niche is created. This predicted digitalised niche acts as a framework to which the LIG & LGM datasets can be projected on to. Inferences towards conditions and potential suitable habitats during earlier glacial maximum/interglacial periods can then be made based off the LIG and LGM predictions; hence, the potential suitable habitat during a glacial cycle in general can be predicted.

Paleoclimate and speciation within Africa

Evolutionary theories for speciation events within Africa state that important evolutionary changes during Pliocene-Pleistocene interval (the last 5.3 Ma) were potentially mediated by climate variability within Africa, due to recurrent arid-humid climate cycles (deMenocal, 2004). Analysis of mammal fossil data indicate changes in African faunal assemblage and possibly speciation during the Pliocene-Pleistocene, with suggested varied and open habitats at 2.9–2.4 Ma and after 1.8Ma (deMenocal, 2004). These intervals correspond to the emergence of the hominid genes *Homo*. Together, the African faunal and paleo-climatic records suggest three restricted intervals 2.9-2.4 Ma, 1.8-1.6 Ma and 1.2–0.8 Ma, when shifts towards drier African conditions were accompanied by changes to African faunal assemblages and perhaps speciation. These large oscillations in climate producing large changes in flora and fauna can also be traced back into the Tertiary (Hewitt, 1996).

Some consistent patterns of paleoclimate variability have been identified within subtropical Africa (deMenocal, 2004). The first is that orbital-scale African climate variability continued throughout the

Pliocene and in some cases extended into the Miocene (5.3 – 23.03 Ma) and Oligocene (23.03 Ma+) (deMenocal, 1991). The second was that large-amplitude African aridity cycles corresponded to the amplification of high-latitude glacial cycles (deMenocal, 1995; deMenocal et al., 1993) and another was evidence for 10^5 - 10^4 year 'packets' of high and low-amplitude paleoclimate variability being paced by orbital eccentricity (deMenocal, 1995).

Marine dust records are used to determine climatic conditions through time, and are used in many studies as a template towards paleo climates (Knowles et al., 2007). Due to wind patterns across North Africa, sediment is carried from the North/North West of Africa then released into the sea. Sediment such as dirt, vegetation and pollen are swept out to sea, where it is deposited on the seafloor. These core samples give insights to how an environment may have been altered in the past due to the recorded fluctuations in dust and pollen for example, whether it is an indication to changes in vegetation (pollen), or an increase in dust indicating aridity (deMenocal, 2004). Any anomalies can then be examined to see if there is correlation between a species divergence time and proposed climatic oscillations. Marine dust records are essentially core sample records that enable the study of Plio-Pleistocene African climates at a large range of timescales (10^6-10^3 years), (Trauth et al., 2009). Marine sediments accumulating off the western and eastern margins of subtropical North Africa have provided some of the most convincing evidence for oscillating arid-humid climate cycles in North Africa, with progressive step-like increases in African aridity during the late Neogene and Pleistocene (deMenocal, 2004). This African paleoclimate variability emphasises the importance of climatic instability as a mechanism for natural selection (deMenocal, 2004). These fluctuations in dust/pollen provide evidence to suggest changes in vegetation, which have been shown to affect a lizard's success within an area (Dickman et al., 1999).

Although some of the suggested climatic fluctuations fall earlier/later than the divergence times of some species discussed in the study, they are still evidence to suggest climatic oscillations throughout time that some species may have been subjected to. Oscillations that have potential to alter a species suitable habitat and therefore distribution. Many more shifts between arid and humid phases have also been suggested (Beghin et al., 2015; Chumakov, 1984; DeMenocal, 2004; Griffin, 2002; Grootes et al., 1993; Kuechler et al., 2018). These changes in climate have the potential to effect a species distribution through fragmentation, isolation, diversification and possibly extinction in parts.

Quaternary ice ages are known to have had genetic consequences for many species (Bell et al., 2010; Bush & Oliveira, 2006; Hewitt, 2000; Knowles et al., 2007; Walstrom et al., 2012). From the beginning of the Quaternary (2.4 Ma) until 0.9 Ma ice sheets have advanced and receded with an approximate 41-Ka cycle (linked to the Earths Axial tilt in relation to its plane of orbit around the sun). Thereafter

these cycles have been approximately 100-Ka (linked to the Earths eccentricity, simply, the shape of the Earth's orbit around the Sun) and have become increasingly dramatic, which involved changes of as much as 7–15°C over a few decades, which then lasted hundreds of years. The Croll-Milankovitch theory tries to explain this cycle and proposes that the Earth's orbit around the Sun is the "pacemaker" of the ice ages (Hewitt, 2000). The main orbital eccentricity has a 100-Ka cycle, deviation in the Earth's axial tilt has a 41-Ka cycle, and precession due to the axial wobble has a 19-23-Ka cycle. All these changes to the Earth's rotation around the Sun adjust the insolation of the Earth and the energy it receives. A lot of this energy is diffused/transported by the oceanic circulation system and the interaction of orbital variation and currents leads to significant climate changes. There is also evidence to suggest that prior to 2.8 Ma subtropical African climate varied at roughly 20-Ka cycles (corresponding to Earths precession), which has been linked to African monsoonal variability. Humid–arid shifts in climate are also suggested to oscillate in approximately 21Ka cycles in North Africa during the Miocene, corresponding to Earth's precession (Hewitt, 2000), again supplying potential evidence for habitat disruption due to humid-arid oscillations (Griffin, 2002). After 2.8 Ma, the African climate varied primarily at longer 41 Ka cycles. After 1 Ma eolian variability shifted towards the longer and larger 100 Ka cycles.

The "variability selection hypothesis" and "habitat specific hypothesis/savannah hypothesis" emphasize the importance of either environmental instability or paleoclimatic variability as a mechanism for natural selection (deMenocal, 2004). Climatic fluctuations had the potential to make species extinct over large areas of their habitat, to disperse species to new locations and to reduce species to small refugia. This pattern of populations retreating/fragmented into refugia then expanding is likely to occur several times, due to the glacial patterns brought on by the Earth's circumnavigation of the Sun. During times when populations were restricted/isolated to small refugia evolutionary independence will occur due to the restriction of gene flow. This could potentially lead to fewer genes successfully exchanged between the two populations during secondary contact. Eventually the two populations will become incompatible with each other, as they cannot mate successfully, resulting in the two populations regarded as two separate species (Hewitt, 2000). This type of speciation is known as allopatric. Allopatric speciation is based on physical isolations that effectively create barriers to gene flow, in many cases geographic isolation is the cause of the genetic and ecological divergence. The allopatric speciation concept was originally proposed by Mayr (1942), and appears relevant to many northwest African lizards.

Geographical isolation can occur in various ways, one is the dispersal across a physical barrier such as the ocean between two islands, followed by colonization of new habitat (Censky et al., 1998). The

second is vicariance, in which a species range is fragmented by the appearance of a new physical barrier, which could be as small as a road or as large as the ocean, or in the case of a species in this study (*A. impalearis*) the Atlas Mountains (Brown et al., 2002). A species range could also be split due to drastic changes in environmental variables, whether they are biotic or abiotic. Changes in environment could be caused by glacial patterns that expanded and receded ice sheets worldwide, this pattern of glacial expansion then retraction could cause vicariance speciation through the creation of a natural barrier. Vicariance can be a slow process, such as the rise of a mountain (Atlas), to rapid events such as manmade disruptions (road) or lava flow.

Six species were included in this study, but population divergence times are unavailable for every species, as genetic studies have not been done to estimate such times (*T. weigmanni*). Genetic studies in Morocco and northwest Africa have identified genetic splits within *Chalcides mionecton*, *Agama impalearis, Messalina olivieri, Chalcides ocellatus* and *Saurodactylus mauritanicus*, all of which are used in this study. Most species have recognised, distinct subspecies that originate from a common ancestor (Barata, 2012; Brown et al., 2012; Harris & Rato, 2008; Kapli et al., 2014; Kornilios et al., 2010). The estimated divergence times that are available for the northwest African species in this study date between 1-15 Ma. Reliable climatic data that dates back 15 Ma are unavailable (Elith & Leathwick, 2009), hence the LIG and LGM datasets are used as the climate models to act as a surrogate to earlier interglacial and glacial maximum periods which surround the divergence times of species discussed.

Paleoclimate studies around suggested speciation times

The Calabrian is a subdivision of the Pleistocene Epoch, and covers the time period for the suggested divergence of *C. mionecton* (1.43 Ma) and *C. ocellatus* (1.6-3 Ma), (Brown et al., 2012; Kornilios et al., 2010). Two periods of arid conditions suggested by deMenocal (2004) roughly coincide with these estimated divergence times. DeMenocal (2004) indicates fundamental shifts in African climate around 1.8 Ma and 1.0 Ma (deMenocal, 2004). Fundamental shifts in the African climate such as the ones described above could cause major changes to a landscape. Natural barriers such as rivers or cut off islands due to "hyper" humid phases discussed earlier may have also been possible. The African faunal and paelo-climatic studies suggest two other restricted intervals (1.8-1.6 Ma, 1.2-0.8 Ma) around the divergence times of *C.mionecton* and *C. ocellatus* (1.6-3 Ma, 1.43 Ma). These shifts toward increasingly variable, drier African conditions were accompanied by some changes in African faunal assemblages (deMenocal, 2004). The apparent increase in African aridity at 1.5 Ma matches the progressive

vegetation shift from trees and shrubs to tropical grasses (Trauth et al., 2009; Seegalen et al., 2007). Changes in vegetation are a strong indicator of a drastic change in environment/climate, theoretically causing changes in habitat with potential to force a species to migrate/disperse or fragment.

Within the North-western African clade the hotter/wetter climate of the middle Pliocene (3.6-2.5 Ma), suggested by Willis et al. (1999) quickly became colder and drier during 2.5-1.8 Ma (Webb and Bartlein, 1992). Fragmentation of plant and animal populations have been suggested to correspond to these fluctuations, and supports the major diversification events of several animal species (Mouline et al., 2008) with the evolution of arid-adapted fauna suggested to correspond to this aridification shift (deMenocal, 2004). Two species have estimated divergence times around the suggested climatic oscillations between the Mid-Upper Pliocene, *M. olivieri* (2.8 Ma) and *C. ocellatus* (1.6-3 Ma). These fundamental shifts in African climate are suggested to have initiated bursts of biotic change within Africa and are viewed to have favoured the evolution of arid-adapted fauna (deMenocal, 2004). Between 3.6 and 2.4 Ma, with a marked rise between 2.6 and 2.4 there is a documented remarkable decrease in the closed woodland and forest species and an increase in grassland species. This suggests climate change caused significant shifts in vegetation. Many of the first appearances between 2.5–3 Ma were grazing species, this supports the idea that faunal changes were linked to increased aridity which corresponds to expanding grasslands (Vrba, 1995). Arid-adapted taxa are evident, with faunal changes between 2.9 and 2.4 Ma.

One example of divergence being possibly linked to climatic oscillations within northwest Africa is a study produced by Brown et al. (2012). This study suggests there is no current evidence of unsuitable habitat between the two populations of *C. mionecton,* with a continuous pattern presently observed across the geographic space between the two subspecies (the southernmost four-digit sample from Cap Rhir and the adjacent five-digit sample from Taghazoute). This observed distribution pattern is an indication of secondary contact, with the two forms coming into contact without hybridization, an indicator towards significant genetic differences between the two populations, possibly due to allopatric speciation and the effect of restricted gene flow. This appears more likely than the hypothesis that divergence is due to an on-going restriction to gene flow mediated by a physical barrier such as the Atlas Mountains, as suggested by (Brown et al., 2002) for *A. impalearis*. Continuous patterns of distribution are observed for many of the species discussed in this study. This suggests that similar scenarios to *C. mionecton* may have occurred for the other species, which suggests speciation via allopatry.

As mentioned, it is unknown when all studied species/subspecies diverged (e.g no estimates are available for *T. weigmanni*). Therefore, any record/study that indicates a climatic fluctuation within NW Africa could have influenced a species discussed in this study. The fluctuation could have been the direct climatic event that initially caused populations to diverge, or another climatic oscillation that the already diverging populations were subjected to. Under the Croll-Milkankovitch theory of the Earth's orbit around the Sun corresponding to changes in the Earth's climate, we can infer that many climatic fluctuations will have occurred (minor and major), that species in this study may have been subjected to at some point.

Last Interglacial Climate.

The LIG is known as the last period of global climate warmth before the Holocene period, which occurred between ~120 Ka–40 Ka. However, there is some discussion on how long the duration of this period was (Cronin et al., 1999). The last interglacial is commonly understood as an interval with climate warm as or warmer than today (Kukla et al., 2002), with temperature sensitive species found in northerly regions reflecting this. Global sea levels were also suggested to be 4-6m higher (Cronin et al., 1999). A reconstruction of the LIG climate (CLIMAP, 1984) calculated a mean annual temperature increase of around 1°C. However, another study suggests that at the very beginning of the interglacial, between ~130 Ka - 127 Ka, the temperature climbed as much as 3°C higher than today (Kukla et al., 2002). Seasonality of precipitation was also predicted during the early stage of the LIG (ca. 130–119 Ka), with mineralogical, macrofossil, and pollen records suggesting a shift from mire to lacustrine conditions. This was simultaneous with the expansion of sclerophyllous vegetation (evergreen) and the presence of acicular aragonite, which indicates highly evaporative summer conditions (Milner at al., 2012). Milner (2012) also suggested that after the early LIG, summer temperatures and evaporation decreased and precipitation became less seasonal. At the end of the last interglacial period, over 100,000 years ago, glaciers grew, sea levels dropped, and deserts expanded (Kukla et al., 2002). These changes in temperature and precipitation help support the theory of arid–humid transitions and changes in vegetation around earlier periods, produced by P. DeMenocal (2004). The same transition occurred many times earlier, linked to periodic shifts of the Earth's orbit around the sun (Hewitt, 2000; Kukla et al., 2002).

Last Glacial Maximum Climate

Paleoclimate work suggests the LGM occurred approximately 21 Ka BP, with a duration of 23-18 Ka, and was generally associated with drier/cooler conditions than today (Otto-Bliesner et al., 2006). The LGM endured conditions considered generally drier in both hemispheres, with an 18% decrease in atmospheric perceptible water and an average precipitation decreasing from 2.49 mm per day to 0.25 mm per day. These annual mean changes in precipitation reflect seasonal changes associated with the Milankovitch cycle of solar insolation (Otto-Bliesner et al., 2006). A precipitation time series shows a negative shift starting at about 30 Ka BP and ending at the LGM, with estimated precipitation values are about 15-20% lower than today (Gasse et al., 1994). A global cooling of 4.5°C was estimated to have occurred with greater glacial cooling in higher latitudes but a 2°C decrease in tropical regions. A sea surface temperature (SST) decrease of 1.7°C and a topical land decrease of 2.6°C on average have also been suggested (Otto-Bliesner et al., 2006). With lower tropical land and sea-surface temperatures suggested (Gasse et al., 1994). Maley (1996) used both palynological and biogeographical evidence to construct a map detailing the limits of lowland rain forest habitat during the Quaternary at high latitudes of Africa. It exposed how different the region was at this time, with grassland and savannahs cloaking the forested areas that are around today. The work produced by Maley, (1996) indicates very different habitat in comparison to what is found today, which supports changes in flora, topography and suitable habitat throughout Africa, supported by DeMeoncal (2004) that suggested changes in flora are recorded from arid adapted vegetation to humid adapted vegetation throughout Africa's paleoclimate history.

Opinions on the precipitation within north Africa/Morocco during the LGM differ between studies/models, some models predict wetter mean annual conditions (MRI, IPSL, GISS, CNRM and CCSM4), whilst others predict annual drier LGM conditions within Morocco (FGOALS, MIROC). The differences found between these studies may be the cause of certain climatic indicators being sensitive to seasonal precipitation. The reason seasonal precipitation may affect climatic indicators is because of the proposed fluctuation of precipitation within Morocco during the LGM (Beghin et al., 2015). Beghin et al (2015) suggests higher precipitation within Morocco is based from multi proxy data, the study also indicates evidence of both semi-arid and temperate taxa found in the marines core, which supports the confliction of paleoclimate studies. This reflects the diversity of local climate and altitudes within a larger region, suggesting that precipitation is spatially dependent. The higher precipitation in Morocco/North Africa is due to the large-scale effect of the southward shift of the North Atlantic jet stream during the LGM (Beghin et al., 2015), supported in other paleoclimate studies (Kuechlet et al., 2018; Otto-Bliesner et al., 2006). Increased annual precipitation in northern Africa is

associated with increased monsoonal precipitation during July–October. Other paleoclimate studies also support this arid-humid fluctuation if north Africa/Morocco (deMenocal, 2004; Otto-Bliesner et al., 2006; Gasse& Campo, 1994).

Climates associated between and within glacial extremes.

In the African tropics and subtropics deglaciation (from the LGM) is said to have started around 17Ka (DeMenocal, 1995). This time is considerably earlier than the time glaciation occurred in the North Atlantic regions, indicating that the effects of glaciation were less severe in the tropics. A paleoclimate study indicated that the last deglacial period in Africa was not only sooner, but happened in a series of sudden conversions between arid and humid environments (Gasse et al., 1994). These sudden changes between arid and humid environments had the potential to have drastic changes to an environment over a short period.

Major climatic events are suggested around 17-16, 15-14.5 and 11.5-11 Ka (Gasse et al., 1994); these arid-humid oscillations indicate that drastic climatic changes occurred within glacial cycles as well as between. DeMenocal, (1995) suggested that the North Western Sahara would have experienced much wetter conditions than today during interglacial periods, highlighting environmental changes with potential to alter a species suitable habitat. As deglaciation continued climatic conditions become wetter through North Africa when compared to today. Major dry spells are also recorded between ~11-9.5 Ka, ~8-7Ka and ~4-3Ka (Gasse et al., 1994). These records again provide evidence for abrupt climatic fluctuations within deglaciation periods, and support theories of potential species divergence for species in this study due to climatic oscillations. Evidence also suggests conditions were much wetter than today in the North Western Sahara around 6Ka. A time period that fits between two dry spells previously mentioned, again portraying this pattern of arid-humid fluctuations.

Rationale

Reviewed publications concerning the genetics of study species within this study have estimated speciation events within Moroccan Lizards that have occurred during the last 1.4–15 million years (Brown et al., 2002; Brown et al., 2012; Harris & Rato, 2008; Kapli et al., 2015; Kornilios et al., 2010). These study species show either intraspecific divergence or are recognised as recently-diverged sister species. Currently very few modelling studies have addressed the cause of the speciation events within Morocco, with some studies suggesting that speciation events could be climatically mediated through range shifts (Brown et al., 2012; DeMenocal, 2004; Knowles et al., 2007), whilst other papers suggest vicariance due to changes in physical barriers such as mountains (Brown et al., 2002).

Paleo-climate studies have shown numerous climatic oscillations to have occurred in North Africa (Beghin et al., 2015; Chumakov, 1984; deMenocal, 2004; Trauth et al., 2009; Seegalen et al., 2007; Webb & Bartlein, 1992), with some of the suggested oscillations corresponding to estimated speciation times (Brown et al., 2002; Brown et al., 2012; Carretero et al., 2005; Harris & Rato, 2008; Kapli et al., 2015; Kornilios et al., 2010; Mattiucci et al., 2001). This supports theories of speciation events within North Africa to be climatically mediated (deMenocal, 2004; Otto-Bliesner, 2006).

The aim of this study was to identify if changes in climate associated with LIG and LGM conditions had potential to change a species distribution, which may have led to climatically mediated divergence events. Climatic data that predate the LIG and LGM are considered too coarse to be effective in modelling past species distributions. Therefore, the LIG and LGM datasets are used as a surrogate to earlier glacial maximum and interglacial climates, under the Croll-Milankovitch theory that proposes the Earth's orbit around the Sun is the "pacemaker" of the ice ages and the climate associated with them. Under the Croll-Milankovitch theory any patterns of fragmentation/isolation predicted are said to have been potentially repetitive patterns over millions of years.

Methods

Species Studied

Six species were selected for this study that all have distributions within Morocco. Previous studies have shown genetic divergence within lizards in this region (Brown et al., 2002; Brown et al., 2012; Harris & Rato, 2008; Kapli et al., 2015; Kornilios et al., 2010). This subsequently led to all six species to having identified subspecies. The six species possess quite different population sizes, locations and ranges within Morocco for variance between species and the habitat they inhabit. Some species have a large population with a broad geographical range *(A. impalearis)* whilst others may have a small population with a narrow geographical range (*T. weigmanni*). The distribution patterns between sister subspecies in the study also differ. Some share a continuous pattern with each other (*C. mionecton*) whilst others are geographically isolated from one another (*C. ocellatus*). Lizards are ectotherms and rely on their surroundings to keep warm, they are considered sensitive to temperature and other abiotic variables such as precipitation (Hertz et al., 1982), variables that are easily quantifiable and easier to source for modelling (WorldClim). Lizards are also considered to have low dispersal abilities which is a potential limitation for SDM's, and often specific ecological niches, making them accurate indicators of past climatic conditions (Camargo et al., 2010). Paleoclimate studies also provide evidence of climatic oscillations around the speciation times for many species in this study (Chumakov, 1984; deMenocal, 2000; deMenocal, 2004; Griffin, 2002; Trauth et al., 2009). Paleoclimate work on NW Africa will therefore help with the interpretation of results.

Species observed within Morocco were chosen for this study because distributional data were available for this region. Each species was modelled using the same environmental layers and projected onto the same environmental layer (projection layers) this is done so all species will be affected by the same historical processes, beneficial in highlighting general historical events/general patterns that may have affected species in north Africa. The observational data for this study originated from Bons & Geniez (1996) derived from survey data.

The observation maps of the species used is this study are produced below and a table of information produced after.

- *Chalcides mionecton*

Chalcides mionecton
- subsp. *mionecton*, 4 doigts
- subsp. *mionecton*, 5 doigts
- subsp. *trifasciatus*
Altitude: 0 - 700 (1200) m
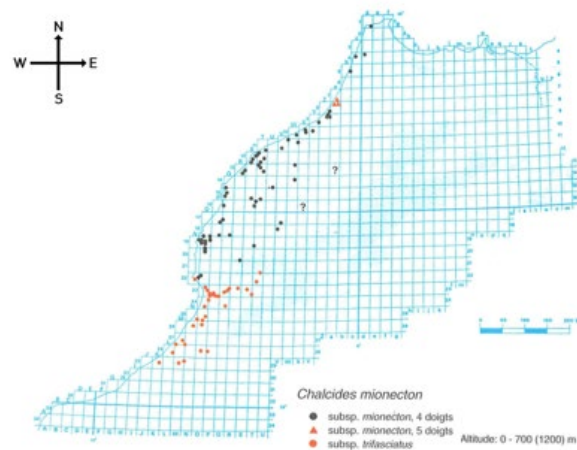
(Figure 1.0). Distribution map for *Chalcides mionecton*, sampled from (Bons & Geniez, 1996).

C. mionecton was selected for this study as the distribution for this species is narrow, suggesting a small niche for the species. A recent study suggests that this species split into a distinct subspecies, *C. m. mionecton* and *C. m. trifasciatus* (Brown et al., 2012), with allopatric speciation suggested as the cause. The geographic variation in digit number has been used to describe two subspecies: The subspecies *C. mionecton mionecton* Böttger 1874 is referred to as the northern form because it is applied to populations from Tangiers in the north to Cap Rhir ~700 km to the south-west (Schleich et al. 2006). It is characterised by possession of four digits. The southern form, *C. mionecton trifasciatus* Chabanaud 1917, has five digits and has been recorded from Cap Rhir south to Foum Assaka (Bons, 1959). The southern subspecies possesses 5 digits, but this state was also common in individuals from the extreme north. Only four digit specimens were found in between, giving a disjunctive pattern. Anatomically, four-toed *C. mionecton mionecton* possess a 5th metatarsal bone but lack distal or proximal phalanges for this toe.

- *Chalcides ocellatus*
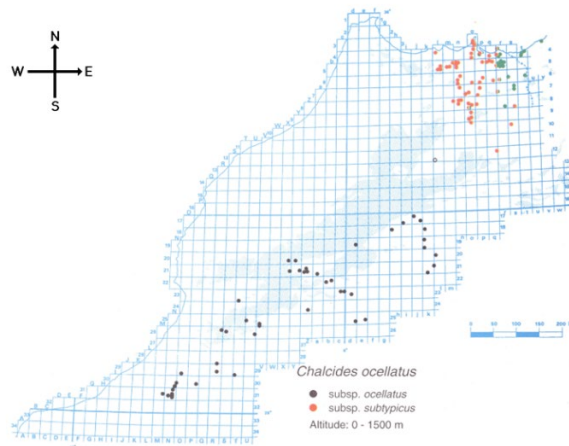
(Figure 1.1). Distribution map for *Chalcides ocellatus*, sampled from Bons & Geniez, (1996).

C. ocellatus comprises well-defined sub species, with *C. ocellatus subtypicus* observed on the North coast of Morocco and *C. ocellatus ocellatus* observed inland in the South of Morocco. Similar to *S mauritanicus* the two subspecies are distant from one another with clear differences between them in terms of habitat i.e. coastal to inland, North-South, showing no continuous pattern.

- *Messalina olivieri* (Audouin, 1829)
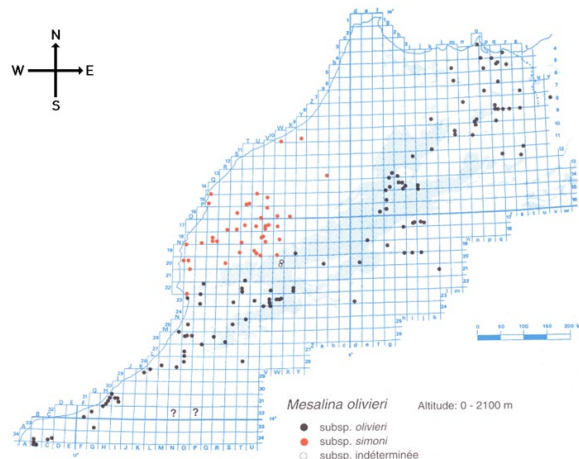
(Figure 1.2). Distribution map for *Mesalina oliveri* sampled from Bons & Geniez (1996).

In Morocco *M. olivieri olivieri* and *M. olivieri simoni* have a continuous pattern geographically, an indication that under certain conditions the species inhabit areas close to one another and gene flow

would be possible, a considerable difference in range size as well as population size between these two subspecies. The present continuous pattern is similar to *C. mionecton*.

- *Saurodactylus mauritanicus*
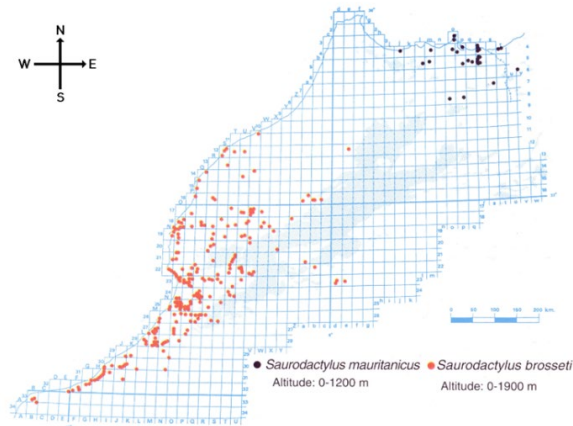
(Figure 1.3). Distribution map for *Saurodactylus mauritanicus*, sampled from Bons & Geniez (1996).

S. mauritanicus was chosen for this study as it has quite a small distribution, particularly the subspecies *mauritanicus*. Presently the observed data shows that the two subspecies are distant from one another, an indication of a difference in niches between the two subspecies. This observed distribution is different to many of the species in this study such as *C. mionecton* and *M. olivieri*, which have a continuous pattern, a variation in size of distributions and a variation in patterns of distribution; this is why it was chosen for this study. *S. mauritanicus* are an extremely agile species and found only in sandy habitat.

- *Trogonophis weigmanni* (Kaup, 1830)
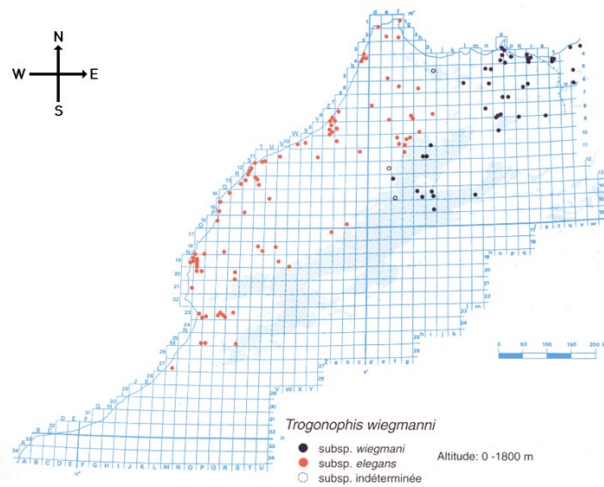
(Figure 1.4). Distribution map for *Trogonophis wiegmanni*, sampled from Bons & Geniez (1996).

T. weigmanni was selected as the species' current distribution possesses both coastal and inland habitats for both subspecies and resides in small areas suggesting a specific niche. The two populations are observed next to one another producing a continuous pattern, suggesting under certain climatic conditions gene flow between the two populations may have been possible, a distributional pattern the ancestral species may have had. A study suggests the cause of divergence could be due to the rise of the Atlas Mountains during the Miocene (Sánchez & Escoriza, 2014).

- *Agama impalearis*
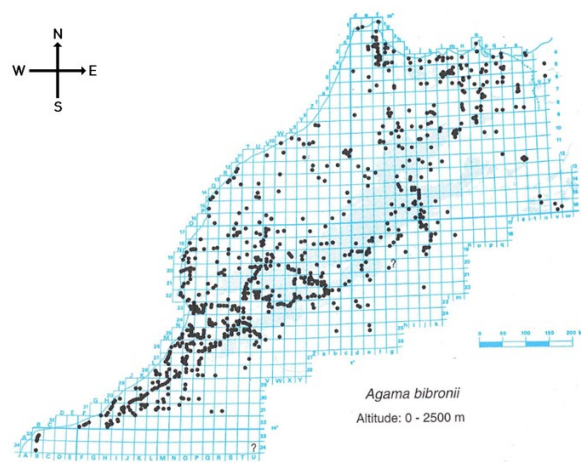
(Figure 1.5). Distribution map for *Agama impalearis* sampled from Bons & Geniez (1996).

The above figure shows Agama *impalearis* as one species, without the population being split into sister subspecies. A study indicates genetic divergence within the population (Brown et al., 2002), which suggests the two populations diverge around the Atlas Mountains in Morocco, the population was there for split manually around the Atlas Mountains for this study.Below is an image of how the split has been interpreted in this study. (The pink points represent Northern *A. impalearis* occurrence points while the yellow points represent Southern *A. impalearis* occurrence points)
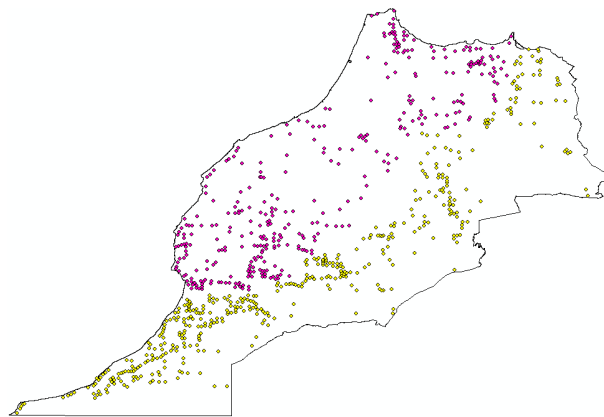
A. impalearis Boettger *, 1874* was chosen for this study as the two populations of the subspecies span a large area geographically and therefore could incorporate a large variance of environmental variables. *A. impalearis* is presently widespread across Morocco suggesting a much broader niche in comparison to *C. mionecton*. This sort of diversity is important for the study as it allows inferences to be made about how different species with different sized distributions and constraints would have managed under LIG and LGM conditions. Genetic work on *A. impalearis* has also shown a genetic split, indicating past disturbances in the species population such as fragmentation (Brown et al., 2002).

(Table 2). The table below gives particular information on species discussed in this study. Ecology information was not available for every subspecies, with only ecological information on the species only. Divergence times also not known for every species in this study.

Species	Range description	Elevation	Divergence	Ecology
Mesalina olivieri olivieri	Morocco, Algeria, Western Sahara, Tunisia, Libya, Egypt, Sinai, Israel, Jordan, S Iraq, N Saudi Arabia	Found from sea level up to 2,100 m	2.8 Ma	Habitat of the species is on flat terraces of open ground on soil with stones and shrubs; sandy or loamy soils with Frankenia thymifolia and Zygophyllum album; Halfa grass steppes; and sandy regions with rocks in the Saharan region. *M. olivieri* is considered an opportunistic feeder, their diet consists of all insect groups in the biotope; spiders, mites (*Acarina*); snails.
Mesalina olivieri simoni	Morocco, Algeria, Western Sahara, Tunisia, Libya, Egypt, Sinai, Israel, Jordan, S Iraq, N Saudi Arabia	Found from sea level to above 3000 ft.		Inhabits rocky plateaus with sparse vegetation. As it has only recently been differentiated from M. olivieri, details on ecology are still lacking (Boettger, 1881, Yousefkhani et al. 2015).
Chalcides ocellatus	This very widespread species occurs in northern and northeastern Africa and west Asia, and in a few parts of southern Europe.	Found from sea level up to 2,500 m asl (Morocco)	1.6 - 3 Ma (2.4 Ma)	Often associated with arid to moist sandy, open areas or places with sparse scrub. It is found in coastal dunes and stabilized vegetated dunes further inland and It is found in orchards, vineyards, fields, oases, rural gardens and urban areas. *C. ocellatus* can remain active during the hottest hours even in June and July. They remain active into dusk with an active temperature range of 28 - 37˚C.
Agama Impalearis	This species ranges from the Atlantic coastal region of Western Sahara, through most of Morocco to northern Algeria (Trape *et al.* 2012).	Found from sea level up to 2,500 m	9 Ma	Found in rocky areas, where it is active by day and forages for athropods, although it will also feed on small lizards and on plants.(Trape *et al.* 2012). In Morocco it has also been reported from Mediterranean vegetation, steppe, and areas of suitable habitat at the margins of cultivated land
Chalcides mionecton	This species is largely found on the Atlantic coastal area of Morocco. It is a lowland species that is known to occur up to 700 m asl.	A lowland species that is known to occur up to 700 m asl.	1.4 Ma	Found in open areas of loose, slightly moist sand and soil, where it may be found under rocks and at the base of sparse vegetation. It can be found at the margins of fields, and in courtyards, gardens and parks. It is also known from the Argan Forest.
Trogonophis wiegmanni	This species ranges from western Morocco (including Ceuta, Melilla and the Chafarinas Islands [Spain]), eastwards through northern Algeria into northwestern Tunisia.	Can be found from sea level up to 1,900 m asl.	Unknown	Found in moist or dry soil that is covered by stones, rocks and other ground cover. It can be found close to roadsides, in traditionally cultivated areas, grassland, in oak forest and oak-juniper forests, in steppe habitat and in sandy patches without vegetation.
Saurodactylus mauritanicus mauritanicus	This species is found in northeastern Morocco, northwestern Algeria, and the Spanish territories of Melilla, the Isla de Alborán and the Charfarinas Archipelago.	Can range from sea level up to 1,200m asl.	15 Ma	Found in semi-arid and arid rocky areas, with a sloping gradient. Animals can be found in piles of stones and sheltering in ground vegetation.
Saurodactylus mauritanicus brosseti	This species ranges from northern coast of Western Sahara, along the southwestern coast of Morocco, extending inland in Morocco to the western slopes of the Atlas Mountains and the Dra Valley as far east as Zagora. There are some isolated inland population in northeastern Western Sahara and also north of the Atlas Mountains in Morocco. It might occur in extreme western Algeria although its presence here needs to be confirmed.	Occurs from sea level up to 1,900 m asl.		Found in various stony or rocky areas, including degraded agricultural lands (where it can be found under stones). It has also been recorded from semi-forested areas.

Observation records

The observational data for this study were based on one book (Bons & Geneiz, 1996). Observational records were geo referenced in ArcGIS by matching a Moroccan (georeferenced) shape file with the distribution maps from Bons & Geneiz (1996), once the two images are aligned points can be marked in ArcGIS at every occurrence point, these marked points can then be assigned the correct longitude and latitude co-ordinates. ArcGIS was also used to create the study/sample area (Morocco) under a georeferenced digital layer so the program predictions/background points were sampled only from the areas required. The geographic location of an animal is vital to the model as the model extracts the values of each predictor variable (abiotic) used in the study at each observation point, variable importance is calculated based the variables at each location.

Climate data

The current climate data was sourced from worldclim.org which has a timeline of 1950-2000, at a resolution of 30 arc seconds in the format of an ESRI file. Two Paleoclimate datasets used in this study were also sourced from worldclim.org.

The last interglacial (LIG; ~120,000 - 140,000 years BP) sourced from (Otto-Bliesner et al., 2008), and the last glacial maximum (LGM; ~21,000 years BP) was created from Paleoclimate Modelling Intercomparison Project Phase II (PMIP2). All bioclim variables were used in this study, the Bioclim variables available for all time periods are the same, these are:

BIO1 = Annual Mean Temperature

BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp))

BIO3 = Isothermality (BIO2/BIO7) (* 100)

BIO4 = Temperature Seasonality (standard deviation *100)

BIO5 = Max Temperature of Warmest Month

BIO6 = Min Temperature of Coldest Month

BIO7 = Temperature Annual Range (BIO5-BIO6)

BIO8 = Mean Temperature of Wettest Quarter

BIO9 = Mean Temperature of Driest Quarter

BIO10 = Mean Temperature of Warmest Quarter

BIO11 = Mean Temperature of Coldest Quarter

BIO12 = Annual Precipitation

BIO13 = Precipitation of Wettest Month

BIO14 = Precipitation of Driest Month

BIO15 = Precipitation Seasonality (Coefficient of Variation)

BIO16 = Precipitation of Wettest Quarter

BIO17 = Precipitation of Driest Quarter

BIO18 = Precipitation of Warmest Quarter

BIO19 = Precipitation of Coldest Quarter.

Correlation between bioclim variables may be an issue when modelling and bias could be introduced. This issue has been analysed through a PCA to determine the correlation between variables. See appendix for details of the PCA using bioclim data.

Data Formats

MaxEnt is designed to integrate smoothly with GIS programs such as Arc GIS or GRASS GIS, by using the ESRI ASCII grid format (text files representing raster data). MaxEnt will also take input in simple csv file format, generally the format environmental and species occurrence data are sourced in. MaxEnt insists on species names in the first column along with geocoordinates (longitude and latitude). Followed by environmental variables that the niche will be determined by in the second column saved as .csv files. Projection layers allow MaxEnt to predict a species distribution over different environmental scenarios, whether that is a past, future or hypothesised environment. The format for these files is also .csv files, like the environmental layer (environmental variables).

Environmental variables were sourced in a BIL format. These files need to be converted to ASCII files that are more amenable. Two conversions were done through DivaGIS; BIL to grid (GRD) format, and GRD to ASCII. To geo-reference the observational data the program ArcGIS was used. This program allows the user to set markers onto a georeferenced digital layer and then identify coordinates from the markers that represent places a lizard/animal was observed. ArcGIS is also used to convert ASCII files, the form MaxEnt uses for output results. ArcGIS converts ASCII files to raster which is better for presentational purposes as reclassifications of images can be done from a raster dataset.

Study Area
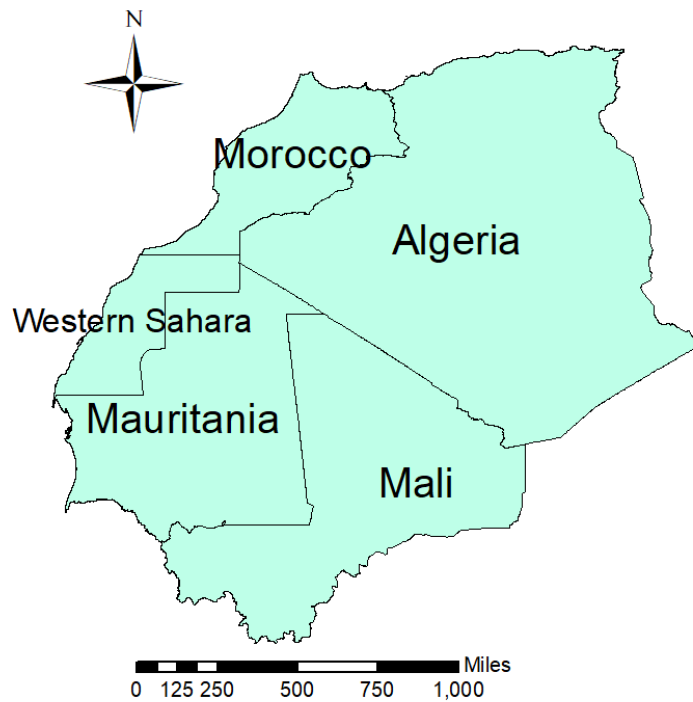
(Figure 2). Image to show the countries that are included in the study area. Another image has also been produced that shows the study area overlaid over NW Africa to show surrounding countries (Appendix P).

Morocco is the country all species observation data was sourced, with a large body of information on the distributions of species within Morocco, although other areas poorly known (Bons & Geniez, 1996; Schleich, 1996). Morocco was therefore used as the area to estimate the species *potential* niche, to use for modelling and hindcasting (Bons & Geniez, 1996). When then extrapolating over the LIG and LGM climate scenarios a larger geographical range was chosen (NW Africa), comprising of Algeria, Mali, Mauritania Western Sahara and Morocco. For another representation of the geographical area chosen to hindcast the niche over can be found in the appendix (Appendix Q). NW Africa was the area used to project the LIG and LGM predictions to allow for range expansions/dispersion beyond current distributions. The area of Morocco was used when hindcasting one species in this study rather than the NW Africa region and that was *Chalcides mionecton,* the reason for this was that there is no observational data for this species anywhere else other than Morocco, an indication that the species has always resided in Morocco. Morocco is one of the most biodiverse region in North Africa (Bons & Geniez, 1996), which has a Mediterranean and sub-Saharan climate, with a current mean precipitation ranging from 300 to 600mm (Michard et al., 2008). Morocco is also mountainous, this coupled with varied precipitation and climates makes it an interesting region for biodiversity and speciation.

Modelling

MaxEnt was used in this study due to its efficiency with presence-only modelling using background data to produced background points or pseudo absences randomly generated (from study area) (Phillips et al., 2004; Phillps et al., 2006). MaxEnt was also used due to advantages the maximum entropy principle has with incomplete datasets. This is advantageous given that this study used only abiotic variables and presence only data. This statistical method performed well in model comparison studies (Anderson et al., 2006), which produces generally better results than other presence-only SDM methods (Heikkinen et al., 2006). MaxEnt also performs well even when the number of observation records available are small (Wisz et al., 2008). For research and studies on species distribution, MaxEnt probability distributions follow a concise mathematical definition. This makes it is easy to understand and therefore amenable for interpretation and analysis. Allowing more validation techniques and methods to be performed on the models helps ensure a high level of model validation. With current research using MaxEnt a catalogue of advice and literature is available that describes any disadvantages/advantages within their research/model and different approaches to consider with MaxEnt under different scenarios. It's beneficial to use a popular modelling method, for additional help from other researches and studies. In this report it is assumed the SDMs measure the *potential* distribution, as we cannot incorporate all the variables into the model (proximal/biotic variables for example).

In this study subspecies were modelled both separately and modelled combined with sister subspecies (to act as an ancestral species). When combining the subspecies together the climatic tolerances of this population may be less extreme and could resemble the climatic tolerances of an ancestral species before speciation. Two model types were used as no definitive way (model subspecies individually or model subspecies together) has been given to suggest one method being better than the other and so by doing both model types comparisons between these two models can be made and similar patterns highlighted.

Assumptions made

Niche conservation has been assumed in this project, in other words the assumption is made that a species environmental tolerances to both abiotic and biotic are identical to their ancestors. This assumption is usually necessary, as information on the environmental tolerances of an ancestral species are generally unknown and unavailable to source. *Niche conservation* is

useful for species distribution as a niche can be derived using the present population, and be set as a predicted ancestral niche. Used to project onto past climatic scenarios.

Another assumption made is the assumption that ice ages/glacial patterns are similar to one another in relation to environmental impact and duration (Chumakov, 1984; Kukla et al., 2002; Otto-Bliesner et al., 2006). This assumption is based on the Croll-Milankovitch theory that suggests the regular variations in the Earth's orbit around the Sun are the "pacemakers" of the ice-age cycles. The assumption that earlier glacial patterns are similar to the LGM and LIG aids in the analysis of how a species may distribute in time periods where data are not available and allow more accurate datasets such as the LIG and LGM to act as a surrogate for other/earlier interglacial and glacial maximum climates.

A source habitat is also assumed in this study, this is the assumption that the current habitat of a species allows the population to thrive, and that a species occupies the maximum land presently available to the species (equilibrium between species and habitat, including equilibrium with predators and competitions). This assumption is made as it becomes difficult to try quantify whether a habitat is source or not and to what extent, due to factors such as competition and predation etc, being notoriously hard to quantify/source. A species ability to adapt to changes in the degree of predation and competition is a factor usually reserved for analysis or assumed negligible, mainly because this interaction is hard to quantify and introduce to a model. However if data are available predation and competition are an additional consideration in outlining a species true ability to reach an equilibrium state throughout suitable habitat (Holt & Gaines, 1992; Kawecki, 1995; Holt, 1996; Holt & Gomulkiewicz, 1997). This assumption does not take into account dispersal or migration, therefore the model has no way to measure the species ability to move across the landscape.

MaxEnt options

MaxEnt is able to run a model multiple times, averaging the results from all models created, improving the statistical validity of a model prediction. Executing multiple runs also provides a way to measure the amount of variability in the model. For this study, the number of replicates was 15, suggested in Phillips (2005). The random test percentage option allows users to evaluate model performance in MaxEnt. This setting allows MaxEnt to use a certain percentage of your presence data to evaluate the models performance. Without setting data aside MaxEnt will randomly select a requested percentage from the observational data. For this study, 75% was used to build the model and 25% set aside for testing. This allows the program to use statistical

analysis such as regression and AUC. These percentage choices are typical values for this type of analysis and considered optimal (Phillips, 2005). The run type used for this study was the bootstrap method. This was due to the small occurrence records for some subspecies in this study (see appendix N). For this study, the threshold value chosen was one that corresponds to equal sensitivity and specificity, and is one that is considered a robust method (Liu, 2005). In this study, the background concept is preferred which uses background points rather than pseudo absences as it requires fewer assumptions and has some coherent statistical methods for dealing with the "overlap" between presence and background points (e.g. Ward et al. 2009; Phillips and Elith, 2011). The number of background points was 1000, randomly sampled in MaxEnt from the background data distributed over Morocco. Morocco was the decided area from where to sample the background points because the extent of Morocco is equivalent to the extent as the observed presence points (Bons & Geniez, 1996). A Cross validation run was also performed to see how the two different run types compared under the same dataset. This method didn't perform as well due to the model type underperforming with small datasets (sample sizes) (Wisz et al, 2008), a comparison table to compare different run types/feature selection has been made (tables 6-8). For this project, I decided to use the default settings (auto features). This automates the task of choosing feature types, using an empirical algorithm based on sample size. Auto features was chosen due the variance of population size between subspecies in this study, the sample size and features selected for each subspecies can be found in the appendix (see appendix N). A model was also produced for comparison using only the hinge feature; this was done to see how the models compared with the auto features model (see table 7). The hinge only threshold was used as a comparison as it is suitable for datasets with a small amount of observational data. Hinge features provide a generalization of linear and threshold features. A model the uses only hinge features produce complex but smoothed response curves that are much like GAMs (Elith et al., 2010). A sum of hinge features is always a piece-wise linear function, so if only hinge features are used, the MaxEnt exponent is piece-wise linear.

Model Validation

For this study the AUC, jackknife, sensitivity/specificity analysis and null hypothesis evaluation were the model validation techniques used as these four statistics cover most factors concerning a distribution model, that are available to extract from this model. As MaxEnt used presence only data, the AUC scores define the relationship between presences vs background points, rather than presence vs absences (this is because absences are unknown), these are said to be slightly inflated values (Yackulic et al., 2013). Therefore, although the AUC scores do indicated good model performance in this study, the AUC scores have been primarily used to compare

different models within this study. As mentioned elsewhere the AUC coupled with other model validation procedures is optimal. PCA and niche overlap analysis were also performed as a niche comparison tool to identify differences between the niches of species and respective sister subspecies.

Principle component Analysis (PCA)

In this study PCA was performed on subspecies separately. For this the program SPSS was used. A PCA was done to see if there are any differences between sister subspecies in terms of variable importance. Potentially indicating niche divergence. PCA was performed in this study to further describe and quantify differences in niches between subspecies. Data extraction was performed at all observation points for each subspecies, exact values of all 19 bioclim variables were extracted at these points to help portray a picture of the constraining factors of each subspecies. A correlation matrix was produced for each species and respective subspecies (See appendix G-L). A scree plot of the eigenvalues was plotted to see how many components are needed explain the variance within the data. Scatter graphs were produced for each species plotting one component against another to see how they interact in the component space; these graphs help visually see similarities and differences within the component space, these images can be viewed later in the results section.

Another PCA was performed as a method to measure correlation between the bioclim variables (See Appendix O). Data extraction was executed on every point a bioclim variable was located within the study area (NW Africa). SPSS was used to handle the large dataset and a correlation matrix and scree plot where used to determine the amount of components that were significant. Four components were significant judged by the scree plot, which explained 93% of the cumulative variance. The components group variables with similar importance that explains the majority of variance but minimises the correlation between variables. These components were then used as the environmental layer in MaxEnt instead of the individual Bioclim variables. These components were used as the environmental layers to see if there was any significant difference in MaxEnt outputs between models that used the bioclim variables and the components that were comprised of bioclim variables (when correlation between variables had been addressed and an effort made to try reduced the correlation). The same settings were used in MaxEnt as the original model.

Niche overlap

For each species and respective subspecies, Schoener's D values have been produced from tables that present niche breadth overlap values using ENM tools. MaxEnt is used in conjunction with ENMtools to define niches of subspecies and to use the Schoener's D statistic to determine the amount of niche overlap, these values range from 0 (niche models have no overlap) to 1 (niche models identical). These tables can be found in the appendix and the mean values calculated and put in a table found in the results section.

Null hypothesis

A null hypothesis is a method used to determine whether the model produces results that are considered better than random. To perform a null hypothesis a number of random species locations have to be produced in ENM tools then saved. Each data set having the same number of presence points as the original species' data set. These datasets with the randomly selected species locations now act as other independent datasets of specie observations. The 99 randomly calculated observation datasets are considered a good number with the 100th dataset being the actual species dataset, obtained from (Bons & Geniez, 1996). The 99 randomly calculated occurrence data sets can then be used in MaxEnt where an evaluation of AUC values can take place.

For this study, I produced 99 random sample datasets using the program ENMTools, these datasets were then run through MaxEnt, only 1 replicate for each dataset instead of the initial 15 and the amount of iterations was kept the same. Histograms have been produced for each of the species; each histogram contains the AUC values of the randomly sampled observations (99), calculated using ENMTools shown in grey, and the actual AUC value from the authentic dataset coloured red (see results). The actual AUC value that was calculated using Maxent essentially acts as 1%, and can therefore be compared with the randomly observed 99%. Producing histograms for each species containing the 99 random datasets shows how the model has performed against a random sample.

Sensitivity and specificity analysis

To perform sensitivity and specificity analysis a threshold had to be chosen, there are many methods on choosing a threshold with reasons for each, some of these methods are: Fixed value; where a value is chosen based on no previous statistical work, usually this value would be 0.5. Lowest predicted; a method of selecting a threshold that corresponds to the lowest logistic value that is assigned to a recorded presence and the fixed sensitivity method; a method of choosing a threshold that corresponds to a 95% accuracy of the true positive rate. For this study, the threshold value chosen was one that corresponds to equal sensitivity and specificity and is one that is considered a robust method (Liu, 2005). Equal sensitivity and specificity means that the model uses the threshold to essentially allow the model to perform equally well at both predicting true positives as well as true negatives. The MaxEnt outputs were used to obtain the values needed for the sensitivity and specificity analysis (true presence, true absence, false presence and false absence) and then calculations were performed using the necessary formulae. Sensitivity and Specificity analysis was performed on the pseudo ancestral species and on individual subspecies (see appendix). An equal sensitivity and specificity threshold was also used to change the MaxEnt outputs to a binary colour code, where the options are present or absent, again this threshold was derived from the MaxEnt outputs.

Results

Jack knife

Table 3. Jackknife results produced through MaxEnt as a measure of variable importance.

Species	Variable with biggest contribution towards model	Percentage contribution	Variable with biggest contribution that isn't represented by other variables
Jack knife results for variable importance			
A. impalearis	Maximum tempertaure of warmest month (Bio 5)	20.2	Mean Diurnal Range
C. mionecton	Temperature annual range (Bio 7)	39.5	Precipitation of Driest Month
S. mauritanicus	Isothermality (Bio 3)	24.5	Seasonal precipitation
C. ocellatus	Annual precipitation (Bio 12)	45.9	Precipitation of warmest quarter
M. olivieri	Precipitation of coldest quarter (Bio 19)	35.6	Precipitation of warmest month
T. wiegmanni	Precipitation of coldest quarter (Bio 19)	38.9	Mean temperature of wettest quarter

From examining the Jackknife results (subspecies modelled together to create potential ancestral niche) the variables that are considered to contribute highly towards the explanation of a certain species habitat differ. These results can then provide evidence and reasoning to potential differences in the predicted suitable habitat for different species.

Table 3.1. Table below shows the jackknfe results from MaxEnt when subspecies were treated individually.

Jack knife results for variable importance (Subspecies)		
Species	**Variable with biggest contribution towards model**	**Percentage contribution**
Northern *A. impalearis*	Precipitation of wettest mounth	23.1
Southern *A. impalearis*	Precipitation of coldest quarter	31.1
C. m. mionecton	Temperature annual range	34.3
C. m. trifasciatus	Temperature seasonality	29.4
S. m. mauritanicus	Mean Diurnal Range	48.2
S. m. brosseti	Isothermality	19
C. o. ocellatus	Annual precipitation	32.2
C. o. subtypicus	Mean Diurnal Range	37.7
M. o. olivieri	Precipitation of wettest mounth	15.2
M. o. simoni	Precipitation of coldest quarter	28.3
T. w. elegans	Precipitation of coldest quarter	34.9
T. w. wiegmanni	Precipitation of warmest quarter	47.3

From comparing the jackknife results of the two different model types, it's clear that in some of the models it seems a certain subspecies has dominance within the model. This observation is based on certain subspecies models showing the same variable importance as the whole subspecies models, while the other species does not.

Null hypothesis Evaluation

Graphs to present Null hypothesis evaluation

Agama impalearis
Actual AUC value: 0.7399
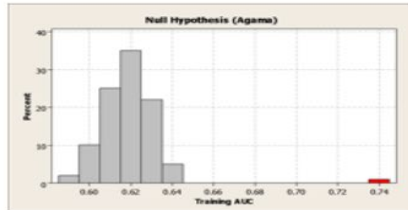
(Figure 2.1). Graph to represent the AUC values from the Null hypothesis evaluation of *Agama impalearis*.

Trogonophis wiegmanni
Actual AUC value: 0.9136
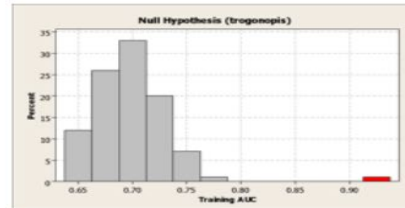
(Figure 2.2). Graph to represent the AUC values from the Null hypothesis evaluation of *Trogonophis wiegmanni*.

Chalcides mionecton
Actual AUC value: 0.9526
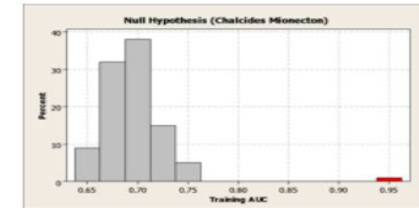
(Figure 2.3). Graph to represent the AUC values from the Null hypothesis evaluation of *Chalcides mionecton*.

Chalcides ocellatus
Actual AUC value: 0.9212
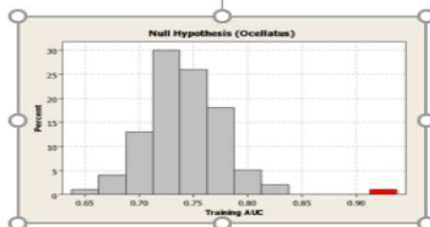
(Figure 2.4). Graph to represent the AUC values from the Null hypothesis evaluation of *Chalcides ocellatus*.

Mesalina olivieri
Actual AUC value: 0.932
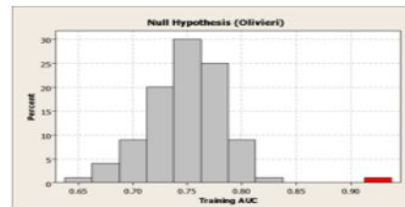
(Figure 2.5). Graph to represent the AUC values from the Null hypothesis evaluation of *Mesalina olivieri*.

Saurodactylus mauritanicus
Actual AUC value: 0.9095
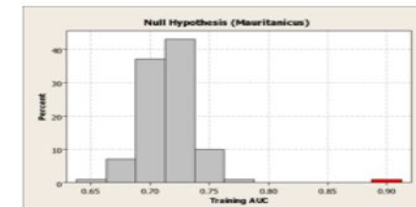
(Figure 2.6). Graph to represent the AUC values from the Null hypothesis evaluation of *Saurodactylus mauritanicus*.

(Figure 3) Graphs to highlight the differences in AUC values when compared with random datasets produced for the null hypothesis evaluation.

The majority of the AUC values produced for the null hypothesis are above 0.90 for the actual model, therefore showing a high level of statistical significance, an indicator that the models produced in this study via MaxEnt are strong, evidence they adhere to the initial constraints apposed on them. What this means in terms of this study is that the predicted models should behave in the same way the original data (observations) interacts with the data (environmental), allowing for a better analysis to be performed. The Null hypothesis shows all models perform better than random which shows validity towards the program MaxEnt. Other studies also suggest using the null-hypothesis as an extra step in model validation (Raes & Steege, 2007).

Sensitivity and Specificity Analysis

For this study, the threshold value chosen was one that corresponds to equal sensitivity and specificity and is one that is considered a robust method (Liu, 2005). The threshold value differs for each species and can be derived from the results MaxEnt produces. Sensitivity and specificity analysis was also performed on subspecies individually, those tables of values can be found in the appendix (Appendix M).

Table 4. Table to represent the sensitivity and specificity scores, along with false positive/negative rate values for each species.

Species	Equal sensitivity & Specificity threshold	Sensitivity	Specificity	False positive rate (α)	False negative rate (β)
C. mionecton	0.2	0.947	0.866	0.133	0.052
A. impalearis	0.29	0.946	0.287	0.103	0.822
M.saurodactylus	0.23	0.957	0.756	0.244	0.433
M. olivieri	0.25	0.929	0.804	0.196	0.071
C. ocellatus	0.3	0.933	0.770	0.230	0.067
T. weigmnanni	0.18	0.944	0.700	0.301	0.056

Sensitivity (true presence) values were high for all species in this study, a mean of 94%, the specificity (true negative) values were significantly lower with a mean of 69%, the reason for this drop is the considerably lower value produced for *A. impalearis*. This could be due to the vast range in which *A. impalearis* inhabits, potentially making it statistically difficult for the programme MaxEnt to define a true negative scenario under the constraints and information known on *A. impalearis* as a whole species. The low specificity of *A. impalearis* may also be due

to human error when splitting the observation records manually using a GIS program, leading to the misidentification of a subspecies. Appendix M compares the sensitivity and specificity scores between subspecies and ancestral whole species in table format.

Niche overlap analysis

Table 5. Table to show the Schoener's D values for each species. Figures derived through ENMtools in conjunction with MaxEnt to calculate niche overlap between subspecies and respective sister subspecies. Based on the range of values classed by Rödder, D & Engler (2011), the niche overlap scores for species in this study indicate very limited to low niche overlap, with all values falling within the range 0-0.4.

A measure of niche overlap between sister subspecies; Schoener's D value			
Species		Mean	S.D
Chalcide mionecton		0.268	0.164
Chalcide ocellatus		0.079	0.005
Messalina olivieri		0.399	0.008
Trogonophis weigmanni		0.162	0.009
Saurodactylus mauritanicus		0.184	0.013
Agama impalearis		0.299	0.013

Table 6. Tables below compare model performance with three models that have different environmental layers; the original environmental layer comprising of the bioclim data, a model using 3 principle components derived from Bioclim data and another using 4 components derived from Bioclim data.

Comparison of environmental variable selection in MaxEnt

Original BioClim data as environmental variables			PCA: 3 Components used as environmental variables			PCA: 4 Components used as environmental variables		
Species	Auto Features		Species	Auto Features		Species	Auto Features	
	AUC	S.D		AUC	S.D		AUC	S.D
C. m. mionecton	0.970	0.006	*C. m. mionecton*	0.950	0.010	*C. m. mionecton*	0.949	0.007
C. m. trifasciatus	0.980	0.003	*C. m. trifasciatus*	0.927	0.130	*C. m. trifasciatus*	0.958	0.007
C. o. ocellatus	0.942	0.014	*C. o. ocellatus*	0.866	0.025	*C. o. ocellatus*	0.900	0.014
C. o. subtypicus	0.976	0.006	*C. o. subtypicus*	0.938	0.008	*C. o. subtypicus*	0.963	0.007
A. i. north	0.890	0.004	*A. i. north*	0.840	0.012	*A. i. north*	0.851	0.009
A. i. south	0.891	0.006	*A. i. south*	0.807	0.012	*A. i. south*	0.842	0.008
T. w. elegans	0.918	0.012	*T. w. elegans*	0.879	0.010	*T. w. elegans*	0.880	0.011
T. w. weigmanni	0.964	0.007	*T. w. weigmanni*	0.889	0.026	*T. w. weigmanni*	0.933	0.012
S. m. mauritanicus	0.982	0.002	*S. m. mauritanicus*	0.963	0.011	*S. m. mauritanicus*	0.981	0.005
S. m. broseseti	0.867	0.004	*S. m. broseseti*	0.899	0.009	*S. m. broseseti*	0.918	0.006
M. o. olivieri	0.921	0.011	*M. o. olivieri*	0.826	0.020	*M. o. olivieri*	0.870	0.012
M. o. simoni	0.949	0.008	*M. o. simoni*	0.889	0.021	*M. o. simoni*	0.906	0.017

As you can see from the tables above, the AUC values and the standard deviation values are highest for the model that used the bioclim variables as the environmental layer. All other settings were the same in MaxEnt; 25% of data to testing, 15 replicates, 1000 iterations and equal sensitivity and specificity threshold applied. The areas of predicted habitat calculated in MaxEnt for models using the principle components were all in areas similar to the model that used the bioclim variables as the environmental layer, however reduced in size. Judging by the AUC scores produced in MaxEnt the results deemed less accurate. SPSS outputs for the PCA work on bioclim variables can be found in the appendix.

Table 7. Two tables comparing some feature choices. The "Auto feature" option used to produce the actual models, using an empirical algorithm based on sample size (chosen for this study). The second table using "Hinge" only features. For details on what features were selected using the "auto feature" option see appendix N.

Comparison of MaxEnt feature options							
Model Run	**Subspecies**	**AUC**	**S.D**	**Model Run**	**Subspecies**	**AUC**	**S.D**
MaxEnt: **Bootstrap**, 25% test, 15 replicates, equal spencificity & sensitivity threshold, **Auto features**	*C. m. mionecton*	0.970	0.006	Maxent: **Bootstrap**, 25% test, 15 replicates, equal spencificity & sensitivity threshold, **Hinge only features**	*C. m. mionecton*	0.968	0.005
	C. m. trifasciatus	0.980	0.003		*C. m. trifasciatus*	0.976	0.004
	C. o. ocellatus	0.942	0.014		*C. o. ocellatus*	0.942	0.011
	C. o. subtypicus	0.976	0.006		*C. o. subtypicus*	0.975	0.005
	A. i. north	0.890	0.004		*A. i. north*	0.848	0.009
	A. i. south	0.891	0.006		*A. i. south*	0.853	0.007
	T. w. elegans	0.918	0.012		*T. w. elegans*	0.920	0.011
	T. w. weigmanni	0.964	0.007		*T. w. weigmanni*	0.961	0.008
	S. m. mauritanicus	0.982	0.002		*S. m. mauritanicus*	0.989	0.003
	S. m. broseseti	0.867	0.004		*S. m. brosetti*	0.905	0.006
	M. o. olivieri	0.921	0.011		*M. o. olivieri*	0.856	0.013
	M. o. simoni	0.949	0.008		*M. o. simoni*	0.947	0.015

Generally, the model that used the auto feature option performed better than the model that used the hinge only feature. Higher AUC scores where predicted for *S. m. brosseti* (0.038), *S. m. mauritanicus* (0.007) and *T. w. elegans* (0.002) using the hinge only feature. The suitable habitat predicted by MaxEnt when using the hinge only feature were very similar to the predicted habitat MaxEnt produced using the auto feature option.

Table 8. Two tables comparing two run type options from MaxEnt. The first table are values that relate to the Bootstrap run type. The second table are values that relate to the Cross-validation run type. All other options in MaxEnt were identical.

Model Run	Subspecies	AUC	S.D			AUC	S.D
Maxent: **Bootstrap**, 25% test, 15 replicates, equal spencificity & sensitivity threshold, **Auto features**	*C. m. mionecton*	0.970	0.006	Maxent: **Cross-validation**, 25% test, 15 replicates, equal spencificity & sensitivity threshold, **Auto features**	*C. m. mionecton*	0.951	0.016
	C. m. trifasciatus	0.980	0.003		*C. m. trifasciatus*	0.958	0.028
	C. o. ocellatus	0.942	0.014		*C. o. ocellatus*	0.836	0.062
	C. o. subtypicus	0.976	0.006		*C. o. subtypicus*	0.962	0.017
	A. i. north	0.890	0.004		*A. i. north*	0.810	0.004
	A. i. south	0.891	0.006		*A. i. south*	0.819	0.014
	T. w. elegans	0.918	0.012		*T. w. elegans*	0.874	0.056
	T. w. weigmanni	0.964	0.007		*T. w. weigmanni*	0.930	0.070
	S. m. mauritanicus	0.982	0.002		*S. m. mauritanicus*	0.983	0.014
	S. m. broseseti	0.867	0.004		*S. m. brosetti*	0.875	0.029
	M. o. olivieri	0.921	0.011		*M. o. olivieri*	0.774	0.046
	M. o. simoni	0.949	0.008		*M. o. simoni*	0.895	0.077

As you can see from the table above the AUC and standard deviation values are higher from the model that used the Bootstrap rather than the Cross-validation run type. Cross-validation was the model run used as the comparison as it is generally considered a good option to use and is the default option within MaxEnt.

Table 9. Table below shows the AUC and S.D values produced for the whole species dataset through MaxEnt. The values produced are generally lower than when the subspecies are treated individually. However these AUC scores are still considered significant. All other MaxEnt options aside from the observational data were kept the same.

Whole species (Ancestral species) Model		
Species	Auto Features	
	AUC	S.D
C. mionecton	0.952	0.003
C. ocellatus	0.904	0.014
A. impalearis	0.782	0.006
T. wiegmanni	0.936	0.009
S. mauritanicus	0.927	0.007
M. olivieri	0.903	0.013

Principle component analysis

The Principal component analysis (PCA) results are also below for each species to highlight the differences in niches between subspecies.

- ***Chalcides mionecton***

For *Chalcides mionecton* three components were selected and examined that accounted for the majority of the variance (89%): PC1 (44% of variance), PC2 (27% of variance), PC3 (18% of variance).

Variables with high positive loadings on PC1 are associated with precipitation during the dry periods of the year. These variables are precipitation of warmest quarter (0.111), precipitation of driest quarter (0.109) and precipitation of driest month. Variables with high negative loadings are associated with temperature during the coldest period of the year. These variables are mean temperature of coldest quarter (-0.114), mean temperature of wettest quarter (-0.111), minimum temperature of coldest month (-0.105) and mean annual temperature (-0.93).

Variables with high loadings on PC2 are associated with temperature during the warmer periods of the year. These variables are maximum temperature of warmest month (0.19), mean temperature of driest quarter (0.185) and mean temperature of warmest quarter (0.184).

Variables with heavily weighted loadings on PC3 are associated with precipitation during the wetter period of the year. These variables are precipitation of wettest month (0.254), precipitation of wettest quarter (0.246), precipitation of coldest quarter (0.245) and annual precipitation (0.216). See Appendix G for the component score coefficient matrix.
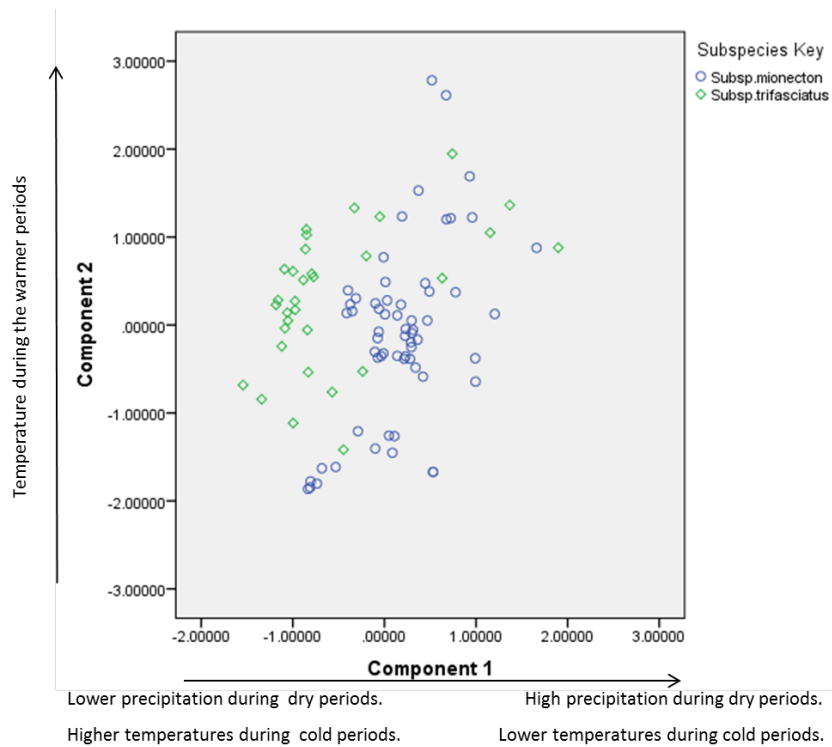
Figure 4: Component scores of component 1 (PC1) v component 2 (PC2) for PCA on *Chalcides mionecton*.

There is considerable overlap between *C. mionecton mionecton* and *C. mionecton trifasciatus* in the component space. Small differences can be seen in figure 4, subspecies *mionecton* has a broader range across PC2, which is heavily loaded by temperature during warmer periods, while subspecies *trifasciatus* has a broader range across PC1, associated with precipitation during dry periods and temperature during cold periods, although cluster more to the right of component space.
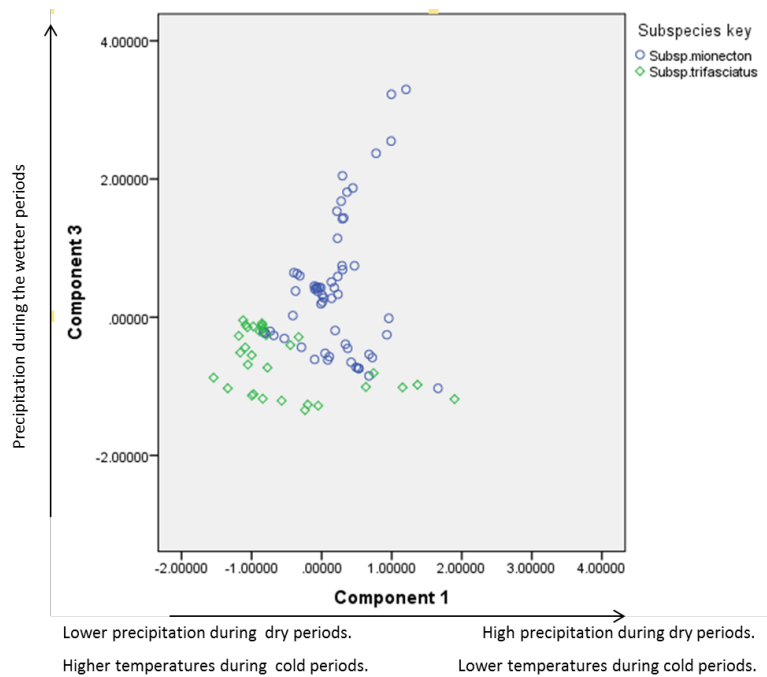
Figure 4.1: Component scores of component 1 (PC1) v component 3 (PC3) for PCA on *Chalcides mionecton.*

Separation between *C. mionecton mionecton* and *C. mionecton trifasciatus* is also visible along PC3. This indicates *C. mionecton mionecton* is associated with higher precipitation particularly during the wetter periods. The separation between subspecies seen in the PCA (Figure 4 and 4.1) correspond to the low overlap in niche breath (Appendix A).

- ***Chalcides ocellatus***

For *C. ocellatus* two components were selected for analysis: PC1 (54% of variance) and PC3 (10% of variance. PC2 shows separation along PC1 as do the other two graphs however for PC2 the range for both species are very similar.

Variables with high positive loadings on PC1 are associated with temperature during the warmest periods of the year. These variables are mean temperature of warmest quarter (0.9), annual mean temperature (0.9) and maximum temperature of warmest month (0.83). For PC1 variables with high negative loadings on PC1 are associated with precipitation. These variables are precipitation of coldest quarter (-0.86), precipitation of wettest quarter (-0.89), precipitation of wettest month (-0.89), annual precipitation (-0.93) and precipitation of warmest quarter (-0.88).

Variables with high loadings on PC3 are associated with seasonal differences and the associated temperature and precipitation differences. These are isothermality (0.319), mean temperature of driest quarter (0.315) and precipitation seasonality (0.232). This component has one heavily weighted negative value, mean temperature of wettest quarter (-0.273). See Appendix H for the component score coefficient matrix.
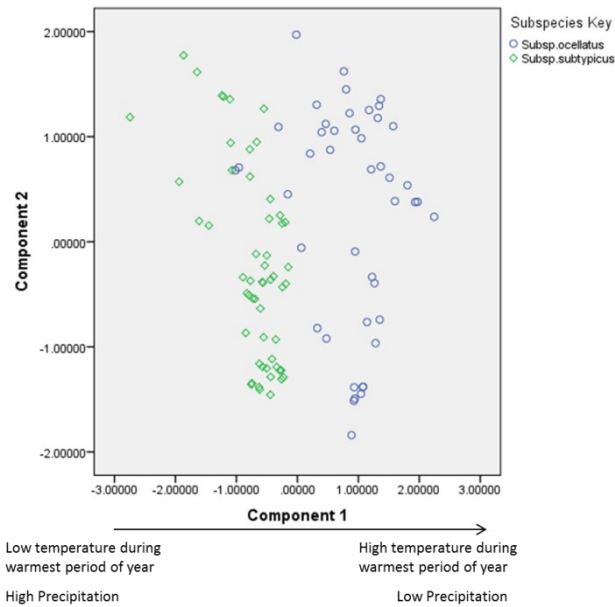
Figure 4.12: Component scores of component 1 (PC1) v component 2 (PC2) for PCA on *Chalcides ocellatus.*

The two subspecies show clear separation along PC1. The separations show that *C. ocellatus subtypicus* is associated with higher precipitation and cooler temperatures during the warmer periods of the year. *C. ocellatus ocellatus* is associated with a warmer climate with less precipitation.
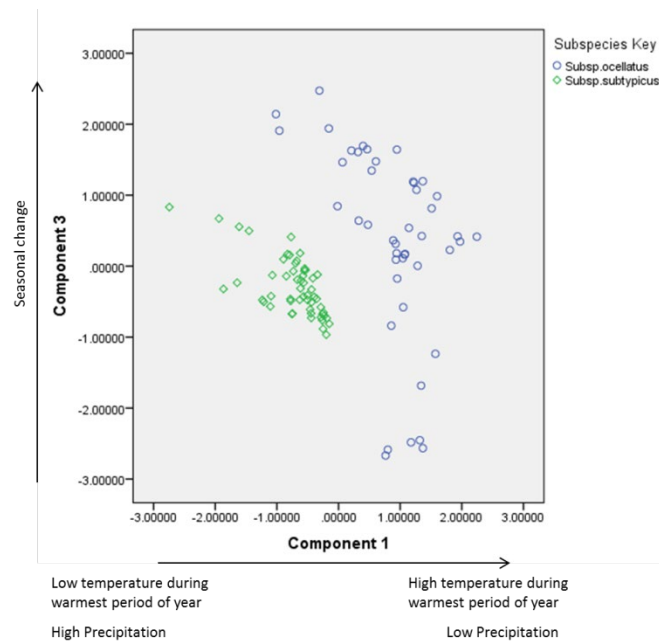
Figure 4.13: Component scores of component 1 (PC1) v component 3 (PC3) for PCA on *Chalcides ocellatus.*

It is clear that the ranges of the two subspecies along PC3 differ, heavily loaded by seasonal change (Figure 4.13). For *C. ocellatus subypicus* the range is quite narrow when compared to *C. ocellatus ocellatus,* indicating an association to higher sensitivity to seasonal change for *C. ocellatus subtypicus*. The separation between subspecies indicated through the PCAs (Figure 4.12 and 4.13), correlate to the lack of niche equivalency in the niche overlap analysis.

- ***Messalina olivieri***

For *Messalina olivieri* 3 principle components were selected for analysis PC1 (53% of variance), PC2 (25% of variance) and PC3 (11% of variance) representing a total of 89%.

Variables with high positive loadings on PC1 are associated with precipitation during the warmest periods of the year. These are precipitation of driest quarter (0.086), precipitation of warmest quarter (0.085) and annual precipitation (0.082). The negatively loaded variables correlated with temperature during the coldest period of the year. These variables are minimum temperature of coldest month (-0.094), mean temperature of coldest quarter (-0.096) and mean annual temperature (0.086).

Variables with high positive loadings on PC2 are associated with temperature during the warmest periods of the years. These variables are maximum temperature of warmest month (0.191) and mean temperature of warmest quarter (0.184).

Variables with high positive loadings on PC3 are associated with precipitation during the wettest and coldest part of year. These variables are precipitation coldest quarter (0.258), precipitation of wettest quarter (0.239), precipitation of wettest month (0.234) and mean temperature of driest quarter (0.218). Those which are negatively correlated with large loadings relate to the precipitation during the driest part of the year, the largest variable being precipitation of driest month (-0.25). See Appendix I for the component score coefficient matrix.
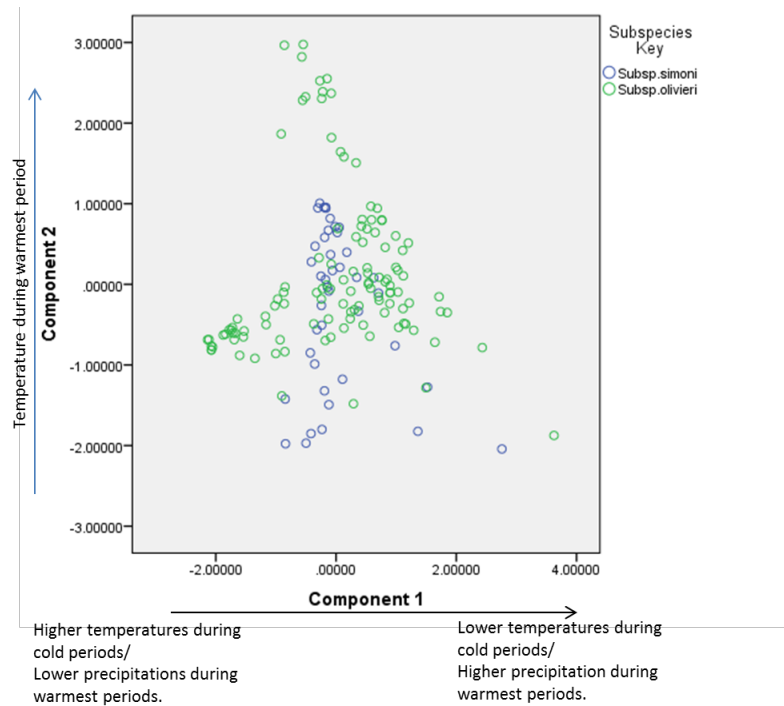
Figure 4.14: Component scores of component 1 (PC1) v component 2 (PC2) for PCA on *Messalina olivieri.*

Some discrimination between subspecies can be seen from the above graph, PC2 in particular. This component represents higher temperature during the warmest part of the year. *M. olivieri olivieri* shows a broad range of scores along component 1 in comparison to *M. olivieri simony.* This indicates subspecies *olivieri* are found in environments with broader ranges of temperatures during cold periods and precipitation during warm periods.
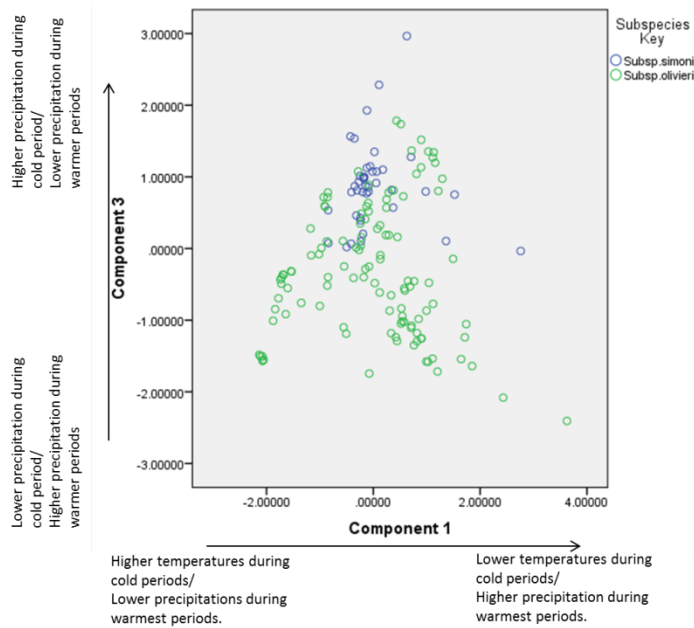
Figure 4.15: Component scores of component 1 (PC1) v component 3 (PC3) for PCA on *Messalina olivieri.*

There is separation between subspecies *simoni* and *olivieri* along PC3. This indicates *M. olivieri simoni* is associated with wetter conditions during colder parts of the year and lower precipitation during the warmer periods in comparison to *M. olivieri olivieri*. This is represented by the areas of no overlap in the top and bottom of the component space.

- ***Trogonophis weigmanni***

For *Trogonophis weigmanni* 3 components were selected for analysis and further discussion due to the percentage of the variance explained through each component. PC1 (51% of variance), PC2 (21% of variance) and PC3 (19% of variance), a collective representation of 91%.

Variables with high positive loadings on PC1 are associated with precipitation during the warmer periods of the year. These are precipitation of driest quarter (0.1), precipitation of warmest quarter (0.098) and precipitation of driest month (0.097.) The negative loadings in this component are associated with temperature during the colder periods of the year. These are mean temperature of coldest quarter (-0.102), Minimum temperature of coldest month (-0.101) and annual mean temperature.

Variables with high loadings on PC2 are associated with precipitation during the wetter colder periods of the year. These are precipitation of coldest quarter (0.238), precipitation of wettest quarter (0.237), precipitation of wettest month (0.236) and annual precipitation (0.214).

Variables with heavily weighted loading on PC3 are associated with temperatures during the warmer periods of the year. These are mean temperature of driest quarter (0.254), mean temperature of warmest quarter (0.249) and maximum temperature of warmest month (0.247). See Appendix J for the component score coefficient matrix.
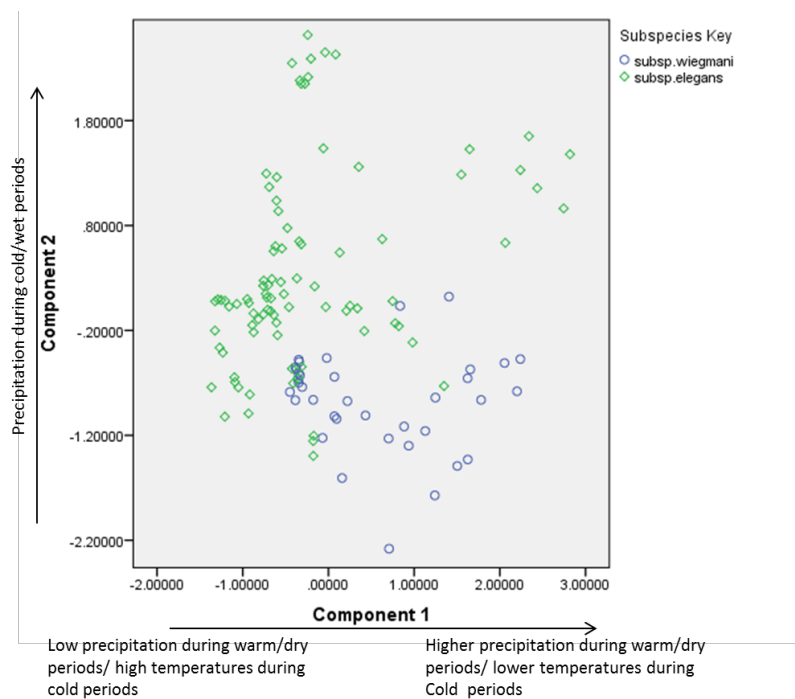
Figure 4.16: Component scores of component 1 (PC1) v component 2 (PC2) for PCA on *Trogonophis wiegmanni.*

The above graph shows separation along component 2. *T. wiegmanni elegans* has a cluster of points near top left of the component space where overlap between subspecies is visible. This indicates that *T. wiegmanni elegans* is found in environmental conditions of higher precipitation during the colder periods, which would suggest *T. wiegmanni wiegmanni* is associated with lower precipitation during the same periods. Along component 1 there is a lot of overlap between the two subspecies which makes it difficult to derive anything definitive from the component.
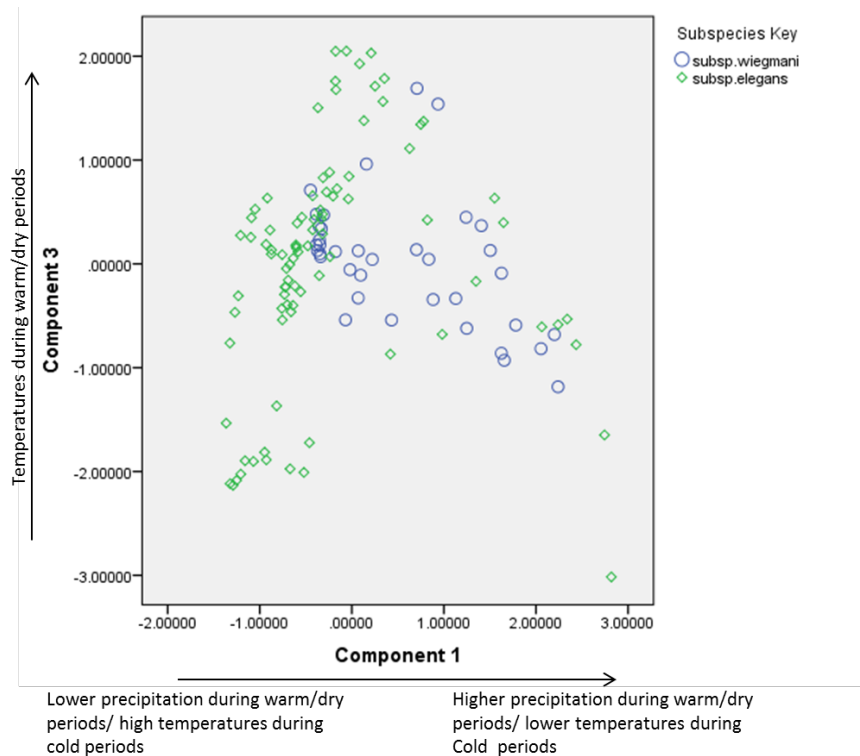
Figure 4.17: Component scores representing Component 1 (PC1) v Component 3 (PC3) for PCA on *Trogonophis wiegmanni.*

The range of *T. wiegmanni elegans* along PC3 spans that of *T. wiegmanni wiegmanni*. This would indicate that *T. wiegmanni elegans* is found in environments with higher temperature variation, in comparison to *T. wiegmanni wiegmanni* which is associated to higher temperatures during warmer periods.

- ### *Saurodactylus mauritanicus*

For *Saurodactylus mauritanicus* 3 components were selected for analysis PC1 (59% of variance), PC2 (24% of variance) and PC3 (7% of variance).

Variables with high positive loadings on PC1 are associated with precipitation, particularly during the warmer parts of the years. These are annual precipitation (0.082), precipitation of driest quarter (0.081) and precipitation of warmest quarter 0.081. The negative loadings for PC1 are associated with temperature, particularly the colder/wetter parts of the year. These are mean temperature of coldest quarter (-0.086), minimum temperature of coldest month (-0.082) and mean temperature of wettest quarter (-0.081).

Variables with high loadings on PC2 are all positively correlated and are associated with temperatures during the warmer/drier parts of the year. These are mean temperature of warmest quarter (0.204), max temperature of warmest month (0.203) and mean temperature of driest quarter (0.197). See Appendix K for the component score coefficient matrix.
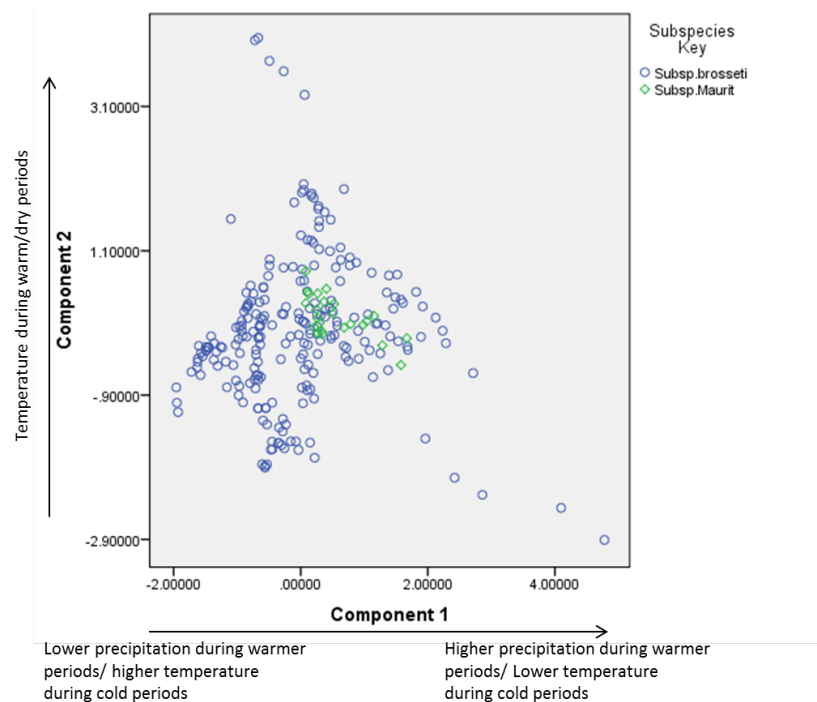
Figure 4.18: Component scores representing Component 1 (PC1) v Component 2 (PC2) for PCA on *Saurodactylus mauritanicus*.

S. mauritanicus. mauritanicus and *S. mauritanicus. brosseti* overlap within the component space on both PC1 and PC2 showing no definitive separations. *S. mauritanicus brosseti* is clustered more toward the left of the component space along PC1, although spans the width of the component space. This indicates that *S. mauritanicus brosseti* is found in environmental conditions where temperatures fluctuate more, and has more tolerance to changes in temperature when compared to *S. mauritanicus mauritanicus.* Along PC2, *S. mauritanicus brosseti* spans the component space whereas *S. mauritanicus mauritanicus* clusters towards the middle of the component space, again showing that *S. mauritanicus brosseti* is tolerable to a broader range of temperatures during warmer periods, whereas *S. mauritanicus mauritanicus* is not as comfortable with low or high temperatures when compared to *S. mauritanicus brosseti.* The difference in the amount of sample points however makes it quite difficult to determine how much variation there is between the two sister subspecies. The size

difference in observation records for subspecies *mauritanicus* may be a problem in the PCA, potentially unable to determine anything definitive due to the size difference (Wisz et al., 2008).

- ***Agama impalearis***

For *Agama impalearis* 3 components were selected for analysis PC1 (47% of variance), PC2 (28% of variance) and PC3 (13% of variance).

Variable with high positive loadings on PC1 are associated with precipitation, particularly during the warmer parts of the year. These are precipitation of driest quarter (0.905), precipitation of warmest quarter (0.892) and precipitation of driest month (0.798). The negatively loaded variables are associated with temperature, particularly the colder/wetter parts of the year. These are mean temperature of coldest quarter (-0.956), minimum temperature of coldest month (-0.886), annual mean temperature (-0.879) and mean temperature of wettest quarter (-0.803).

Variables with high loadings on PC2 are all positively correlated and are associated with temperatures during the warmer/drier parts of the year. These are maximum temperature of warmest month (0.911), mean temperature of warmest quarter (0.79), temperature annual range (0.764) and temperature seasonality (0.755).

Variable with high loadings on PC3 are all positively correlated and correspond to precipitation, particularly during the colder/wetter parts of the year. These are precipitation of coldest quarter (0.617), precipitation of wettest quarter (0.616), precipitation of wettest month (0.606) and annual precipitation (0.486). See Appendix L for the component score coefficient matrix.
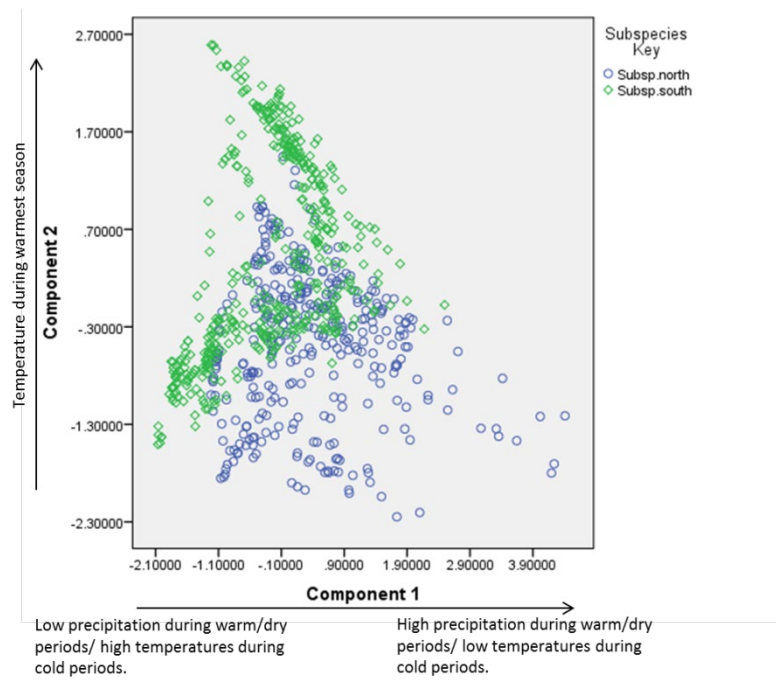
Figure 4.19: Component scores representing component 1 (PC1) v component 2 (PC2) for PCA on *Agama impalearis.*

The above scatter diagram shows some separation of clusters. PC1 shows a lot of overlap between subspecies along this component, however it does seem to show *A. impalearis. north* to be in environments where precipitation is higher during the warmer periods with lower temperatures during colder periods, indicated by lack of overlap near the end of the component space along PC1. PC2 shows more separation within the component space; along PC2 it shows *A. impalearis south* nearer the top with a horizontal separation between subspecies. An indication *A. impalearis south* is associated with warmer temperatures with emphasis on seasonal changes, when compared with *A. impalearis north*. Both PC1 and PC2 highlight *A. impalearis south* to be associated to warmer drier conditions.
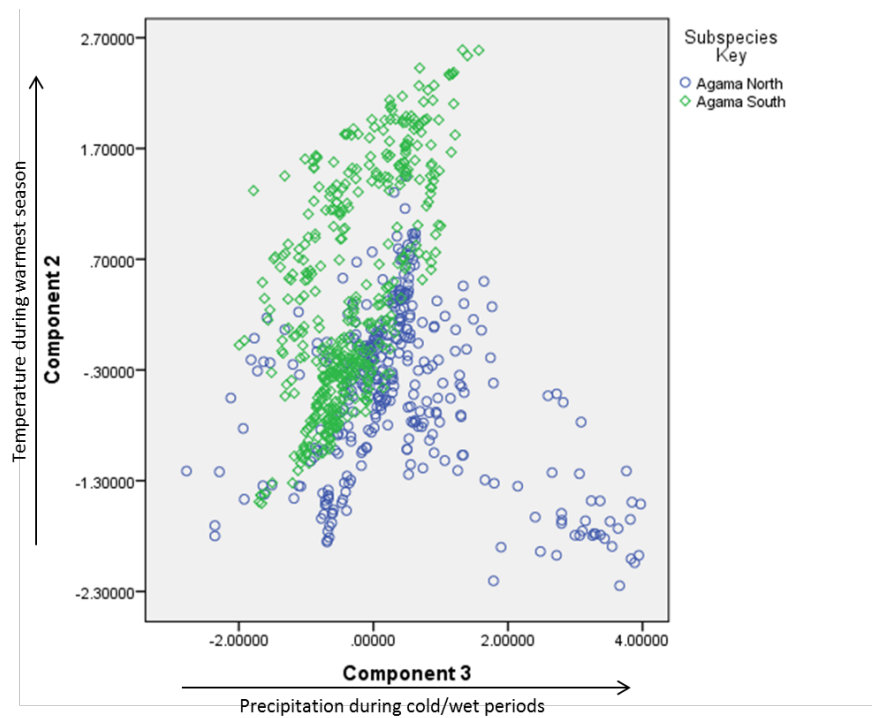
Figure 4.2: Component scores representing component 3 (PC3) v component 2 (PC2) for PCA on *Agama impalearis*.

The above scatter diagram shows a clear separation along PC3, this indicates *A. impalearis north* is associated with higher precipitation particularly during the colder/wetter parts of the year in comparison to *A. impalearis south*, indicated by the lack of overlap near the end of PC3. Figure 4.2 also indicates that *A. impalearis south* can tolerate higher temperatures in comparison to *A. impalearis north*, derived from the slight separation along PC2, more evident separation in Figure 4.19.

MaXent outputs

An equal sensitivity and specificity threshold has been added to all images to give a binary presence/ non-presence presentation, the threshold used was the calculated equal training sensitivity and specificity threshold (Liu, 2005).
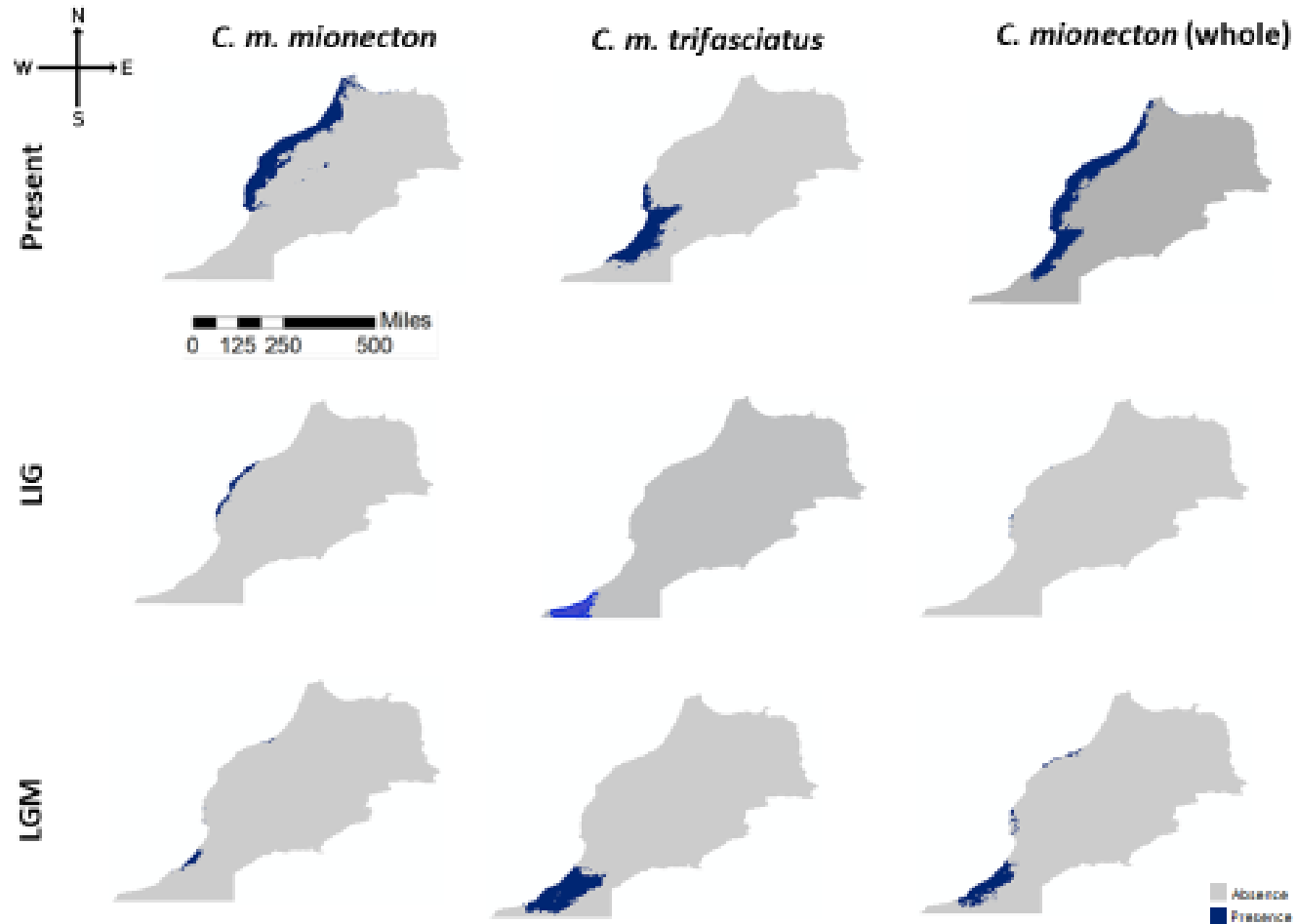
Figure 4.30 : Predicted suitable habitat for C. mionecton and subspecies under present, LIG and LGM conditions.
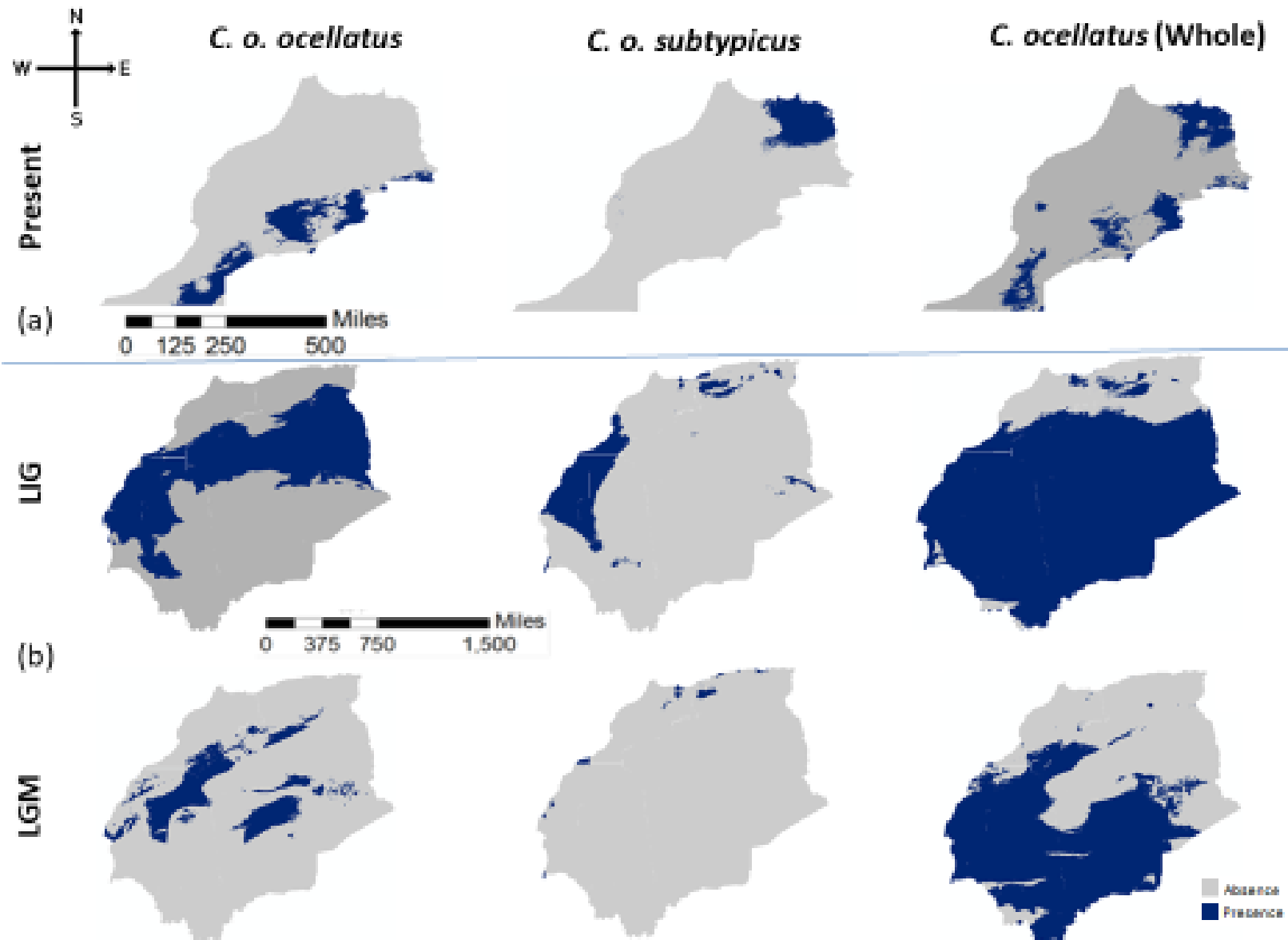
Figure 4.31: Predicted suitable habitat for *C.ocellatus* and subspecies under present (a), and LIG and LGM conditions (b).
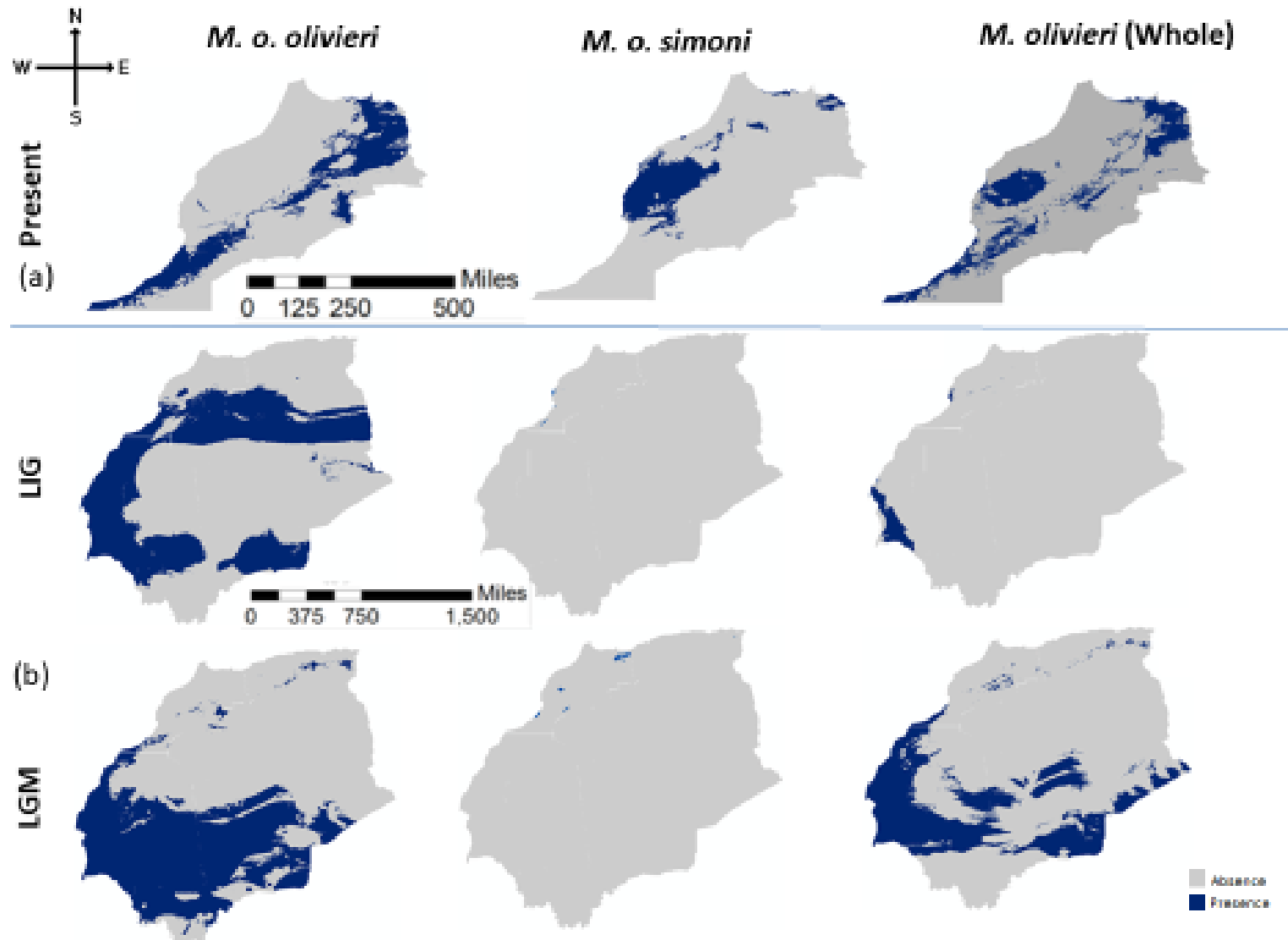
Figure 4.32: Predicted suitable habitat for *M. olivieri* and subspecies under present (a), and LIG and LGM conditions (b).
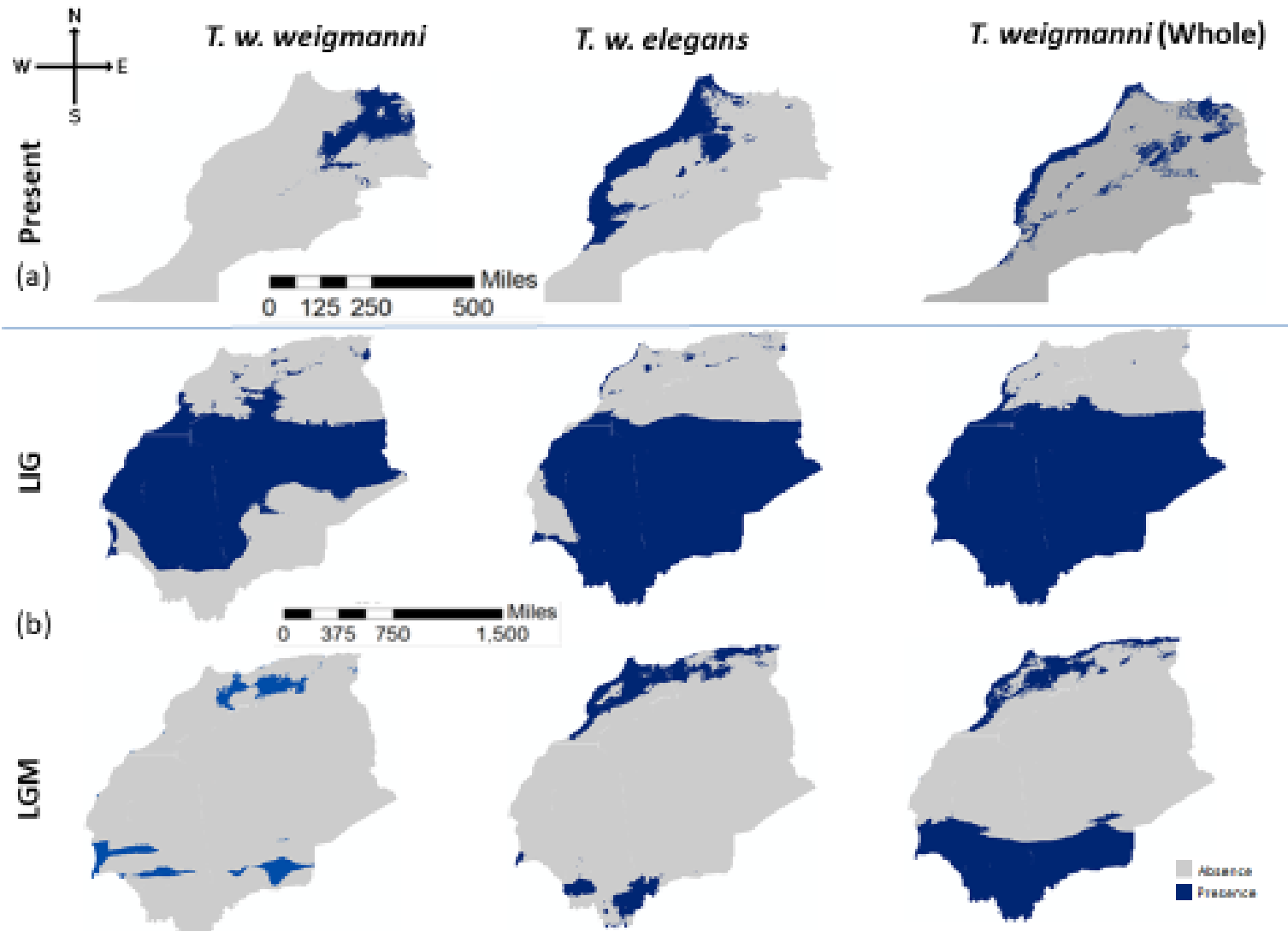
Figure 4.33 : Predicted suitable habitat for *T. weigmanni* and subspecies under present (a), and LIG and LGM conditions (b).
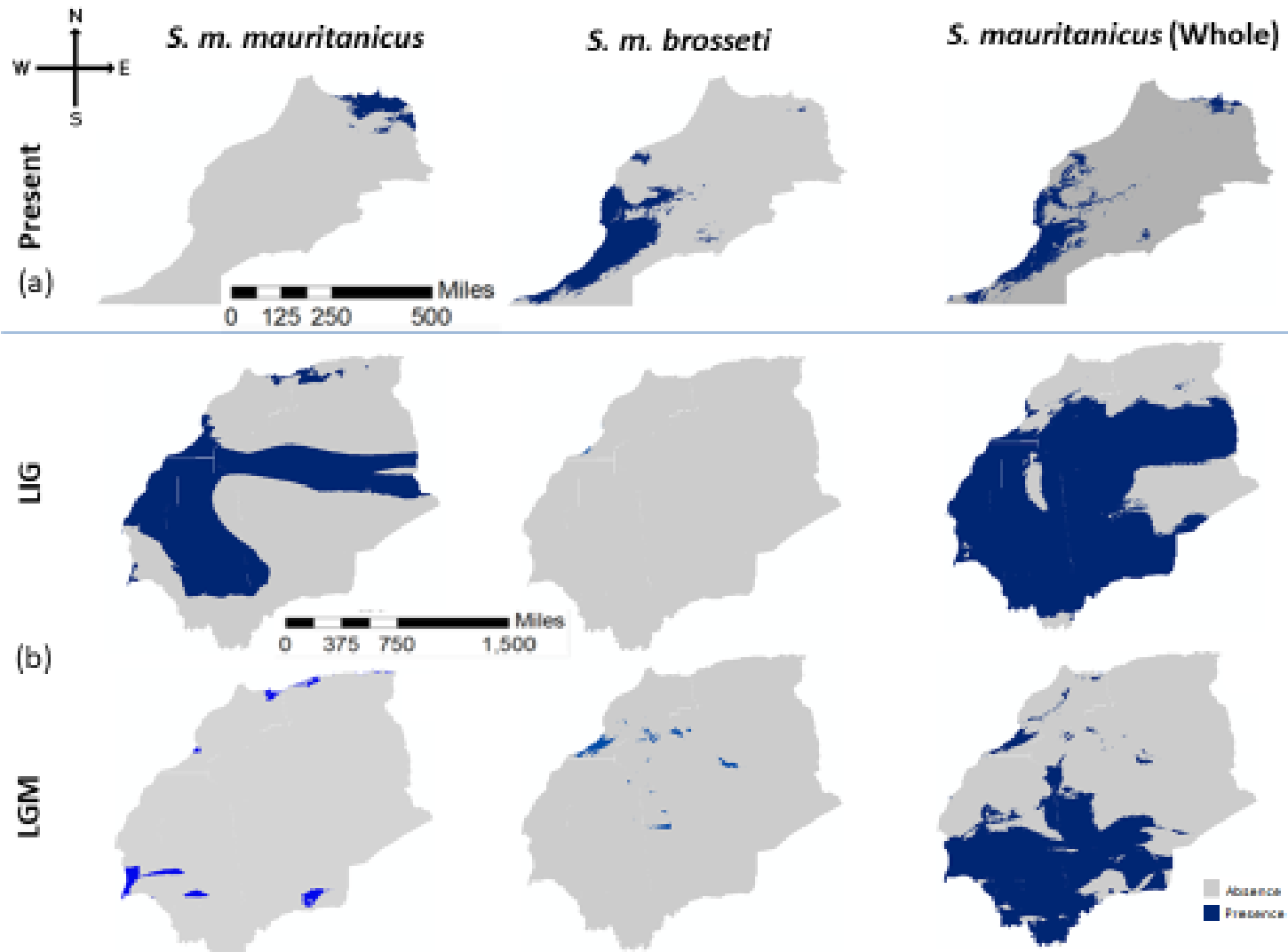
Figure 4.34: Predicted suitable habitat for *S. mauritanicus* and subspecies under present (a), and LIG and LGM conditions (b).
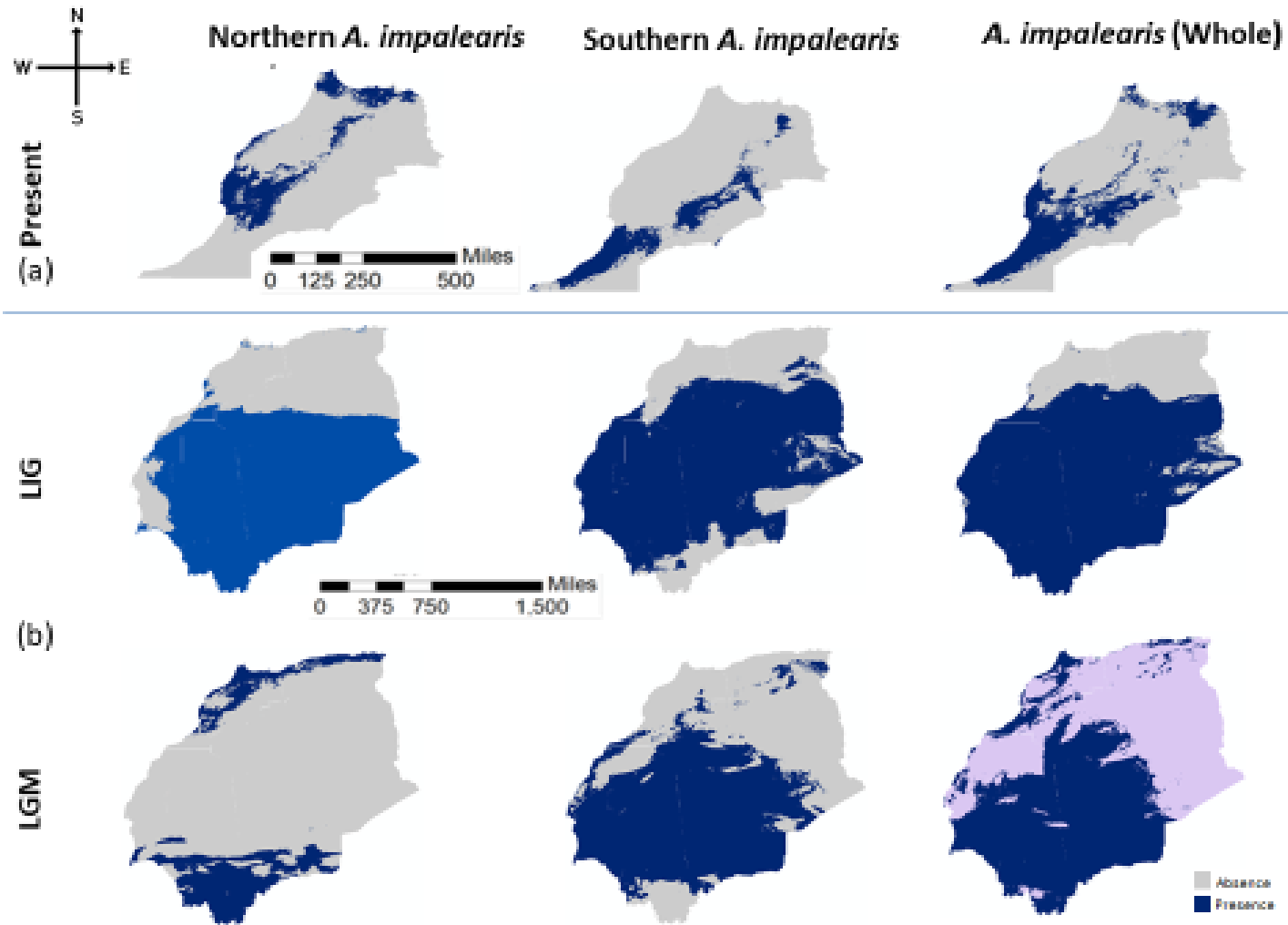
Figure 4.35: Predicted suitable habitat for A. *impalearis* and subspecies under present (a), and LIG and LGM conditions (b).

Current climate predictions

When the niches of sister subspecies were compared together for each species little overlap was calculated (Schoener's D), (See Table 5). Current observation maps support this with no overlap observed between species and respective subspecies. Some subspecies showing distinct geographical differences in suitable habitat (Figures 1.0-1.5), though continuous patterns are observed for most species (*T. weigmanni, C. mionecton, A. impalearis and M. olivieri*). PCA work also supports differences between the niches of sister subspecies, highlighted through separations within the component space (see figures 4 – 4.2).

Under current climatic conditions small overlap of suitable habitat was predicted for both *M. olivieri* and *S. mauritanicus,* with Schoeners D values of (D= 0.4 and 0.2 respectively, although values are still considered to represent low overlap). The current observational maps show a continuous pattern for *M. olivieri* and isolated populations for *S. mauritanicus*. Overlap between the sister subspecies and respective subspecies occurred in the component space of the PCA for *S. maurtianicus*, highlighting niche similarities.

C. mionecton, T. weigmanni and *A. impalearis* had predicted continuous distributions of suitable habitat under current climate conditions, with Schoener's D values of (D = 0.28, 0.17 and 0.3 respectively). The current observation maps also show a continuous pattern between sister subspecies, supporting MaxEnt outputs. The PCA results for these species and respective subspecies highlight differences in the component space portraying differences in niche/variable importance.

C. ocellatus had no overlap or continuous pattern of suitable habitat predicted under current climate conditions, with the two sister subspecies populations' isolated from one another (Schoeners D value for this species was the lowest D = 0.09). The observational data for this species and respective subspecies supports the prediction, as currently there is no overlap of populations or continuous pattern observed. PCA work also supports the lack of population/ niche overlap observed, as little overlap predicted in the component space, highlighting differences in variable importance.

C. mionecton showed a continuous pattern of suitable habitat when subspecies were modelled separately and as a whole, this predicted suitable habitat supports findings from the PCA, which show overlap in the component space (Figure 4), with one of the higher values from the niche overlap calculations (D= 0.27).

When species were modelled as a whole (subspecies together) under current climatic conditions, the suitable habitat predicted showed disjunct distributions for *C. ocellatus*, *T. weigmanni*, *M. olivieri*, *S. mauritanicus* and *A. impalearis*, with areas of predicted habitat in similar locations to those predicted by the models when subspecies were modelled individually.

LIG predictions

Under LIG climate conditions, the suitable habitat predicted for *C. ocellatus, T.weigmanni, S. mauritanicus* and *A.impalearis* and respective subspecies showed similar scenarios, with predicted overlap between areas of suitable habitat. Isolated populations were also predicted in areas that coincide with the geographical area that particular sister subspecies currently inhabits (*C. m. mionecton, C. m. trifasciatus, C. o. subtypicus, T. w. elegans, S. m. mauritanicus and Northern A. impalearis)*. PCA results support the predicted isolations, shown by the separations in component space between subspecies. No overlap of suitable habitat was predicted for *C. mionecton* (D = 0.28) and respective subspecies under LIG conditions, with both subspecies completely isolated from one another in areas geographically similar to the current observed populations of the individual subspecies. The predicted suitable habitat for *M. olivieri* under LIG conditions showed no isolated populations (D = 0.4), with one sister subspecies' predicted habitat overlapping the other (Figure 4.32). Supported by the PCA, which shows some overlap in the component space (Figure 4.14).

When sister subspecies were modelled together to create a surrogate ancestral species the predicted habitats under LIG conditions showed either minor or major fragmentation for all species. Isolated habitats were predicted for *C. mionecton, C. ocellatus, A. impalearis* and *M. olivieri* in areas geographically similar to areas a sister subspecies currently inhabits *(C. m. mionecton, C. o. subtypicus, Northern A. impalearis* and *M. o. simoni*).

LGM predictions

The predicted suitable habitats under LGM conditions show overlap and fragmentation for all species and their respective subspecies, although the size of the predicted overlap of habitat differs for each species. Species' that only had small areas of overlap predicted where *C. mionecton*, *C. ocellatus*, *M. olivieri* and *S.mauritanicus*. Isolations of predicted suitable habitat during the LGM predicted within these species (*C. m. mionecton, C. o. subtypicus, M .o. simoni, S. m. mauritanicus* and *S. m. brosseti)*. These areas of isolated suitable habitat predicted for the subspecies listed, coincide with areas that the subspecies currently inhabit, largely in North Africa/Morocco. Under LGM conditions, *T. weigmanni* and *A. impalearis* have larger areas of niche overlap predicted between subspecies, though isolated habitats also predicted between subspecies for Northern *A. impalearis*. During the LGM *T.*

weigmanni has large areas of suitable habitat overlap, though areas of no overlap between subspecies are also predicted. Both the subspecies of *T. weigmanni* predicted a north south division, so although overlap is predicted between populations the geographical distance between the populations could be significant, as gene flow would not occur between the northern and southern overlapping populations.

When sister subspecies were modelled together to create a surrogate ancestral species, the predicted habitats under LGM showed disjunct/fragmented distributions for all species (*C. mionecton, C. ocellatus, A. impalearis, S. mauritanicus, T. weigmanni* and *M. olivieri*). Isolated habitats also predicted for *C. mionecton*, *C. ocellatus*, *M. olivieri*, *T. weigmanni* and *S. mauritanicus,* again these predicted isolations are in areas geographically similar to areas a sister subspecies currently inhabits, largely in North Africa/Morocco *(C. m. mionecton, C. m. trifasciatus, C. o. subtypicus, M. o. simoni* and *S. m. brosseti).* The geographical split of suitable habitat for T. weigmanni was a large north-south division under LGM conditions.

All species had predicted habitat isolations for at least one subspecies in at least one climate model/scenario, with the whole species models also supporting isolated suitable habitat. Habitat isolation predicted for *C. o. subtypicus* under both climate models (LIG & LGM), with the whole species models supporting a LIG isolation. Habitat isolation predicted for *T. w. elegans* under LIG, with the whole species model supporting isolations during the LGM. Habitat isolation predicted for *S. m. mauritanicus* under both climate models (LIG & LGM), with the whole species model supporting isolations during the LGM. Habitat isolation predicted for *C. m. mionecton* under both climate models (LIG & LGM), with whole species model supporting isolations under both climate models. Habitat isolations predicted for Northern *A. impalearis* under both climate models (LIG & LGM), with whole species model supporting isolations in both climate models. Finally, habitat isolation predicted for *M. o. simoni* under LGM conditions (LIG & LGM), with whole species model supporting isolation under both climate scenarios. Below is a table to portray this information in a concise manner.

Table 10. Overview of the MaxEnt predictions of suitable habitat.

Subspecies modelled individually				
Species	**Schoener's D**	**Current observed distribution**	**LGM predictions**	**LIG predictions**
C. ocellatus	0.09	Isolation between subspecies	Areas of **isolation** between subspecies. **Fragmentation** within subspecies.	Areas of **isolation** between subspecies, overlap also between subspecies.
T. weigmanni	0.17	Continous pattern between subspecies	**Overlap** between subspecies with **fragmentation** within each subspecies.	Overlap between subspecies, with small areas of **isolation** between subspecies.
S. mauritanicus	0.2	Isolation between subspecies	**Isolation** between subspecies with **fragmentation** within each subspecies.	Areas of overlap between subspecies, with area **isolation** between subspecies.
C. mionecton	0.28	Continous pattern between subspecies	Areas of **isolation** between subspecies, with overlap between subspecies.	**Isolated areas** between subspecies.
A. impalearis	0.29	Continous pattern between subspecies	Areas of **isolation** between subspecies, with overlap between subspecies.	Areas of **isolation** between subspecies, with overlap between subspecies.
M. olivieri	0.4	Continous pattern between subspecies	**Isolation** between subspecies with **fragmentation** within each subspecies.	Large areas of overlap between subspecies with small areas of **isolation**.
Subspecies modelled together (ancestral species)				
Species	**Schoener's D**	**Current observed distribution**	**LGM predictions**	**LIG predictions**
C. ocellatus	0.09	Isolation between subspecies	**Framented** distribution, small areas of **isolation**.	**Fragmented** distribution, **isolated** in parts.
T. weigmanni	0.17	Continous pattern between subspecies	**Fragmented** distribution, large areas of **isolation**.	**Continous** distribution with small areas of **isolation**.
S. mauritanicus	0.2	Isolation between subspecies	**Fragmented** distribution, small areas of **isolation**.	**Continous** distribution with small areas of **isolation**.
C. mionecton	0.28	Continous pattern between subspecies	**Fragmented** distribution, areas of **isolation**.	**Fragmented** distribution, small areas of **isolation**.
A. impalearis	0.29	Continous pattern between subspecies	**Fragmented** distribution, areas of **isolation**.	Large **continous** distribution with small **isolated** areas.
M. olivieri	0.4	Continous pattern between subspecies	**Fragmented** distribution, areas of **isolation**.	**Fragmented** distribution, areas of **isolation**.

Discussion

All of the reptile taxa discussed in this study either show intraspecific divergence or are recognized as recently-diverged sister species (Brown et al., 2002; Brown et al., 2012; Carretero et al., 2005; Harris & Rato, 2008; Kapli et al., 2015; Kornilios et al., 2010; Mattiucci et al., 2001). Currently very few modelling studies have addressed the cause of the speciation events within Morocco, with some studies suggesting that speciation events could be climatically mediated through range shifts (Brown et al., 2012; DeMenocal, 2004), whilst other papers suggest vicariance due to changes in physical barriers such as mountains (Brown et al., 2002). The rationale for modelling species distributions over past glacial climates (LGM & LIG) was to identify opportunities/climatic events that help explain the cause of divergence that led to speciation. For example, if isolations/fragmentations due to climatic oscillations are predicted for studied species, then this may explain speciation events for other species within Morocco/NW Africa.

The distribution models created for this study show evidence of isolations/fragmentations in every case, under either both of the past climate scenarios or at least one of them (LIG or LGM). Under LGM conditions, geographically isolated areas of suitable habitat were predicted between separately modelled subspecies for all study species, except *T. w. weigmanni* and *T. w. elegans*. Geographically isolated populations were also predicted under LGM conditions when the whole species was modelled (*C. mionecton, T. weigmanni, S. mauritanicus. M. olivieri and A. impalearis*). This indicates that the climate (precipitation/temperature levels) associated with the LGM (and earlier glacial maxima) had a stronger impact on a species range when compared to precipitation/temperature levels during the LIG (and earlier interglacial). Other species distribution studies have also suggested climate associated with the LGM can cause divergence within a species (Barnes et al., 2002; Leonard et al., 2000; Rossetto et al., 2012; Walstrom et al., 2012). Nevertheless, the LIG models produced in this study also predicted geographically isolated areas of suitable habitat, irrespective of whether the subspecies were modelled separately, (*C. mionecton, C. ocellatus, T. weigmanni, A. impalearis & S. mauritanicus*) or modelled together (*C. mionecton, C. ocellatus, A. impalearis & M. olivieri*).

LGM refugia for most species were found in northern Africa (North of study area), which was the area of highest rainfall during that period (Beghin et al, 2015). The higher precipitation in this area was due to the large-scale effect of the southward shift of the North Atlantic jet stream during the LGM (Beghin et al., 2015). Therefore, during a time when precipitation is considered to have generally decreased (Otto-Bliesner et al., 2006), Morocco/north Africa may have provided refugia, supported by the higher precipitation in that area. Subspecies *C. m. mionecton, C. o. subtypicus, T. w. elegans*, *M. o. olivieri* and northern *A. impalearis* all preferred higher precipitation when compared to their sister subspecies, corresponding to isolated refugia in northern Africa during the LGM. Hence, rainfall mediated allopatry or range fragmentations could have been a driver for divergence in these species. The subspecies S*. m. mauritanicus* similarly had refugia predicted in northern Africa; however, multivariate analysis did not detect a preference to higher precipitation when compared to its sister subspecies.

Precipitation variability would have likely had a major influence on the ecological landscape in terms of available food, vegetation, cover or changes that effect prey (direct or indirect) (Beghin et al., 2015; Gasse & Campo, 1994; Griffin, 2002; Otto-Bliesner et al., 2006). Cores samples analysed by DeMenocal (2004) showed an increase in dust corresponded to changes in pollen, showing that an increase in aridity has a direct relationship to changes in vegetation, this therefore supports the idea that changes to habitat would correspond to changes in precipitation. Other studies have shown rainfall to have a major role in a lizard's daily behaviour, and suggested that it can enhance survival, growth and possibly clutch size and hatching success (Dickman et al., 1999). Ryan et al. (2016) suggests that rainfall is the biggest contributing factor to changes in a lizard's microhabitat, not temperature, due to changes in cover/vegetation that arise from this. These changes in vegetation/cover have been linked to the success and demise of certain lizards within the same environment (Dickman et al. 1999). This study cannot determine the exact impact of rainfall and/or associated vegetation changes; however, paleoclimate studies coupled with phylogenetic studies suggest that it would have produced a habitat change large enough to fragment distributions (deMenocal, 2004; Kornilios et al., 2010). This is based off estimated divergence times corresponding with recorded changes in precipitation (deMenocal, 2004; Mouline et al., 2008). Meaning, times associated with increased aridity, a species may have contracted into potentially isolated refugia. This may help explain the differences found in the niche comparison methods between subspecies used in this study, because of the evolutionary independence of the reduced/isolated populations.

Temperature variables are the second type of climatic variable available in the Bioclim data. It may be expected that lizards would also be effected by temperature, particularly as reptiles are ectotherms and very sensitive to this (Ihlow et al., 2012; Kapli et al., 2014; Shine et al., 2002; Spellerberg, 1972). Female reptiles of many taxa select nest sites based on physical cues (thermal) that relate to incubation success (Bragg et al., 2000). Increases in temperature have also been shown to correlate to increases in female body size, and, as fecundity is strongly reliant on female body size, clutch size and total reproductive output increase (Chamaille-james et al., 2005).

Niche evaluation work in this study showed sister subspecies to indicate differences in temperature tolerance between one another. However in this study precipitation seemed to be the significant driver, based on the areas on isolation corresponding to areas of higher precipitation. Other paleoclimate studies suggest that the impact of changes in temperature are indirectly and largely due to associated changes in precipitation (Araújo et al., 2006). Changes in temperature are likely to cause changes to a species range, whether it is directly related to the temperature change, or the associated impact of changes in precipitation. If populations become isolated during times in refugia then the genetic diversity is reduced in the separate populations. If gene flow is dramatically reduced or ceases with another population, each isolate follows independent evolutionary trajectory, which may potentially lead to speciation (Hewitt, 2000).

Deglaciation must have followed the LGM and other glacial maxima as generally temperatures and precipitation began to increase (LIG/earlier interglacial). Increased temperatures and precipitation could provide opportunities for species to expand from refugia and disperse (potential for secondary contact), and for population sizes to increase (Chamaille-james et al., 2005), due to potential increases in suitable vegetation/habitat. Not every species in this study had predicted range expansion that could lead to secondary contact, with certain species predicting range contraction during the LIG (*C .mionecton* and, *M. olivieri*). The idea of range expansion following deglaciation is supported in other studies (Kutch & Tan., 2005; Lee-Yaw et al., 2008; Zamudio & Savage., 2003), and is one mechanism that can cause secondary contact.

The current observed distributions for four of the species (Bons & Geniez, 1996) show a continuous distribution across two divergent subspecies, yet currently (at least in the cases of *C. mionecton* and *A. impalearis:* (Brown et al., 2002; Brown et al., 2012) there is limited gene flow. This suggests secondary contact. Secondary contact occurs when two species/subspecies that originated during divergence remeet due to changes to a natural barrier or climatic changes

increasing a species potential range. Between the glacial extremes (LGM & LIG), patterns of range contraction and expansion were predicted in this study. This pattern could have potentially repeated itself numerous times over several million years and could explain how continuous patterns with no genetic introgression are currently observed. The duration and the frequency of the isolations determine the accumulation of genetic divergence (Hewitt, 2000): The longer the isolation, the greater the genetic distance between populations, leading to fewer genes exchanged during subsequent secondary contact. Therefore, after many glacial cycles (with repeated isolations), the two populations would have diverged to a point where they cannot mate successfully and so are regarded as two separate species.

The divergence times for species in this study range between 1-15Ma (Brown et al., 2002; Brown et al., 2012; Kapli et al., 2015; Harris & Rato, 2008; Kornilios et al., 2010). For speciation to take place, the divergent populations/genomes must have remained largely separate through many range changes. The Croll-Milankovitch theory proposes that the Earth's orbit around the Sun is the "pacemaker" of the ice ages (Hewitt, 2000), which states the main orbital eccentricity has a 100-Ka cycle. Therefore, in 15 Ma there would have been roughly 150 glacial cycles brought on by orbital oscillation (Hewitt, 2004). A subspecies such as *S. m. brosseti,* that has an estimated divergence time of 15 Ma from *S. m. mauritanicus* (Rato & Harris, 2008), may have been subjected to repeated fragmentation/isolation events prior to 15 Ma that were climatically mediated, with roughly 150 glacial cycles since 15 Ma that will have occurred. These 150 glacial cycles would create plenty of potential isolation events followed by secondary contact scenarios, until eventually, the separate isolated populations have diverged to a point where they cannot successfully share genes, as they are unable to mate successfully, due to potential habitual or morphological differences. This leads to the pattern that is currently observed by some of the species in this study (*A. impalearis, C. mionecton*) and possibly (*T. weigmanni, M. olivieri*) where the current climate enables a continuous pattern between subspecies to occur (secondary contact scenario) yet gene flow does not. (Genetic work to determine genetic introgression only available for *A. imaplearis* and *C. mionecton*).

The genetic diversity within populations can change drastically following an isolation event. Before the isolation event, a species may have had a continuous distribution that spanned a large geographical range. Within the species range (prior to fragmentation/isolation), there could be varied selection pressures brought on by differences in the environment within the extremes of the habitat. Meaning, a population (prior to isolation) may have large genetic diversity. A reduction in population size caused by fragmentation/isolation will reduce the genetic diversity. This reduction in population size will increase genetic drift within populations

leading to greater divergence between populations. Therefore, the predicted isolations will have caused reduced populations, attributed to the changes in climate. These reduced populations would therefore have reduced genetic diversity leading to divergence through genetic drift or differing adaptive pressures between the isolated populations. For example, consider genes that relate to predator avoidance. If the range of predation experienced in the larger continuous population are varied, then the genetic diversity in relation to predator avoidance is likely to be large. Therefore, if a population is split then the genetic diversity in relation to predator avoidance may have reduced, due to the smaller population and hence less individuals to share genes with. If the genetic diversity is reduced then divergence will occur quicker.

The observed distributions of all the species in this study correspond well to the niche overlap detected by this study and the suitable habitat predicted by MaxEnt. The observed distributions for *T. wiegmanni, C. mionecton, A. impalearis and M. olivieri* (Bons & Geniez, 1996) all portray a continuous pattern between subspecies, an indication that gene flow could have been possible whenever climate conditions were similar to those at present. When modelling subspecies separately (for each species) the MaxEnt models support the current distribution maps (Bons & Geniez, 1996), with continuous patterns predicted for *T. weigmanni, C. mionecton, A. impalearis* and *M. olivieri* under current climatic data with no overlap, thus supporting the hypothesis of potential gene flow under certain climates. The above species were found to have low niche overlap (Rödder & Engler, 2011).

In contrast, the current distribution maps for *C. ocellatus* and *S. mauritanicus* did not show a continuous distribution (Bons & Geniez, 1996). They were found to have very low niche overlap scores indicating little or no overlap (Rödder & Engler, 2011). *C. ocellatus* and *S. mauritanicus* were not predicted to be continuous under current climatic data, corresponding well with current observed distributions (Bons & Geniez, 1996) and the low niche overlap values for both these species. It is important to note that the degree of niche overlap does not seem to be associated with the tendency for a continuous distribution or not.

When species were modelled as a whole (subspecies together), fragmented distributions were predicted for all species except *C. mionecton* under current climatic conditions. This contradicts the majority of the observed distributions (*C. ocellatus and S. mauritanicus* excluded), and the models that treated subspecies independently. The reason the whole species models may contradict with the subspecies models could be that when the two subspecies are put together as a pseudo ancestral species, the niches of two sister subspecies create a broad/varied niche and so encompasses a larger suitable habitat with varying ecological constraints. An example of

this could be M*. olivieri* as the two sister subspecies inhabit different geographical/potentially ecological areas*; M. o. olivieri* is observed inland within Morocco and M*. o. simoni* observed along the coast of Morocco. Evidence of conflicting niches can be seen in the jackknife, PCA and niche overlap results between the two model types (whole species/subspecies), (See results). The jackknife results indicated a particular subspecies to have a strong weighting within the whole species models, this was suggested by a certain subspecies jackknife result corresponding to the whole species jackknife result.

Model Evaluation

Model validation methods performed in this study support good model performance for MaxEnt. The null hypothesis results showed how the MaxEnt model out performed models where the data are random. The majority of AUC scores produced in MaxEnt are of high significance. The sensitivity analysis showed good model performance from MaxEnt in both whole species and subspecies models, and the niche overlap (Schoeners D) calculations support the findings from the PCA which highlighted differences between sister subspecies.

The jackknife results are noteworthy. It seems that the variables deemed most important in the whole species models, were the same for some of the subspecies models, this indicates that one subspecies may have higher weighting within the whole species (ancestral) model. However, the reason for this higher weighting/dominance of a certain subspecies is unclear, as the sample size does not seem to be the contributing factor so there does not seem to be a definitive correlation between sample size and dominance of a subspecies. There also does not seem to be a definitive correlation between AUC scores and sample sizes which has been highlighted as a potential problem in other papers (Wisz et al., 2008).

Limitations and future work

In this study, we have predicted the past distribution of the species based on modelling of the current niche. We have used past distributions to identify areas of isolation which may have led to divergence events (opportunities for divergence). Therefore, if we assume that the past distributions are an accurate reflection of the niche during the time when speciation took place we must therefore assume niche conservation between the niche of the species now and the niche at the time of the divergence event. If we assume niche conservation, we must also assume that the divergence was due to evolutionary/genetic changes on characteristics other than those which influence the climatic niche of the species (Wiens et al., 2010). Under this

assumption we can identify past isolations through modelling, but this study can tell us nothing about the nature of the divergence which may have occurred.

We have also made the assumption that conditions within the LIG and LGM are representative of other interglacial and glacial maximum conditions which occurred earlier. This is under the assumption of the Croll-Milankovitch theory that suggests the regular variations in the Earth's orbit (Eccentricity, Axial tilt and Precession) around the Sun are the pacers of the ice age cycles. The variations in the Earth's orbit adjust the insolation of the Earth and the energy it receives leading to significant climate changes, this theory has been supported in other studies (Hays et al., 1976; Imbrie & Imbrie, 1980; Kukla & Gavin, 2004; Rossignol, 1983). This is a limitation of the study as the climatological data this assumption is based off are estimates, therefore lacking certainty. By assuming Earths insolation (therefore climate) is mediated by the Earths circumnavigation around the Sun, we can infer that any patterns of isolation or fragmentation predicted under the LGM and LIG climates were repeated numerous times.

The choice of predictor variables was a limitation of this study as Bioclim variables were the only environmental variables used due to the lack of available past data for other variables; this means that results produced in MaxEnt can only be interpreted in terms of temperature and precipitation. This is a limitation of the study as important environmental influences on a species distribution (such as biotic factors) have not been incorporated into the model and therefore have not been considered.

Using other variables to calculate the species niche may be beneficial, for example biotic variables such as observational data for a predator. Currently few studies have attempted to incorporate biotic interactions when hindcasting (Li et al., 2018; Silva et al., 2015; Troy et al., 2007), with more studies incorporating biotic interactions when forecasting (Dahdouh-Guebas et al., 2004; Davis et al., 2012; Hof et al., 2012 ; Keenan et al., 2011; Le Roux et al., 2014; Potter, 2004). I hope that the progression of the use of SDMs will increase the amount of studies aiming to create estimated biotic datasets based on the past. These past biotic dataset can then be applied to SDMs.

Some studies have attempted to predict future patterns of vegetation (Dahdouh-Guebas et al., 2004; Davis et al., 2012; Keenan et al., 2011; Potter, 2004) to help improve model performance when forecasting. The predicted future vegetation pattern equips researchers with more environmental interactions on which to project a species niche. Some paleoclimate studies have also estimated past vegetation patterns (Li et al., 2018; Troy et al., 2007), this could help strengthen predictive studies that hindcast, as extra environmental interactions can be

quantified within the model. Vegetation has been proven to correspond with changes in precipitation (deMenocal, 2004; Griffin, 2002), as precipitation was identified as a main driver in the predicted fragmentations within this study then vegetation data would help validate what sort of impact the changes in precipitation had on the geographical landscape. Incorporation of other biotic variables such as predation when hindcasting could also improve the predictions. Incorporating biotic variables may improve the models ability to calculate the species niche. Depending on what variables are added a *realised* niche may be estimate rather than a *potential*.

Conclusion

This study shows some evidence that ranges of North African reptiles are likely to have changed dramatically over the most recent glacial cycles (LIG & LGM) associated with changes in temperature and precipitation between these two glacial extremes. The divergence times for species discussed in this study were previously shown to be between 1-15 Ma, which predates any reliable climate data, and so the LGM and LIG act as surrogate climates for earlier interglacial and glacial maxima periods. Based on the assumption that glacial cycles have a similar cyclic pattern of occurrence (Croll-Milankovitch theory), we can infer that many previous climatic cycles would lead to fragmentation/isolation events. These fragmentation/isolation events will have occurred within a species between each glacial cycle, creating numerous opportunities over time for isolated populations to diverge from one another due to evolutionary independence, which over many glacial cycles could result in speciation.

Appendix

Appendix A: Table of Schoener's D values for niche overlap on *Chalcides mionecton.*

SPECIES	C.m.trifasciatus 9	C.m.mionecton 0	C.m.mionecton 1	C.m.mionecton 2	C.m.mionecton 3	C.m.mionecton 4	C.m.mionecton 5	C.m.mionecton 6	C.m.mionecton 7	C.m.mionecton 8	C.m.mionecton 9	C.m.trifasciatus 0	C.m.trifasciatus 1	C.m.trifasciatus 2	C.m.trifasciatus 3	C.m.trifasciatus 4	C.m.trifasciatus 5	C.m.trifasciatus 6	C.m.trifasciatus 7	C.m.trifasciatus 8
C.m.trifasciatus 9	1	0.273684375	0.266612791	0.272865444	0.26448066	0.269958032	0.269612741	0.271595012	0.253791017	0.271601258	0.246800409	0.903230327	0.924675936	0.928852035	0.930223112	0.940280566	0.944310017	0.924464019	0.902027999	0.911233959
C.m.mionecton 0	+	1	0.965170086	0.951753911	0.948971002	0.936740791	0.950230128	0.950324169	0.923188513	0.92992537	0.918488328	0.288402268	0.286253555	0.285197685	0.281981901	0.277334058	0.260436328	0.282878675	0.245307524	0.28187035
C.m.mionecton 1	+	x	1	0.956201785	0.950981756	0.93986391	0.954989697	0.957767384	0.927804184	0.929201246	0.923548589	0.279570022	0.278195048	0.276882481	0.273966475	0.270401009	0.251704279	0.275057766	0.236794722	0.273021478
C.m.mionecton 2	+	x	x	1	0.936452951	0.949122325	0.973517993	0.969080958	0.914481182	0.953873351	0.910092779	0.287656258	0.287001469	0.284714281	0.281346853	0.277346013	0.260547123	0.282639013	0.245806137	0.281596438
C.m.mionecton 3	+	x	x	x	1	0.935676192	0.942858653	0.934075249	0.936883652	0.925589066	0.9144813	0.277397962	0.273451439	0.274293458	0.27147732	0.267332626	0.249721195	0.271857281	0.234309877	0.270483086
C.m.mionecton 4	+	x	x	x	x	1	0.953750843	0.950299814	0.927160302	0.945268919	0.914625707	0.280726493	0.283976939	0.279151108	0.275809989	0.275928207	0.258889498	0.278350329	0.238292082	0.276022372
C.m.mionecton 5	+	x	x	x	x	x	1	0.964487696	0.910590942	0.953963	0.904359243	0.284368575	0.284310404	0.282124723	0.278718622	0.274359304	0.258236016	0.279742446	0.243340382	0.27940781
C.m.mionecton 6	+	x	x	x	x	x	x	1	0.915208758	0.952971016	0.910615497	0.285743332	0.2863915	0.282953995	0.279803333	0.276081303	0.259839218	0.28178026	0.24382306	0.280058473
C.m.mionecton 7	+	x	x	x	x	x	x	x	1	0.900831283	0.924208243	0.261954926	0.257697324	0.258095246	0.255167076	0.257141422	0.237770365	0.257469993	0.217262893	0.254502784
C.m.mionecton 8	+	x	x	x	x	x	x	x	x	1	0.905418302	0.286521175	0.289708786	0.284198078	0.2807756	0.277724408	0.262856778	0.283509477	0.244181461	0.281711215
C.m.mionecton 9	+	x	x	x	x	x	x	x	x	x	1	0.255250997	0.258938989	0.254797342	0.251886251	0.252903145	0.233700867	0.254338855	0.214304102	0.250191633
C.m.trifasciatus 0	+	x	x	x	x	x	x	x	x	x	x	1	0.89837474	0.900541785	0.891580456	0.907504692	0.899186193	0.890687131	0.880047467	0.897830743
C.m.trifasciatus 1	+	x	x	x	x	x	x	x	x	x	x	x	1	0.950542104	0.947144402	0.930055559	0.920077525	0.947946849	0.905797723	0.914909543
C.m.trifasciatus 2	+	x	x	x	x	x	x	x	x	x	x	x	x	1	0.980048611	0.942356596	0.915625826	0.967535887	0.916611991	0.930937873
C.m.trifasciatus 3	+	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.939169819	0.913424825	0.96972836	0.912724959	0.927387801
C.m.trifasciatus 4	+	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.923475677	0.939441929	0.898918876	0.936064217
C.m.trifasciatus 5	+	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.910803345	0.918544953	0.912895206
C.m.trifasciatus 6	+	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.90718669	0.923229135
C.m.trifasciatus 7	+	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.911354898
C.m.trifasciatus 8	+	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1

Appendix B: Table of Schoener's D values for niche overlap on *Chalcides ocellatus.*

SPECIES	Subsp.subtyp	Ocellatus_	Ocellatus_	Ocellatus_	Ocellatus_	Ocellatus_	Ocellatus_	Ocellatus_	Ocellatus_	Ocellatus_	Ocellatus_	Subsp.sub	Subsp.sub	Subsp.sub	Subsp.sub	Subsp.sub	Subsp.sub	Subsp.sub	Subsp.sub	Subsp.sub
Subsp.sub	1	0.085847	0.070849	0.077531	0.077117	0.078479	0.082029	0.07971	0.081702	0.076038	0.075035	0.954404	0.943201	0.959649	0.94811	0.950408	0.936616	0.955451	0.967968	0.953336
Ocellatus_	x	1	0.92794	0.934316	0.920776	0.929763	0.930114	0.920587	0.932618	0.931097	0.928817	0.084489	0.087069	0.085103	0.08459	0.088274	0.080676	0.089714	0.08595	0.091965
Ocellatus_	x	x	1	0.91383	0.900761	0.921046	0.896693	0.905827	0.926787	0.914478	0.922426	0.068947	0.072152	0.070334	0.068771	0.073161	0.066938	0.074192	0.071315	0.07723
Ocellatus_	x	x	x	1	0.938972	0.945436	0.933841	0.928738	0.944946	0.949704	0.951648	0.0762	0.078695	0.076175	0.075628	0.079116	0.072001	0.080315	0.077379	0.083389
Ocellatus_	x	x	x	x	1	0.938062	0.930451	0.927065	0.937083	0.939977	0.93136	0.076063	0.078343	0.075885	0.075482	0.078925	0.072095	0.080061	0.076989	0.083103
Ocellatus_	x	x	x	x	x	1	0.915214	0.931047	0.957646	0.958074	0.944284	0.076651	0.079888	0.07756	0.076483	0.080337	0.073825	0.081529	0.078785	0.084721
Ocellatus_	x	x	x	x	x	x	1	0.918063	0.918428	0.918593	0.921218	0.080918	0.083148	0.080556	0.080287	0.08369	0.076832	0.084699	0.081763	0.087742
Ocellatus_	x	x	x	x	x	x	x	1	0.932411	0.941891	0.9354	0.07743	0.080986	0.078754	0.076989	0.081049	0.074899	0.082427	0.079987	0.085982
Ocellatus_	x	x	x	x	x	x	x	x	1	0.959834	0.942742	0.08016	0.083015	0.080792	0.079798	0.083527	0.076513	0.085001	0.081793	0.088041
Ocellatus_	x	x	x	x	x	x	x	x	x	1	0.954661	0.073928	0.077453	0.075188	0.073906	0.077775	0.071468	0.07904	0.076358	0.082427
Ocellatus_	x	x	x	x	x	x	x	x	x	x	1	0.073146	0.07604	0.073906	0.072011	0.076133	0.069869	0.077474	0.075099	0.081329
Subsp.sub	x	x	x	x	x	x	x	x	x	x	x	1	0.942174	0.934522	0.934152	0.932147	0.920803	0.938902	0.948189	0.955912
Subsp.sub	x	x	x	x	x	x	x	x	x	x	x	x	1	0.934479	0.930143	0.926633	0.934639	0.932904	0.954248	0.945103
Subsp.sub	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.945751	0.949656	0.937653	0.966724	0.949288	0.936992
Subsp.sub	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.946515	0.933053	0.943827	0.939968	0.928812
Subsp.sub	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.951156	0.957722	0.939812	0.938223
Subsp.sub	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.934402	0.934955	0.92073
Subsp.sub	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.944259	0.940967
Subsp.sub	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.955617
Subsp.sub	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1

Appendix C: Table of Schoener's D values for niche overlap on *Messalina olivieri.*

SPECIES	Olivieri.ol	Olivieri si	Olivieri si	Olivieri.si	Olivieri.si	Olivieri.si	Olivieri.si	Olivieri.si	Olivieri.si	Olivieri.si	Olivieri.si	Olivieri.ol	Olivieri.ol	Olivieri.ol	Olivieri.ol	Olivieri.ol	Olivieri.ol	Olivieri.ol	Olivieri.ol	Olivieri.ol
Olivieri.ol	1	0.40323	0.404668	0.399985	0.402779	0.39881	0.406869	0.408878	0.39698	0.403356	0.406489	0.913827	0.929267	0.929364	0.928071	0.924459	0.919607	0.926344	0.927145	0.903383
Olivieri.si	x	1	0.897323	0.923291	0.912186	0.91562	0.897628	0.921996	0.898259	0.910415	0.908835	0.39134	0.398217	0.409336	0.406617	0.407747	0.389088	0.399156	0.390124	0.402653
Olivieri.si	x	x	1	0.910698	0.943809	0.902345	0.921792	0.926272	0.918454	0.936272	0.93502	0.390304	0.399237	0.410137	0.407082	0.407961	0.38885	0.400423	0.391671	0.397356
Olivieri.si	x	x	x	1	0.924452	0.937215	0.911616	0.926323	0.907755	0.934514	0.932986	0.388842	0.394475	0.405388	0.403198	0.404054	0.386139	0.397505	0.388865	0.394537
Olivieri.si	x	x	x	x	1	0.918133	0.926541	0.939783	0.916146	0.950216	0.951243	0.388372	0.398118	0.40851	0.404927	0.406188	0.385957	0.399692	0.390759	0.394849
Olivieri.si	x	x	x	x	x	1	0.901786	0.931748	0.897996	0.912546	0.931987	0.387507	0.395437	0.404029	0.402437	0.40314	0.384542	0.397226	0.388838	0.392254
Olivieri.si	x	x	x	x	x	x	1	0.922283	0.922856	0.943778	0.940298	0.397574	0.400965	0.410034	0.408351	0.410559	0.394495	0.402267	0.394608	0.401856
Olivieri.si	x	x	x	x	x	x	x	1	0.908193	0.924225	0.937685	0.397786	0.404126	0.416033	0.413151	0.414896	0.394088	0.407498	0.398726	0.405522
Olivieri.si	x	x	x	x	x	x	x	x	1	0.925402	0.923841	0.384741	0.391368	0.401189	0.39923	0.399548	0.384267	0.392067	0.382258	0.38841
Olivieri.si	x	x	x	x	x	x	x	x	x	1	0.946259	0.390699	0.397581	0.407325	0.405163	0.406118	0.388286	0.398983	0.390952	0.395692
Olivieri.si	x	x	x	x	x	x	x	x	x	x	1	0.394401	0.401713	0.411901	0.409186	0.411211	0.39159	0.404221	0.395244	0.399717
Olivieri.ol	x	x	x	x	x	x	x	x	x	x	x	1	0.914455	0.923362	0.924298	0.925133	0.935981	0.929379	0.936636	0.899163
Olivieri.ol	x	x	x	x	x	x	x	x	x	x	x	x	1	0.920877	0.930557	0.916861	0.918108	0.936916	0.927235	0.889151
Olivieri.ol	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.930468	0.926478	0.920008	0.933653	0.934749	0.89929
Olivieri.ol	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.932598	0.921853	0.933599	0.938801	0.902936
Olivieri.ol	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.921026	0.921749	0.928364	0.900287
Olivieri.ol	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.928832	0.922462	0.893161
Olivieri.ol	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.934709	0.896773
Olivieri.ol	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.898592
Olivieri.ol	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1

Appendix D: Table of Schoener's D values for niche overlap on *Trogonophis weigmanni.*

SPECIES	Trogonopl	subsp.wie	subsp.wie	subsp.wie	subsp.wie	subsp.wie	subsp.wie	subsp.wie	subsp.wie	subsp.wie	subsp.wie	Trogonopl	Trogonopl	Trogonopl	Trogonopl	Trogonopl	Trogonopl	Trogonopl	Trogonopl	Trogonopl
Trogonopl	1	0.147761	0.157844	0.13737	0.146266	0.154643	0.14906	0.15793	0.159944	0.154271	0.152846	0.938416	0.918325	0.921099	0.917767	0.923196	0.926768	0.934333	0.928597	0.921508
subsp.wie	x	1	0.931879	0.923893	0.919635	0.920848	0.916844	0.926128	0.903618	0.891806	0.950211	0.161777	0.16185	0.162207	0.156177	0.166072	0.148719	0.166643	0.16205	0.157466
subsp.wie	x	x	1	0.918633	0.917546	0.952247	0.940118	0.946647	0.920544	0.886708	0.935617	0.17198	0.171793	0.172357	0.163812	0.17498	0.158501	0.176327	0.171789	0.166271
subsp.wie	x	x	x	1	0.921716	0.920817	0.941394	0.929463	0.90331	0.884805	0.925163	0.149406	0.1508	0.150106	0.14384	0.151895	0.137001	0.15479	0.149908	0.143968
subsp.wie	x	x	x	x	1	0.924052	0.921066	0.93286	0.923496	0.919511	0.938724	0.161371	0.16091	0.160825	0.153016	0.162845	0.148209	0.164584	0.161282	0.155936
subsp.wie	x	x	x	x	x	1	0.94284	0.963678	0.926022	0.898885	0.940592	0.168596	0.168337	0.168603	0.160295	0.171018	0.155308	0.172521	0.168543	0.161937
subsp.wie	x	x	x	x	x	x	1	0.945526	0.91108	0.87735	0.927694	0.162613	0.163203	0.163015	0.15581	0.165473	0.149575	0.167421	0.162439	0.157379
subsp.wie	x	x	x	x	x	x	x	1	0.92719	0.89548	0.94828	0.17223	0.172106	0.172213	0.163769	0.174151	0.159118	0.176036	0.172187	0.165819
subsp.wie	x	x	x	x	x	x	x	x	1	0.934928	0.939358	0.174493	0.174213	0.174387	0.164778	0.176539	0.161248	0.177679	0.17442	0.16624
subsp.wie	x	x	x	x	x	x	x	x	x	1	0.919639	0.168607	0.168002	0.167903	0.159325	0.170184	0.155478	0.171526	0.168696	0.160265
subsp.wie	x	x	x	x	x	x	x	x	x	x	1	0.167278	0.166991	0.167286	0.159164	0.17004	0.153934	0.171359	0.167127	0.161427
Trogonopl	x	x	x	x	x	x	x	x	x	x	x	1	0.921936	0.9199	0.908232	0.933358	0.927895	0.930966	0.931612	0.932369
Trogonopl	x	x	x	x	x	x	x	x	x	x	x	x	1	0.950453	0.914134	0.937496	0.928382	0.95496	0.943676	0.924373
Trogonopl	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.910054	0.924316	0.93675	0.938567	0.93942	0.921697
Trogonopl	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.924612	0.915838	0.921009	0.920297	0.926737
Trogonopl	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.925142	0.939364	0.942542	0.926375
Trogonopl	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.925817	0.949186	0.923776
Trogonopl	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.939015	0.93314
Trogonopl	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.936761
Trogonopl	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1

Appendix E: Table of Schoener's D values for niche overlap on *Saurodactylus mauritanicus.*

SPECIES	S. Brosseti	S. Mauritanicus	S. Mauritanicus	S. Mauritanicus	S. Mauritanicus	S. Mauritanicus	S. Mauritanicus	S. Mauritanicus	S. Mauritanicus	S. Mauritanicus	S. Mauritanicus	S. Brosseti	S. Brosseti	S. Brosseti	S. Brosseti	S. Brosseti	S. Brosseti	S. Brosseti	S. Brosseti	S. Brosseti
S. Brosseti	1	0.16901429	0.167599	0.163191	0.165553	0.16977	0.171068	0.167117	0.16541	0.165494	0.168159	0.9387854	0.9170577	0.906109	0.8811513	0.9482896	0.915323	0.9265506	0.8709499	0.9304866
S. Mauritanicus	x	1	0.944191	0.918825	0.941565	0.944301	0.936003	0.937939	0.951672	0.941857	0.942645	0.1907539	0.2006673	0.189784	0.1601118	0.1765095	0.1863926	0.1959998	0.196999	0.1901893
S. Mauritanicus	x	x	1	0.908111	0.951269	0.957196	0.944311	0.947316	0.94615	0.950656	0.95265	0.1882539	0.1997769	0.187506	0.1595206	0.1747665	0.1850401	0.1942726	0.1953602	0.1882583
S. Mauritanicus	x	x	x	1	0.917402	0.917741	0.908459	0.90985	0.923116	0.918779	0.91191	0.1843013	0.1937837	0.184248	0.1536423	0.17063	0.1792858	0.1896223	0.1898452	0.1837161
S. Mauritanicus	x	x	x	x	1	0.951228	0.949448	0.941866	0.94281	0.942415	0.944323	0.1878346	0.1976129	0.187718	0.1576653	0.17276	0.1832367	0.1932816	0.1950788	0.1872633
S. Mauritanicus	x	x	x	x	x	1	0.949583	0.943583	0.944301	0.949471	0.954155	0.190359	0.2016656	0.189457	0.1609918	0.1772625	0.1869359	0.1956616	0.1967379	0.1903115
S. Mauritanicus	x	x	x	x	x	x	1	0.937055	0.937092	0.944867	0.944455	0.194018	0.2029512	0.193421	0.1631681	0.1784869	0.1887208	0.1990662	0.2011215	0.1931826
S. Mauritanicus	x	x	x	x	x	x	x	1	0.941346	0.938238	0.942063	0.1884154	0.1991384	0.18732	0.1577594	0.1747494	0.1846136	0.1935243	0.1944296	0.1880834
S. Mauritanicus	x	x	x	x	x	x	x	x	1	0.937905	0.944657	0.1870899	0.1974939	0.186371	0.157187	0.1728682	0.1830353	0.1927943	0.1938409	0.1867379
S. Mauritanicus	x	x	x	x	x	x	x	x	x	1	0.943297	0.1874243	0.1968267	0.18656	0.1569994	0.1728675	0.1827138	0.1928501	0.1939198	0.186579
S. Mauritanicus	x	x	x	x	x	x	x	x	x	x	1	0.1896075	0.1996747	0.188697	0.1605238	0.1753237	0.1855758	0.1960893	0.1969033	0.1894169
S. Brosseti	x	x	x	x	x	x	x	x	x	x	x	1	0.9325986	0.898361	0.858279	0.9381673	0.9203707	0.9180998	0.8632549	0.9542839
S. Brosseti	x	x	x	x	x	x	x	x	x	x	x	x	1	0.888805	0.8584929	0.9121668	0.9409346	0.925075	0.8681946	0.9402554
S. Brosseti	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.8712512	0.9017322	0.9121782	0.9201188	0.9241949	0.917141
S. Brosseti	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.8841145	0.8702095	0.892224	0.8623487	0.8758163
S. Brosseti	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.9242131	0.9261491	0.8657351	0.9392428
S. Brosseti	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.9307855	0.8876501	0.946572
S. Brosseti	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.8884024	0.9391595
S. Brosseti	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.8812209
S. Brosseti	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1

Appendix F: Table of Schoener's D values for niche overlap on *Agama impalearis.*

SPECIES	Agama_Sc	Agama_N	Agama_N	Agama_N	Agama_N	Agama_N	Agama_N	Agama_N	Agama_N	Agama_N	Agama_N	Agama_N	Agama_N	Agama_N	Agama_N	Agama_Sc	Agama_Sc	Agama_Sc	Agama_Sc	Agama_Sc	Agama_Sc	Agama_Sc	Agama_Sc	Agama_Sc	Agama_Sc	Agama_Sc	Agama_Sc	Agama_Sc	Agama_So
Agama_South	1	0.287691	0.303991	0.28306	0.298417	0.287035	0.284805	0.293272	0.303683	0.292952	0.305557	0.300478	0.307897	0.290825	0.310539	0.295731	0.875448	0.868834	0.879209	0.859337	0.875377	0.869592	0.841097	0.863582	0.8736	0.859648	0.866072	0.880633	0.872302	0.872103
Agama_North_x	x	1	0.865555	0.867183	0.870813	0.835715	0.865194	0.882076	0.853405	0.878192	0.88251	0.851952	0.863121	0.876226	0.86762	0.867937	0.286561	0.291754	0.295138	0.300138	0.305779	0.295486	0.285213	0.303133	0.278247	0.290226	0.296017	0.293574	0.289876	0.291809
Agama_North_x	x	x	1	0.870648	0.875548	0.869922	0.852404	0.878217	0.876637	0.878914	0.868126	0.866239	0.871524	0.867132	0.88351	0.886126	0.298852	0.304132	0.305982	0.310156	0.317911	0.305152	0.297384	0.31541	0.287087	0.301051	0.310685	0.307454	0.302391	0.303469
Agama_North_x	x	x	x	1	0.876193	0.862029	0.860934	0.866797	0.863177	0.872951	0.885959	0.860006	0.864759	0.857929	0.874601	0.866694	0.278429	0.28598	0.289395	0.287821	0.295331	0.287092	0.279676	0.29433	0.264529	0.28178	0.291181	0.290497	0.285316	0.282743
Agama_North_x	x	x	x	x	1	0.860849	0.861089	0.869823	0.884467	0.874976	0.891971	0.867473	0.888257	0.876249	0.878106	0.886135	0.29719	0.303038	0.30455	0.306175	0.315159	0.306431	0.298772	0.310663	0.283449	0.301301	0.308418	0.303953	0.300699	0.299514
Agama_North_x	x	x	x	x	x	1	0.850111	0.839727	0.849479	0.857832	0.857237	0.863707	0.868503	0.849256	0.857645	0.868469	0.284477	0.293257	0.290028	0.295759	0.302979	0.290961	0.284236	0.299197	0.271392	0.28706	0.29407	0.292736	0.288951	0.288612
Agama_North_x	x	x	x	x	x	x	1	0.843271	0.842073	0.875494	0.857012	0.843843	0.858553	0.850104	0.845846	0.863243	0.280017	0.288307	0.287035	0.291902	0.297971	0.286927	0.279826	0.296493	0.267899	0.283547	0.290631	0.289303	0.284536	0.285826
Agama_North_x	x	x	x	x	x	x	x	1	0.873793	0.864558	0.872169	0.858233	0.865762	0.879375	0.878392	0.86197	0.290537	0.296147	0.298403	0.304395	0.309194	0.300616	0.28916	0.307169	0.280994	0.294096	0.301629	0.298184	0.29395	0.295628
Agama_North_x	x	x	x	x	x	x	x	x	1	0.862325	0.863609	0.859462	0.865394	0.857625	0.883308	0.869143	0.30415	0.309405	0.308129	0.310226	0.322264	0.31295	0.304051	0.3141	0.287508	0.309693	0.313444	0.308268	0.30618	0.306445
Agama_North_x	x	x	x	x	x	x	x	x	x	1	0.866899	0.872297	0.875581	0.875593	0.872853	0.871532	0.288609	0.294469	0.296382	0.299727	0.308041	0.297561	0.290258	0.305258	0.276844	0.292985	0.302563	0.298626	0.293368	0.291884
Agama_North_x	x	x	x	x	x	x	x	x	x	x	1	0.870083	0.878477	0.864881	0.86576	0.876156	0.304336	0.311214	0.309994	0.314517	0.321095	0.312719	0.302118	0.318091	0.291093	0.307058	0.312479	0.312741	0.309251	0.308212
Agama_North_x	x	x	x	x	x	x	x	x	x	x	x	1	0.875999	0.858193	0.868159	0.871968	0.300171	0.307313	0.305701	0.309548	0.317139	0.308489	0.298963	0.314276	0.286567	0.303907	0.309854	0.307004	0.302804	0.301035
Agama_North_x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.865784	0.87616	0.873576	0.306819	0.312799	0.312509	0.314662	0.325331	0.315808	0.307517	0.3183	0.293821	0.310562	0.316747	0.312327	0.308757	0.308535
Agama_North_x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.861176	0.855378	0.288466	0.29417	0.297011	0.300666	0.306576	0.299235	0.289673	0.30416	0.276227	0.29185	0.29993	0.297481	0.294098	0.293273
Agama_North_x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.867797	0.307781	0.31366	0.316039	0.315599	0.326661	0.31806	0.308527	0.322869	0.294211	0.312537	0.319782	0.31701	0.312633	0.312862
Agama_South_x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.293987	0.299477	0.300336	0.304184	0.312345	0.30018	0.292858	0.310169	0.282381	0.297438	0.304341	0.301029	0.296422	0.29685
Agama_South_x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.892846	0.870842	0.862245	0.857844	0.868631	0.857263	0.875889	0.864817	0.85554	0.876525	0.850988	0.879536	0.879574
Agama_South_x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.870457	0.876026	0.862216	0.864577	0.846264	0.871365	0.856283	0.868696	0.875737	0.858977	0.864729	0.876846
Agama_South_x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.863653	0.867261	0.862314	0.832843	0.871042	0.862312	0.853497	0.867	0.862611	0.865082	0.866803
Agama_South_x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.858372	0.862574	0.827954	0.86849	0.858125	0.859171	0.859487	0.847327	0.856997	0.869021
Agama_South_x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.85731	0.842008	0.848236	0.854086	0.861801	0.858905	0.866337	0.853629	0.846419
Agama_South_x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.851994	0.861721	0.856165	0.862994	0.867967	0.852509	0.876614	0.853886
Agama_South_x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.841692	0.841545	0.853088	0.861769	0.837298	0.858547	0.837865
Agama_South_x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.86127	0.8575	0.88094	0.856186	0.877282	0.884877
Agama_South_x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.859546	0.867402	0.855531	0.849981	0.856355
Agama_South_x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.874372	0.850744	0.853107	0.846512
Agama_South_x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.85827	0.862148	0.864519
Agama_South_x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.861731	0.848551
Agama_South_x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1	0.888064
Agama_South_x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	1

Appendix G: Component scores of bioclim variables for *Chalcides mionecton.*

Component Score Coefficient Matrix				
Climatic Variables	Component			
	1	2	3	4
Annual Mean Temperature	-.093	.108	.069	.045
Mean Diurnal Range (Mean of monthly (max temp - min temp))	.061	.144	-.031	.338
Isothermality (BIO2/BIO7) (* 100)	-.041	-.048	-.040	.714
Temperature Seasonality (standard deviation *100)	.078	.145	.003	-.100
Max Temperature of Warmest Month	.023	.190	.034	.048
Min Temperature of Coldest Month	-.105	-.041	.086	-.130
Temperature Annual Range (BIO5-BIO6)	.072	.152	-.022	.102
Mean Temperature of Wettest Quarter	-.111	.030	.053	.018
Mean Temperature of Driest Quarter	-.026	.185	.062	-.036
Mean Temperature of Warmest Quarter	-.028	.184	.063	-.032
Mean Temperature of Coldest Quarter	-.114	.011	.063	.056
Annual Precipitation	.076	-.033	.216	-.014
Precipitation of Wettest Month	.052	-.033	.254	.068
Precipitation of Driest Month	.091	-.036	-.128	.122
Precipitation Seasonality (Coefficient of Variation)	-.089	-.030	.097	.330
Precipitation of Wettest Quarter	.058	-.036	.246	.029
Precipitation of Driest Quarter	.109	-.048	-.020	.020
Precipitation of Warmest Quarter	.111	-.026	-.027	.121
Precipitation of Coldest Quarter	.058	-.037	.245	.055

Appendix H: Component scores of the bioclim variables for *Chalcides ocellatus.*

Component Score Coefficient Matrix				
Climatic Variables	Component			
	1	2	3	4
Annual Mean Temperature	.090	-.065	-.076	.143
Mean Diurnal Range (Mean of monthly (max temp - min temp))	.056	.163	.210	-.045
Isothermality (BIO2/BIO7) (* 100)	.012	-.046	.319	-.569
Temperature Seasonality (standard deviation *100)	.051	.186	-.009	.195
Max Temperature of Warmest Month	.083	.104	.006	.180
Min Temperature of Coldest Month	.021	-.207	-.188	.078
Temperature Annual Range (BIO5-BIO6)	.055	.182	.095	.105
Mean Temperature of Wettest Quarter	.079	.004	-.273	-.143
Mean Temperature of Driest Quarter	.033	-.050	.315	.135
Mean Temperature of Warmest Quarter	.090	.034	-.071	.227
Mean Temperature of Coldest Quarter	.063	-.167	-.092	.087
Annual Precipitation	-.093	-.004	.068	.170
Precipitation of Wettest Month	-.089	-.029	.107	.225
Precipitation of Driest Month	-.065	.138	-.150	-.076
Precipitation Seasonality (Coefficient of Variation)	.060	-.104	.232	.084
Precipitation of Wettest Quarter	-.089	-.019	.123	.232
Precipitation of Driest Quarter	-.079	.099	-.140	-.077
Precipitation of Warmest Quarter	-.088	.038	-.168	-.088
Precipitation of Coldest Quarter	-.086	-.044	.112	.255

Appendix I: Component scores of bioclim variables for *Messalina olivieri.*

Component Score Coefficient Matrix			
Climatic Variables	Component		
	1	2	3
Annual Mean Temperature	-.086	.076	.119
Mean Diurnal Range (Mean of monthly (max temp - min temp))	.067	.119	.105
Isothermality (BIO2/BIO7) (* 100)	-.067	-.101	-.006
Temperature Seasonality (standard deviation *100)	.074	.136	.030
Max Temperature of Warmest Month	.024	.191	.117
Min Temperature of Coldest Month	-.094	-.037	.015
Temperature Annual Range (BIO5-BIO6)	.073	.133	.058
Mean Temperature of Wettest Quarter	-.080	.083	-.084
Mean Temperature of Driest Quarter	-.011	.115	.218
Mean Temperature of Warmest Quarter	-.027	.184	.137
Mean Temperature of Coldest Quarter	-.096	-.011	.067
Annual Precipitation	.082	-.087	.149
Precipitation of Wettest Month	.063	-.114	.234
Precipitation of Driest Month	.077	.008	-.250
Precipitation Seasonality (Coefficient of Variation)	-.084	-.061	.066
Precipitation of Wettest Quarter	.067	-.106	.239
Precipitation of Driest Quarter	.086	-.003	-.191
Precipitation of Warmest Quarter	.085	-.018	-.183
Precipitation of Coldest Quarter	.056	-.121	.258

Appendix J: Component scores of bioclim variables for *Trogonophis weigmanni*.

	Component			
	1	2	3	4
Annual Mean Temperature	-.092	-.009	.123	-.028
Mean Diurnal Range (Mean of monthly (max temp - min temp))	.079	-.020	.116	.470
Isothermality (BIO2/BIO7) (* 100)	-.051	.047	-.077	.764
Temperature Seasonality (standard deviation *100)	.090	-.043	.127	-.028
Max Temperature of Warmest Month	.039	-.043	.247	.115
Min Temperature of Coldest Month	-.101	.018	.016	-.184
Temperature Annual Range (BIO5-BIO6)	.089	-.035	.123	.187
Mean Temperature of Wettest Quarter	-.086	-.065	.044	.048
Mean Temperature of Driest Quarter	-.032	-.047	.254	-.067
Mean Temperature of Warmest Quarter	-.037	-.046	.249	-.077
Mean Temperature of Coldest Quarter	-.102	.012	.040	-.037
Annual Precipitation	.045	.214	.050	-.066
Precipitation of Wettest Month	.013	.236	.059	-.014
Precipitation of Driest Month	.097	-.035	-.052	-.024
Precipitation Seasonality (Coefficient of Variation)	-.085	.116	.000	.184
Precipitation of Wettest Quarter	.014	.237	.054	-.034
Precipitation of Driest Quarter	.100	.009	-.030	-.070
Precipitation of Warmest Quarter	.098	.001	-.030	-.139
Precipitation of Coldest Quarter	.009	.238	.058	-.048

Appendix K: Component scores of bioclim variables for *Saurodactylus mauritanicus*

	Component		
	1	2	3
Annual Mean Temperature	-.077	.090	.167
Mean Diurnal Range (Mean of monthly (max temp - min temp))	.059	.132	-.180
Isothermality (BIO2/BIO7) (* 100)	-.055	-.082	-.274
Temperature Seasonality (standard deviation *100)	.068	.138	-.034
Max Temperature of Warmest Month	.027	.203	.008
Min Temperature of Coldest Month	-.082	-.039	.169
Temperature Annual Range (BIO5-BIO6)	.066	.139	-.099
Mean Temperature of Wettest Quarter	-.081	.038	.105
Mean Temperature of Driest Quarter	-.014	.197	.215
Mean Temperature of Warmest Quarter	-.021	.204	.152
Mean Temperature of Coldest Quarter	-.086	-.005	.148
Annual Precipitation	.082	-.049	.217
Precipitation of Wettest Month	.070	-.077	.304
Precipitation of Driest Month	.072	-.029	-.234
Precipitation Seasonality (Coefficient of Variation)	-.076	-.035	-.119
Precipitation of Wettest Quarter	.073	-.072	.303
Precipitation of Driest Quarter	.081	-.025	-.098
Precipitation of Warmest Quarter	.081	-.015	-.128
Precipitation of Coldest Quarter	.070	-.079	.332

Appendix L: Component scores of the bioclim variables for *Agama impalearis*.

	Component		
	1	2	3
Annual Mean Temperature	-.879	.289	.298
Mean Diurnal Range (Mean of monthly (max temp - min temp))	.514	.733	.132
Isothermality (BIO2/BIO7) (* 100)	-.605	-.435	-.255
Temperature Seasonality (standard deviation *100)	.622	.755	.175
Max Temperature of Warmest Month	.090	.911	.360
Min Temperature of Coldest Month	-.886	-.331	.100
Temperature Annual Range (BIO5-BIO6)	.596	.764	.162
Mean Temperature of Wettest Quarter	-.803	.318	-.006
Mean Temperature of Driest Quarter	-.299	.546	.457
Mean Temperature of Warmest Quarter	-.385	.790	.407
Mean Temperature of Coldest Quarter	-.956	-.148	.151
Annual Precipitation	.700	-.511	.486
Precipitation of Wettest Month	.532	-.581	.606
Precipitation of Driest Month	.798	.072	-.439
Precipitation Seasonality (Coefficient of Variation)	-.742	-.367	.232
Precipitation of Wettest Quarter	.544	-.562	.616
Precipitation of Driest Quarter	.905	-.041	-.242
Precipitation of Warmest Quarter	.892	-.138	-.267
Precipitation of Coldest Quarter	.495	-.606	.617

Appendix M: sensitivity and specificity analysis for whole species and individual subspecies.

Sensitivity & Specificity analysis

	Whole species	Subspecies	
	Chalcides mionecton	*C. m. mionecton*	*C. m. trifasciatus*
Sensitivity	0.947	0.951	0.937
Specificity	0.8662	0.879	0.932

	Whole species	Subspecies	
	Chalcides ocellatus	*C. o. ocellatus*	*C. o. subtypicus*
Sensitivity	0.933	0.911	0.963
Specificity	0.77	0.808	0.905

	Whole species	Subspecies	
	Agama impalearis	*A. i. north*	*A. i. south*
Sensitivity	0.946	0.869	0.839
Specificity	0.286	0.759	0.795

	Whole species	Subspecies	
	Saurodactylus mauritanicus	*S. m. mauritanicus*	*S. m. brosseti*
Sensitivity	0.956	0.823	0.958
Specificity	0.755	0.898	0.952

	Whole species	Subspecies	
	Mesalina olivieri	*M. o. olivieri*	*M. o. simoni*
Sensitivity	0.929	0.834	0.952
Specificity	0.803	0.828	0.915

	Whole species	Subspecies	
	Trogonophis wiegmanni	*T. w. wiegmanni*	*T. w. elegans*
Sensitivity	0.944	0.892	0.944
Specificity	0.699	0.87	0.829

Appendix N: Details of data input within MaxEnt.

Species	Sample size	Background points	Features selected through "Auto" selection
C. m. mionecton	47	10000	Hinge, Linear, Quadratic
C. m. trifasciatus	24	10000	Hinge, Linear, Quadratic
C. o. ocellatus	34	10000	Hinge, Linear, Quadratic
C. o. subtypicus	42	10000	Hinge, Linear, Quadratic
A. i. north	265	10000	Hinge, Linear, Quadratic, Threshold, Product
A. i. south	331	10000	Hinge, Linear, Quadratic, Threshold, Product
T. w. weigmanni	28	10000	Hinge, Linear, Quadratic
T. w. elegans	67	10000	Hinge, Linear, Quadratic
S. m. brosseti	174	10000	Hinge, Linear, Quadratic, Threshold, Product
S. m. mauritanicus	18	10000	Hinge, Linear, Quadratic
M. o. simoni	32	10000	Hinge, Linear, Quadratic
M. o. olivieri	87	10000	Hinge, Linear, Quadratic, Threshold, Product

Appendix O: PCA on bioclim dataset.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	8.529	44.887	44.887	8.529	44.887	44.887
2	5.191	27.320	72.208	5.191	27.320	72.208
3	2.402	12.641	84.848	2.402	12.641	84.848
4	.997	5.248	90.096	.997	5.248	90.096
5	.657	3.455	93.552			
6	.472	2.482	96.034			
7	.344	1.812	97.846			
8	.188	.992	98.838			
9	.112	.588	99.426			
10	.060	.316	99.742			
11	.022	.115	99.857			
12	.008	.045	99.901			
13	.007	.039	99.940			
14	.005	.028	99.968			
15	.002	.012	99.980			
16	.002	.010	99.990			
17	.001	.006	99.996			
18	.001	.004	100.000			
19	2.640E-15	1.389E-14	100.000			

Appendix P: Shapefile of the geographical area to which the species niches were projected on to laid over the top of the study area image.
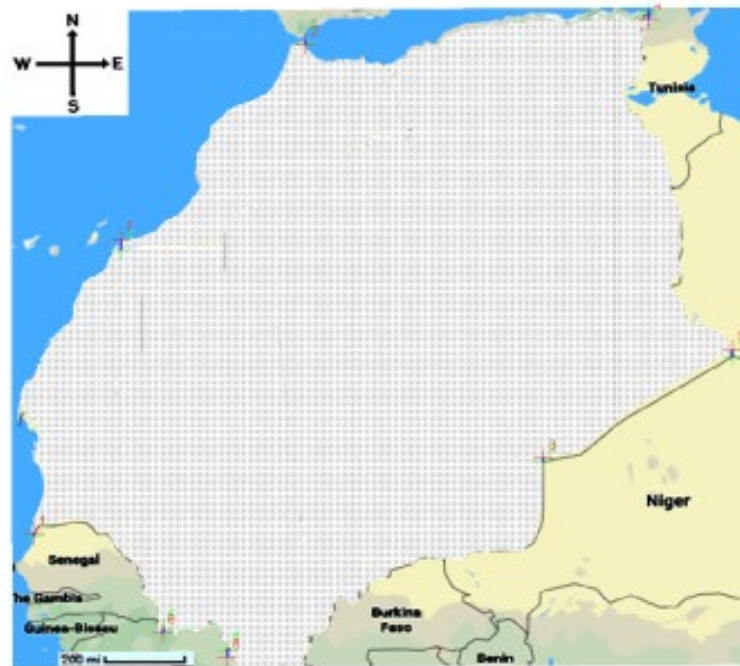
References

Ahmadzadeh, F., Flecks, M., Carretero, M.A., Böhme, W., Ilgaz, C., Engler, J.O., Harris, D.J., Üzüm, N. and Rödder, D., 2013. Rapid lizard radiation lacking niche conservatism: ecological diversification within a complex landscape. *Journal of Biogeography*, *40*(9), pp.1807-1818.

Allouche, O., Tsoar, A. and Kadmon, R., 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of applied ecology*, *43*(6), pp.1223-1232.

Anderson, R.P., Lew, D. and Peterson, A.T., 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological modelling*, *162*(3), pp.211-232.

Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J Hijmans, R., Huettmann, F., R Leathwick, J., Lehmann, A., Li, J., G Lohmann, L. and A Loiselle, B., 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, *29*(2), pp.129-151.

Araujo, M.B. and Guisan, A., 2006. Five (or so) challenges for species distribution modelling. *Journal of biogeography*, *33*(10), pp.1677-1688.

Araújo, M.B., Thuiller, W. and Pearson, R.G., 2006. Climate warming and the decline of amphibians and reptiles in Europe. *Journal of biogeography*, *33*(10), pp.1712-1728.

Araújo, M.B., Pearson, R.G., Thuiller, W. and Erhard, M., 2005. Validation of species–climate impact models under climate change. *Global Change Biology*, *11*(9), pp.1504-1513.

Austin, M.P. and Van Niel, K.P., 2011. Improving species distribution models for climate change studies: variable selection and scale. *Journal of Biogeography*, *38*(1), pp.1-8.

Austin, MP., 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches. *Ecological modelling*, *200*(1), pp.1-19.

Austin, M.P., 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological modelling*,*157*(2), pp.101-118.

Barnes, I., Matheus, P., Shapiro, B., Jensen, D. and Cooper, A., 2002. Dynamics of Pleistocene population extinctions in Beringian brown bears. *Science*, *295*(5563), pp.2267-2270.

Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S.P., Peterson, A.T., Soberón, J. and Villalobos, F., 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, *222*(11), pp.1810-1819.

Beghin, P., Charbit, S., Kageyama, M., Combourieu-Nebout, N., Hatté, C., Dumas, C. and Peterschmitt, J.Y., 2016. What drives LGM precipitation over the western Mediterranean? A study focused on the Iberian Peninsula and northern Morocco. *Climate dynamics*, *46*(7-8), pp.2611-2631.

Bell, R.C., Parra, J.L., Tonione, M., Hoskin, C.J., Mackenzie, J.B., Williams, S.E. and Moritz, C., 2010. Patterns of persistence and isolation indicate resilience to climate change in montane rainforest lizards. *Molecular Ecology*, *19*(12), pp.2531-2544.

Benito Garzón, M., Alía, R., Robson, T.M. and Zavala, M.A., 2011. Intra-specific variability and plasticity influence potential tree species distributions under climate change. *Global Ecology and Biogeography*, *20*(5), pp.766-778.

Berendse, F., 1983. Interspecific competition and niche differentiation between Plantago lanceolata and Anthoxanthum odoratum in a natural hayfield. *The Journal of Ecology*, pp.379-390.

Bons, J. and Geniez, P., 1996. Amphibians and reptiles of Morocco. *Barcelona, Asociación Herpetológica Española*.

Boulangeat, I., Gravel, D. and Thuiller, W., 2012. Accounting for dispersal and biotic interactions to disentangle the drivers of species distributions and their abundances. *Ecology letters*, *15*(6), pp.584-593.

Boyle, M., Schwanz, L., Hone, J. and Georges, A., 2016. Dispersal and climate warming determine range shift in model reptile populations. *Ecological modelling*, *328*, pp.34-43.

BRAGG, W.K., FAWCETT, J.D., BRAGG, T.B. and VIETS, B.E., 2000. Nest-site selection in two eublepharid gecko species with temperature-dependent sex determination and one with genotypic sex determination. *Biological Journal of the Linnean Society*, *69*(3), pp.319-332.

Brown, R.P., Suárez, N.M. and Pestano, J., 2002. The Atlas Mountains as a biogeographical divide in North–West Africa: evidence from mtDNA evolution in the Agamid lizard Agama impalearis. *Molecular phylogenetics and evolution*, *24*(2), pp.324-332.

Brown, R.P., Tejangkura, T., El Mouden, E.H., Baamrane, A. and Znari, M., 2012. Species delimitation and digit number in a North African skink.*Ecology and evolution*, *2*(12), pp.2962-2973.

Bush, M.B. and Oliveira, P.E.D., 2006. The rise and fall of the Refugial Hypothesis of Amazonian speciation: a paleoecological perspective. *Biota Neotropica*, *6*(1), pp.0-0.

Camargo, A., Sinervo, B. and Sites Jr, J.W., 2010. Lizards as model organisms for linking phylogeographic and speciation studies. *Molecular Ecology*, *19*(16), pp.3250-3270.

Cahill, A.E., Aiello-Lammens, M.E., Fisher-Reid, M.C., Hua, X., Karanewsky, C.J., Ryu, H.Y., Sbeglia, G.C., Spagnolo, F., Waldron, J.B., Warsi, O. and Wiens, J.J., 2012, October. How does climate change cause extinction?. In *Proc. R. Soc. B* (p. rspb20121890). The Royal Society.

Carpenter, G., Gillison, A.N. and Winter, J., 1993. DOMAIN: a flexible modelling procedure for mapping potential distributions of plants and animals. *Biodiversity & Conservation*, *2*(6), pp.667-680.

Carretero, M.A., Znari, M., Harris, D.J. and Macé, J.C., 2005. Morphological divergence among populations of Testudo graeca from west-central Morocco. *Animal Biology*, *55*(3), pp.259-279.

Carstens, B.C. and Knowles, L.L., 2007. Estimating species phylogeny from gene-tree probabilities despite incomplete lineage sorting: an example from Melanoplus grasshoppers. *Systematic Biology*, *56*(3), pp.400-411.

Cattell, R.B., 1966. The scree test for the number of factors. *Multivariate behavioral research*, *1*(2), pp.245-276.

Censky, E.J., Hodge, K. and Dudley, J., 1998. Over-water dispersal of lizards due to hurricanes. *Nature*, *395*(6702), p.556.

CHAMAILLÉ-JAMMES, S.I.M.O.N., Massot, M., Aragon, P. and Clobert, J., 2006. Global warming and positive fitness response in mountain populations of common lizards Lacerta vivipara. *Global Change Biology*, *12*(2), pp.392-402.

Chakraborty, A., Gelfand, A.E., Wilson, A.M., Latimer, A.M. and Silander, J.A., 2011. Point pattern modelling for degraded presence-only data over large regions. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *60*(5), pp.757-776.

Chapman, A.D., 1999. Quality control and validation of point-sourced environmental resource data. *Spatial accuracy assessment: Land information uncertainty in natural resources*, pp.409-418.

Chumakov, N.M., 1984. The principal glacial events of the past and their geologic significance. *International Geology Review*, *26*(8), pp.869-888.

Condit, R., Pitman, N., Leigh, E.G., Chave, J., Terborgh, J., Foster, R.B., Núñez, P., Aguilar, S., Valencia, R., Villa, G. and Muller-Landau, H.C., 2002. Beta-diversity in tropical forest trees. *Science*, *295*(5555), pp.666-669.

Cronin, T.M., DeMartino, D.M., Dwyer, G.S. and Rodriguez-Lazaro, J., 1999. Deep-sea ostracode species diversity: response to late Quaternary climate change. *Marine Micropaleontology*, *37*(3), pp.231-249.

Dahdouh-Guebas, F., Van Pottelbergh, I., Kairo, J.G., Cannicci, S. and Koedam, N., 2004. Human-impacted mangroves in Gazi (Kenya): predicting future vegetation based on retrospective remote sensing, social surveys, and tree distribution. *Marine Ecology Progress Series*, *272*, pp.77-92.

Davey, C.M., Chamberlain, D.E., Newson, S.E., Noble, D.G. and Johnston, A., 2012. Rise of the generalists: evidence for climate driven homogenization in avian communities. *Global Ecology and Biogeography*, *21*(5), pp.568-578.

Davis, A.P., Gole, T.W., Baena, S. and Moat, J., 2012. The impact of climate change on indigenous arabica coffee (Coffea arabica): predicting future trends and identifying priorities. *PloS one*, *7*(11), p.e47981.

Davis, A.J., Jenkinson, L.S., Lawton, J.H., Shorrocks, B. and Wood, S., 1998. Making mistakes when predicting shifts in species range in response to global warming. *Nature*, *391*(6669), p.783.

Demenocal, P.B., Ruddiman, W.F. and Pokras, E.M., 1993. Influences of high-and low-latitude processes on African terrestrial climate: Pleistocene eolian records from equatorial Atlantic Ocean Drilling Program site 663. *Paleoceanography*, *8*(2), pp.209-242.

DeMenocal, P., 1991. A rock-magnetic record of monsoonal dust deposition to the Arabian Sea: evidence for a shift in the mode of deposition at 2.4 Ma. *Upwelling system: evolution since the Early Miocene. Geological Society*, (64), pp.389-407.

DeMenocal, P.B., 1995. Plio-Pleistocene African climate. *Science (New York, NY)*, *270*(5233), pp.53-59.

DeMenocal, P.B., 2004. African climate change and faunal evolution during the Pliocene–Pleistocene. *Earth and Planetary Science Letters*, *220*(1-2), pp.3-24.

Dickman, C.R., Letnic, M. and Mahon, P.S., 1999. Population dynamics of two species of dragon lizards in arid Australia: the effects of rainfall. *Oecologia*, *119*(3), pp.357-366.

Dirnböck, T., Essl, F. and Rabitsch, W., 2011. Disproportional risk for habitat loss of high-altitude endemic species under climate change. *Global Change Biology*, *17*(2), pp.990-996.

Dupont, L.M., Donner, B., Schneider, R. and Wefer, G., 2001. Mid-Pleistocene environmental change in tropical Africa began as early as 1.05 Ma. *Geology*, *29*(3), pp.195-198

Elith, J., Kearney, M. and Phillips, S., 2010. The art of modelling range-shifting species. *Methods in ecology and evolution*, *1*(4), pp.330-342.

Elith, J., Leathwick, J.R. and Hastie, T., 2008. A working guide to boosted regression trees. *Journal of Animal Ecology*, *77*(4), pp.802-813.

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E. and Yates, C.J., 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and distributions*, *17*(1), pp.43-57.

Elith, J. and Leathwick, J.R., 2009. Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, *40*(1), p.677.

Etterson, J.R. and Shaw, R.G., 2001. Constraint to adaptive evolution in response to global warming. *Science*, *294*(5540), pp.151-154.

Frontier, S., 1976. Decrease of eigenvalues in principal component analysis-comparison with broken stick model. *Journal of Experimental Marine Biology and Ecology*, *25*(1), pp.67-75.

Gasse, F., Cortijo, E., Disnar, J.R., Ferry, L., Gibert, E., Kissel, C., Laggoun-Défarge, F., Lallier-Verges, E., Miskovsky, J.C., Ratsimbazafy, B. and Ranaivo, F., 1994. A 36 ka environmental record in the southern tropics: Lake Tritrivakely (Madagascar) (Un enregistrement de l'environnement depuis 36 ka en zone tropicale sud: le lac Tritrivakely (Madagascar)). *Comptes rendus de l'Académie des sciences. Série IIa, Sciences de la terre et des planètes*, *318*, pp.1513-1519.

Graham, C.H., Elith, J., Hijmans, R.J., Guisan, A., Townsend Peterson, A. and Loiselle, B.A., 2008. The influence of spatial errors in species occurrence data used in distribution models. *Journal of Applied Ecology*,*45*(1), pp.239-247.

Griffin, D.L., 2002. Aridity and humidity: two aspects of the late Miocene climate of North Africa and the Mediterranean. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *182*(1), pp.65-91.

Guisan, Antoine, Thomas C. Edwards, and Trevor Hastie. "Generalized linear and generalized additive models in studies of species distributions: setting the scene." *Ecological modelling* 157.2 (2002): 89-100.

Guisan, A., Tingley, R., Baumgartner, J.B., Naujokaitis-Lewis, I., Sutcliffe, P.R., Tulloch, A.I., Regan, T.J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C. and Martin, T.G., 2013. Predicting species distributions for conservation decisions. *Ecology letters*, *16*(12), pp.1424-1435.

Guisan, A. and Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. *Ecological modelling*, *135*(2), pp.147-186.

Grootes, P.M., Stuiver, M., White, J.W.C., Johnsen, S. and Jouzel, J., 1993. Comparison of oxygen isotope records from the GISP2 and GRIP Greenland ice cores. *Nature*, *366*(6455), p.552.

Hays, J.D., Imbrie, J. and Shackleton, N.J., 1976. Astronomical theory of ice ages confirmed. *Science*, *194*(4270), pp.1121-1132.

Hewitt, G., 2000. The genetic legacy of the Quaternary ice ages. *Nature*, *405*(6789), pp.907-913.

Hewitt, G.M., 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *359*(1442), pp.183-195.

Hertz, P.E., Huey, R.B. and Nevo, E., 1982. Fight versus flight: body temperature influences defensive responses of lizards. *Animal Behaviour*, *30*(3), pp.676-679.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. and Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas.*International journal of climatology*, *25*(15), pp.1965-1978.

Hijmans, R.J., 2012. Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. *Ecology*,*93*(3), pp.679-688.

Hijmans, R.J., Schreuder, M., De la Cruz, J. and Guarino, L., 1999. Using GIS to check co-ordinates of genebank accessions. *Genetic resources and crop evolution*, *46*(3), pp.291-296.

Hof, A.R., Jansson, R. and Nilsson, C., 2012. How biotic interactions may alter future predictions of species distributions: future threats to the persistence of the arctic fox in Fennoscandia. *Diversity and Distributions*, *18*(6), pp.554-562.

Holt, R.D., 1996. Adaptive evolution in source-sink environments: direct and indirect effects of density-dependence on niche evolution. *Oikos*, pp.182-192.

Holt, R.D. and Gomulkiewicz, R., 1997. The evolution of species' niches: a population dynamic perspective. *Case studies in mathematical modelling: ecology, physiology, and cell biology*, pp.25-50.

Hoskin, C.J., Tonione, M., Higgie, M., MacKenzie, J.B., Williams, S.E., VanDerWal, J. and Moritz, C., 2011. Persistence in peripheral refugia promotes phenotypic divergence and speciation in a rainforest frog. *The American Naturalist*, *178*(5), pp.561-578.

Hunter, D.O., Britz, T., Jones, M. and Letnic, M., 2015. Reintroduction of Tasmanian devils to mainland Australia can restore top-down control in ecosystems where dingoes have been extirpated. *Biological Conservation*, *191*, pp.428-435.

Hsü, K.J., Ryan, W.B. and Cita, M.B., 1973. Late Miocene desiccation of the Mediterranean. *Nature*, *242*(5395), pp.240-244.

Ihlow, F., Dambach, J., Engler, J.O., Flecks, M., Hartmann, T., Nekum, S., Rajaei, H. and Rödder, D., 2012. On the brink of extinction? How climate change may affect global chelonian species richness and distribution. *Global Change Biology*, *18*(5), pp.1520-1530.

Imbrie, J. and Imbrie, J.Z., 1980. Modeling the climatic response to orbital variations. *Science*, *207*(4434), pp.943-953.

Jaynes, E.T., 1957. Information theory and statistical mechanics. *Physical review*, *106*(4), p.620.

Jaynes, E.T., 1986. Bayesian methods: General background.

Johnson, C.J. and Gillingham, M.P., 2008. Sensitivity of species-distribution models to error, bias, and model design: an application to resource selection functions for woodland caribou. *Ecological Modelling*, *213*(2), pp.143-155.

Kaliontzopoulou, A., Brito, J.C., Carretero, M.A., Larbes, S. and 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. *Canadian Journal of Zoology*, *86*(9), pp.992-1001.

Kapli, P., Lymberakis, P., Crochet, P.A., Geniez, P., Brito, J.C., Almutairi, M., Ahmadzadeh, F., Schmitz, A., Wilms, T., Pouyani, N.R. and Poulakakis, N., 2015. Historical biogeography of the lacertid lizard Mesalina in North Africa and the Middle East. *Journal of Biogeography*, *42*(2), pp.267-279.

Kawecki, T.J., 1995. Demography of source—sink populations and the evolution of ecological niches. *Evolutionary Ecology*, *9*(1), pp.38-44.

Keenan, T., Maria Serra, J., Lloret, F., Ninyerola, M. and Sabate, S., 2011. Predicting the future of forests in the Mediterranean under climate change, with niche-and process-based models: CO_2 matters!. *Global change biology*, *17*(1), pp.565-579.

Knowles, L.L., Carstens, B.C. and Keat, M.L., 2007. Coupling genetic and ecological-niche models to examine how past population distributions contribute to divergence. *Current Biology*, *17*(11), pp.940-946.

Kornilios, P., Kyriazi, P., Poulakakis, N., Kumlutaş, Y., Ilgaz, Ç., Mylonas, M. and Lymberakis, P., 2010. Phylogeography of the ocellated skink Chalcides ocellatus (Squamata, Scincidae), with the use of mtDNA sequences: A hitch-hiker's guide to the Mediterranean. *Molecular phylogenetics and evolution*, *54*(2), pp.445-456.

Kozak, K.H. and Wiens, J.J., 2010. Accelerated rates of climatic-niche evolution underlie rapid species diversification. *Ecology Letters*, *13*(11), pp.1378-1389.

Kukla, G.J., Bender, M.L., de Beaulieu, J.L., Bond, G., Broecker, W.S., Cleveringa, P., Gavin, J.E., Herbert, T.D., Imbrie, J., Jouzel, J. and Keigwin, L.D., 2002. Last interglacial climates. *Quaternary Research*, *58*(1), pp.2-13.

Kukla, G. and Gavin, J., 2004. Milankovitch climate reinforcements. *Global and Planetary Change*, *40*(1-2), pp.27-48.

Kuchta, S.R. and TAN, A.M., 2005. Isolation by distance and post-glacial range expansion in the rough-skinned newt, Taricha granulosa. *Molecular Ecology*, *14*(1), pp.225-244.

Lantz, C.A. and Nebenzahl, E., 1996. Behavior and interpretation of the κ statistic: Resolution of the two paradoxes. *Journal of clinical epidemiology*, *49*(4), pp.431-434.

Le Roux, P.C., Pellissier, L., Wisz, M.S. and Luoto, M., 2014. Incorporating dominant species as proxies for biotic interactions strengthens plant community models. *Journal of Ecology*, *102*(3), pp.767-775.

Leathwick, J.R., Elith, J. and Hastie, T., 2006. Comparative performance of generalized additive models and multivariate adaptive regression splines for statistical modelling of species distributions. *Ecological modelling*, *199*(2), pp.188-196.

LEE-YAW, J.A., Irwin, J.T. and Green, D.M., 2008. Postglacial range expansion from northern refugia by the wood frog, Rana sylvatica. *Molecular Ecology*, *17*(3), pp.867-884.

Lehmann, A., Overton, J.M. and Leathwick, J.R., 2002. GRASP: generalized regression analysis and spatial prediction. *Ecological modelling*, *157*(2-3), pp.189-207.

Leonard, J.A., Wayne, R.K. and Cooper, A., 2000. Population genetics of Ice Age brown bears. *Proceedings of the National Academy of Sciences*, *97*(4), pp.1651-1654.

Levsen, N.D., Tiffin, P. and Olson, M.S., 2012. Pleistocene speciation in the genus Populus (Salicaceae). *Systematic biology*, *61*(3), pp.401-412.

Liu, C., Berry, P.M., Dawson, T.P. and Pearson, R.G., 2005. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, *28*(3), pp.385-393.

Lomolino, M.V., 2010. Four Darwinian themes on the origin, evolution and preservation of island life. *Journal of biogeography*, *37*(6), pp.985-994.

Lobo, J.M., Jiménez-Valverde, A. and Real, R., 2008. AUC: a misleading measure of the performance of predictive distribution models. *Global ecology and Biogeography*, *17*(2), pp.145-151.

Luchetti, A., Marini, M. and Mantovani, B., 2005. Mitochondrial evolutionary rate and speciation in termites: data on European Reticulitermes taxa (Isoptera, Rhinotermitidae). *Insectes Sociaux*, *52*(3), pp.218-221.

Mahiny, A.S. and Turner, B.J., 2003, September. Modeling past vegetation change through remote sensing and GIS: a comparison of neural networks and logistic regression methods. In *Proceedings of the 7th international conference on geocomputation. University of Southampton, UK*.

Maley, J., 1996. The African rain forest–main characteristics of changes in vegetation and climate from the Upper Cretaceous to the Quaternary. *Proceedings of the Royal Society of Edinburgh. Section B. Biological Sciences*, *104*, pp.31-73

Manel, S., Williams, H.C. and Ormerod, S.J., 2001. Evaluating presence–absence models in ecology: the need to account for prevalence. *Journal of applied Ecology*, *38*(5), pp.921-931.

Mateo, R.G., Croat, T.B., Felicísimo, Á.M. and Munoz, J., 2010. Profile or group discriminative techniques? Generating reliable species distribution models using pseudo-absences and target-group absences from natural history collections. *Diversity and Distributions*, *16*(1), pp.84-94.

Mayr, E., 1942. *Systematics and the origin of species, from the viewpoint of a zoologist*. Harvard University Press.

McKenzie, G.M. and Busby, J.R., 1992. A Quantitative Estimate of Holocene Climate Using a Bioclimatic Profile of Nothofagus cunninghamii (Hook.) Oerst. *Journal of Biogeography*, pp.531-540.

Michard, A., Saddiqi, O., Chalouan, A. and de Lamotte, D.F. eds., 2008. *Continental evolution: The geology of Morocco: Structure, stratigraphy, and tectonics of the Africa-Atlantic-Mediterranean triple junction* (Vol. 116). Springer.

Milner, A.M., Collier, R.E., Roucoux, K.H., Müller, U.C., Pross, J., Kalaitzidis, S., Christanis, K. and Tzedakis, P.C., 2012. Enhanced seasonality of precipitation in the Mediterranean during the early part of the Last Interglacial. *Geology*, *40*(10), pp.919-922.

Misof, B., Liu, S., Meusemann, K., Peters, R.S., Donath, A., Mayer, C., Frandsen, P.B., Ware, J., Flouri, T., Beutel, R.G. and Niehuis, O., 2014. Phylogenomics resolves the timing and pattern of insect evolution. *Science*, *346*(6210), pp.763-767.

Nix, H.A. and Busby, J., 1986. BIOCLIM, a bioclimatic analysis and prediction system. *Annual report CSIRO. CSIRO Division of Water and Land Resources, Canberra*.

Otto-Bliesner, B.L., Brady, E.C., Clauzet, G., Tomas, R., Levis, S. and Kothavala, Z., 2006. Last glacial maximum and Holocene climate in CCSM3.*Journal of Climate*, *19*(11), pp.2526-2544.

Pearman, P.B., Guisan, A., Broennimann, O. and Randin, C.F., 2008. Niche dynamics in space and time. *Trends in Ecology & Evolution*, *23*(3), pp.149-158.

Peterson, A.T. and Holt, R.D., 2003. Niche differentiation in Mexican birds: using point occurrences to detect ecological innovation. *Ecology Letters*,*6*(8), pp.774-782.

Phillips, S.J., 2005. A brief tutorial on Maxent. *AT&T Research*.

Phillips, S.J. and Dudík, M., 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, *31*(2), pp.161-175.

Phillips, S.J., Dudík, M. and Schapire, R.E., 2004, July. A maximum entropy approach to species distribution modeling. In *Proceedings of the twenty-first international conference on Machine learning* (p. 83). ACM.

Phillips, S.J., Dudík, M., Elith, J., Graham, C.H., Lehmann, A., Leathwick, J. and Ferrier, S., 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications*, *19*(1), pp.181-197.

Pineda, E. and Lobo, J.M., 2009. Assessing the accuracy of species distribution models to predict amphibian species richness patterns. *Journal of Animal Ecology*, *78*(1), pp.182-190.

Potter, C., 2004. Predicting climate change effects on vegetation, soil thermal dynamics, and carbon cycling in ecosystems of interior Alaska. *Ecological Modelling*, *175*(1), pp.1-24.

Pulliam, H.R., 2000. On the relationship between niche and distribution. *Ecology letters*, *3*(4), pp.349-361.

Raes, N. and ter Steege, H., 2007. A null-model for significance testing of presence-only species distribution models. *Ecography*, *30*(5), pp.727-736.

Rato, C. and Harris, D.J., 2008. Genetic variation within Saurodactylus and its phylogenetic relationships within the Gekkonoidea estimated from mitochondrial and nuclear DNA sequences. *Amphibia-Reptilia*, *29*(1), pp.25-34.

Robinson, G.R., Holt, R.D., Gaines, M.S., Hamburg, S.P., Johnson, M.L., Fitch, H.S. and Martinko, E.A., 1992. Diverse and contrasting effects of habitat fragmentation. *Science (Washington)*, *257*(5069), pp.524-526.

Rödder, D. and Engler, J.O., 2011. Quantitative metrics of overlaps in Grinnellian niches: advances and possible drawbacks. *Global Ecology and Biogeography*, *20*(6), pp.915-927.

Rossetto, M., Allen, C.B., Thurlby, K.A., Weston, P.H. and Milner, M.L., 2012. Genetic structure and bio-climatic modeling support allopatric over parapatric speciation along a latitudinal gradient. *BMC evolutionary biology*, *12*(1), p.149.

Rossignol-Strick, M., 1983. African monsoons, an immediate climate response to orbital insolation. *Nature*, *304*(5921), p.46.

Ryan, M.J., Latella, I.M., Giermakowski, J.T., Snell, H., Poe, S., Pangle, R.E., Gehres, N., Pockman, W.T. and McDowell, N.G., 2016. Too dry for lizards: short-term rainfall influence on lizard microhabitat use in an experimental rainfall manipulation within a piñon-juniper. *Functional Ecology*, *30*(6), pp.964-973.

Sánchez, A. and Escoriza, D., 2014. Checkerboard worm lizard (Trogonophis wiegmanni) new records and description of its ecological niche in North-Western Africa. *Bulletin de la Societé Herpétologique de France*, *152*, pp.29-36.

Segurado, P. and Araujo, M.B., 2004. An evaluation of methods for modelling species distributions. *Journal of Biogeography*, *31*(10), pp.1555-1568.

Shao, G. and Halpin, P.N., 1995. Climatic controls of eastern North American coastal tree and shrub distributions. *Journal of Biogeography*, pp.1083-1089.

Shine, R., Barrott, E.G. and Elphick, M.J., 2002. Some like it hot: effects of forest clearing on nest temperatures of montane reptiles. *Ecology*, *83*(10), pp.2808-2815.

Silva, D.P., Varela, S., Nemésio, A. and De Marco Jr, P., 2015. Adding biotic interactions into paleodistribution models: a host-cleptoparasite complex of Neotropical orchid bees. *PLoS One*, *10*(6), p.e0129890.

Soberon, J., 2005. Interpretation of models of fundamental ecological niches and species' distributional areas.

Spellerberg, I.F., 1972. Temperature tolerances of southeast Australian reptiles examined in relation to reptile thermoregulatory behaviour and distribution. *Oecologia*, *9*(1), pp.23-46.

Steininger, F.F. and Rögl, F., 1984. Paleogeography and palinspastic reconstruction of the Neogene of the Mediterranean and Paratethys. *Geological Society, London, Special Publications*, *17*(1), pp.659-668.

Stockwell, D., 1999. The GARP modelling system: problems and solutions to automated spatial prediction. *International journal of geographical information science*, *13*(2), pp.143-158.

Stryszowska, K.M., Johnson, G., Mendoza, L.R. and Langen, T.A., 2016. Species Distribution Modeling of the Threatened Blanding's Turtle's (Emydoidea blandingii) Range Edge as a Tool for Conservation Planning. *Journal of Herpetology*, *50*(3), pp.366-373.

Svenning, J.C., Eiserhardt, W.L., Normand, S., Ordonez, A. and Sandel, B., 2015. The influence of paleoclimate on present-day patterns in biodiversity and ecosystems. *Annual Review of Ecology, Evolution, and Systematics*, *46*, pp.551-572.

Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus, B.F., De Siqueira, M.F., Grainger, A., Hannah, L. and Hughes, L., 2004. Extinction risk from climate change. *Nature*, *427*(6970), p.145.

Trabucco, A., Achten, W.M., Bowe, C., Aerts, R.A.F., ORSHOVEN, J.V., Norgrove, L. and Muys, B., 2010. Global mapping of Jatropha curcas yield based on response of fitness to present and future climate. *GCB Bioenergy*, *2*(3), pp.139-151.

Trauth, M.H., Larrasoaña, J.C. and Mudelsee, M., 2009. Trends, rhythms and events in Plio-Pleistocene African climate. *Quaternary Science Reviews*, *28*(5), pp.399-411.

Tylianakis, J.M., Didham, R.K., Bascompte, J. and Wardle, D.A., 2008. Global change and species interactions in terrestrial ecosystems. *Ecology letters*, *11*(12), pp.1351-1363.

Valinia, S., Englund, G., Moldan, F., Futter, M.N., Köhler, S.J., Bishop, K. and Fölster, J., 2014. Assessing anthropogenic impact on boreal lakes with historical fish species distribution data and hydrogeochemical modeling. *Global change biology*, *20*(9), pp.2752-2764.

Vrba, E.S., 1995. The fossil record of African antelopes (Mammalia, Bovidae) in relation to human evolution and paleoclimate.

Walker, P.A. and Cocks, K.D., 1991. HABITAT: a procedure for modelling a disjoint environmental envelope for a plant or animal species. *Global Ecology and Biogeography Letters*, pp.108-118.

Walstrom, V., Klicka, J. and Spellman, G.M., 2012. Speciation in the White-breasted Nuthatch (Sitta carolinensis): a multilocus perspective. *Molecular Ecology*, *21*(4), pp.907-920.

Walther, G.R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T.J., Fromentin, J.M., Hoegh-Guldberg, O. and Bairlein, F., 2002. Ecological responses to recent climate change. *Nature*, *416*(6879), p.389.

Warren, D.L. and Seifert, S.N., 2011. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*, *21*(2), pp.335-342.

Wieczorek, J., Guo, Q. and Hijmans, R., 2004. The point-radius method for georeferencing locality descriptions and calculating associated uncertainty. *International journal of geographical information science*, *18*(8), pp.745-767.

Wiens, J.J., Ackerly, D.D., Allen, A.P., Anacker, B.L., Buckley, L.B., Cornell, H.V., Damschen, E.I., Jonathan Davies, T., Grytnes, J.A., Harrison, S.P. and Hawkins, B.A., 2010. Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology letters*, *13*(10), pp.1310-1324.

Williams, M.A., Dunkerley, D.L., De Deckker, P., Kershaw, A.P. and Stokes, T., 1993. *Quaternary environments*. E. Arnold.

Willis, K.J., Bailey, R.M., Bhagwat, S.A. and Birks, H.J.B., 2010. Biodiversity baselines, thresholds and resilience: testing predictions and assumptions using palaeoecological data. *Trends in Ecology & Evolution*, *25*(10), pp.583-591.

Wilson, C.D., Roberts, D. and Reid, N., 2011. Applying species distribution modelling to identify areas of high conservation value for endangered species: A case study using Margaritifera margaritifera (L.). *Biological Conservation*, *144*(2), pp.821-829.

Wisz, M.S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H. and Guisan, A., 2008. Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, *14*(5), pp.763-773.

Yackulic, C.B., Chandler, R., Zipkin, E.F., Royle, J.A., Nichols, J.D., Campbell Grant, E.H. and Veran, S., 2013. Presence-only modelling using MAXENT: when can we trust the inferences?. *Methods in Ecology and Evolution*, *4*(3), pp.236-243.

Yamaguchi, R. and Iwasa, Y., 2013. First passage time to allopatric speciation. *Interface focus*, *3*(6), p.20130026.

Zamudio, K.R. and Savage, W.K., 2003. Historical isolation, range expansion, and secondary contact of two highly divergent mitochondrial lineages in spotted salamanders (Ambystoma maculatum). *Evolution*, *57*(7), pp.1631-1652.

Zaniewski, A.E., Lehmann, A. and Overton, J.M., 2002. Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. *Ecological modelling*, *157*(2-3), pp.261-280.