# CHAPTER 7

# Predictive species distribution models disagree in ecological and environmental transition zones

# Predictive species distribution models disagree in ecological and environmental transition zones

B.F.N. Erasmus[1], B.J. van Rensburg[1], A.S. van Jaarsveld[2*], T. Sutton[3], A. Turner[4], D. Stockwell[5] and S.L. Chown[6]

[1]Department of Zoology and Entomology, University of Pretoria, Pretoria 0002, South Africa.
[2]Department of Zoology, University of Stellenbosch, Private Bag X1, Matieland 7602, South Africa
[3]Centre for Plant Diversity and Systematics, School of Plant Sciences, The University of Reading, Reading, RG66AS, UK
[4]Cape Nature Conservation, Scientific Services, Private Bag X5014, Stellenbosch, 7599
[5]NPACI/SDSC, MC 050, 9500 Gilman Dr, Bldg 109, La Jolla, CA 92093-0505
[6]Spatial, Physiological and Conservation Ecology Group, Department of Zoology, University of Stellenbosch, Private Bag X1, Matieland 7602, South Africa

* Corresponding author

For submission to: *Ecological Modelling*

## Abstract

As predictive distribution modelling has become more common in response to a need to understand potential effects of environmental change, there has been a concomitant increase in the diversity of models used for such predictive modelling. This study contributes to this body of work by comparing 3 different predictive distribution models: a climate envelope model, BIOCLIM and GARP. Comparisons were done by using a standard set of climatic predictor variables for a set selection of species. Receiver operating characteristic analysis showed good overall model accuracy and little difference in the mean output between models. However, by mapping the distributions of individual species for which models disagree, it was found that although the models perform adequately, they disagreed on the predicted distributions of individual species in the same geographically distinct region. This region was shown to be an area of significant environmental heterogeneity, with climate, vegetation and avian community transitions occurring here. Identifying such areas prior to a distribution modelling exercise and explicitly accounting for this source of model disagreement, will improve the accuracy of model predictions. The methods followed in this study present a simple analysis to identify such transition zones.

## Introduction

The number of investigations using predictive species distribution modelling is increasing rapidly, largely because these models are of considerable value for understanding the likely impacts of environmental change. The models can be used to understand biogeographic patterns (Jeffree and Jeffree 1994, Cowling and Lombard 2002), ascertain the potential distributions of economically important pest species (Bylund 1999, Robertson et al. 2001, Todd et al. 2002, Watt and McFarlane 2002), determine changes in threats associated with infectious and vector-borne diseases (Daszak et al. 2000, Rogers and Randolph 2000, Harvell et al. 2002), predict potential species invasions (Peterson and Vieglais 2001), and understand the ecological and conservation implications of global environmental change ( Pounds et al. 1997, Petchey et al. 1999, Rutherford et al. 1999, Beaumont and Hughes 2002, Feria and Peterson 2002, Hannah et al. 2002, Peterson et al. 2000, 2001, 2002).

By definition, all models, including predictive species distribution models, constitute an incomplete abstraction of reality (see Samways 2003). In consequence, they will fail to capture all sources of variation, whether these are a result of predictor variable or model characteristics. The extent to which this abstraction of reality is useful for understanding natural phenomena is determined by the model's ability to describe variation in the observed data, using predictor variables selected *a priori* by the modeller (Sutherst 2003). Different types of models differ in their ability to describe this variation.

Climate envelope models have their roots in the concept of "envirograms" first put forward by Andrewartha and Birch (1954, 1986). Since the first generation climate envelope models were developed as specific software applications (Sutherst and Maywald 1985), the approach has found wide application (Walker and Cocks 1991, Carpenter et al. 1993, Jeffree and Jeffree 1994, 1996, Robertson et al. 2001, Erasmus et al. 2000, 2002, Van Staden et al. in press). Standardising the approach into readily available software packages (BIOCLIM[1] and CLIMEX[2]) has made it more accessible and fostered a wider user base, not only for these packages (Busby 1986, 1988, 1991, Nix 1986, Lindenmayer et al. 1991, Beaumont and

---

[1] http://cres.anu.edu.au/outputs/anuclim.html
[2] http://www.ento.csiro.au/climex/climex.htm

Hughes 2002)[3], but also for climate envelope models in general. On the other hand, the increase in computing power in standard desktop computers has made it practical to utilise computing-intensive machine learning type models for distribution predictions (Guegan et al. 1998, Kobler and Adamic 2000). Until fairly recently, this complicated methodology was not readily accessible to biologists. A genetic algorithm-based model developed specifically for species distribution prediction from incomplete data (GARP, Stockwell and Peters (1999)) was made available first as a web application[4] and then later as a stand-alone Windows application[5]. An increase in the number of GARP applications (Feria and Peterson 2002, Peterson and Cohoon 1999, Peterson et al. 2000, 2001, 2002, Peterson and Vieglais 2001,) can be expected, given the fact that it is user friendly, easily accessible, well supported, and places a powerful, well-documented computational technique at the disposal of biologists. Consequently, a comparison between the more established climate envelope models that already have a broad user base, represented by BIOCLIM and the simple PCA-based model developed by Jeffries (Jeffree and Jeffree 1994, 1996), and modified by Erasmus (Erasmus et al. 2000), and GARP, which is likely to become more widely used in future, would be useful to ascertain the extent to which these approaches are comparable. Given that these kinds of modelling approaches provide only a broad-brush view of the likely responses of species to climate, strong model disagreement would strengthen the case against them (see Davis et al. 1998, Gaston 2003, Sutherst 2003 for additional discussion). On the other hand, agreement between models would provide additional support for their continued use. Moreover, explicit investigation of lack of congruence between models would improve confidence in the utility of modelling exercises. The aims of this study are therefore to compare the ability of different predictive distribution models to describe observed variation given a fixed set of predictor variables for a selected number of species, and to investigate the spatial attributes of predictor variables that cause between-model disagreement of predicted species distributions. In so doing, this work contributes to the useful and growing body of model evaluations (Fielding and Bell 1997, Manel et al. 2001) and model comparisons (Brito et al. 1999, Manel et al. 1999, Guisan and Zimmermann 2000).

---

[3] See http://www.ento.csiro.au/climex/bibliography.htm for a list of 148 references on the implementation of the original climate envelope model called CLIMEX by Sutherst and Maywald (1985)

[4] http://biodi.sdsc.edu/bsw_home.html

[5] http://www.lifemapper.org/desktopgarp/

**Methods**

*Data*

Distribution data for birds were obtained from the Avian Demography Unit, University of Cape Town[6] (Southern African Bird Atlas Project (SABAP, Harrison et al. (1997)). These data are the result of a directed sampling effort to map the birds of southern Africa (1987-1992), and unlike data for other taxa from the region (Scholtz and Chown 1995; van Jaarsveld et al.1998a,b, Koch et al. 2000), constitute a close approximation of true presence-absence data. From the total data set for southern African birds, a data subset for grid cells with reliable climate data for South Africa was extracted. This subset consisted of 294816 unique records for 748 bird species occurring in 1858 grid cells (~625km$^2$ per grid cell), encompassing South Africa and Lesotho, but excluding the other countries included in the SABAP. For a previous predictive distribution modelling study (see Erasmus et al. 2002 for a species list), 34 species were selected from this South African data subset based on (a) accurate but not necessarily complete distribution data, (b) representative geographic range types (e.g. species restricted to winter or to summer rainfall regions), (c) robust, well-resolved taxonomy and (d) species known from a reasonable number of records. We used the same 34 species in this study. These 34 species have 23047 records in total, and their summed distributions encompass all 1858 grid cells. The number of records for each of these 34 species varies from 49 to 1574, with half the species having distributions of less than a third of the region's total area.

Van Rensburg et al. (2002) showed that mean annual precipitation (mm.yr$^{-1}$) and mean absolute monthly minimum temperature ($^0$C) averaged over the year are significant positive correlates of avian species richness in South Africa (see also Dean 2000). These results for South Africa are supported by similar findings for other taxa in North America (Currie and Paquin 1987, Boone and Krohn 2000a), New Zealand (Leathwick et al. 1998), Costa Rica (Enquist 2002) and sub-Saharan Africa (Jetz and Rahbek 2002). Based on these known relationships between richness, temperature and precipitation at broad scales, we made use of the same mean annual precipitation and mean absolute monthly minimum and maximum temperature data used by Van Rensburg et al. (2002) as predictor variables in this study. South African climate data were provided by the Computing Centre for Water Research

---

[6] http://www.uct.ac.za/depts/stats/adu/index.html

(CCWR, University of Natal, Pietermaritzburg, metadata published in Schulze and Maharaj (1997)). The historic data (30 year monthly and annual means: 1960 – 1990) were re-sampled to 15' x 15' grid cells from interpolated climate surfaces available at a minute by minute resolution to conform to the resolution of the bird data.

*Modelling procedure*

Using these bird distribution and climate data, we implemented three different species distribution prediction models that have been exposed to peer review: Jeffree's multivariate climate envelope model (Jeffree and Jeffree 1994, 1996, Erasmus et al. 2000, 2002, Olwoch et al. in press, Van Staden et al. in press, also see Robertson et al. 2001 for a similar PCA-based model), the BIOCLIM envelope model (Busby 1986, 1988, 1991, Nix 1986, Lindenmayer et al. 1991, Beaumont and Hughes 2002) and GARP, an inferential genetic algorithm-based model (Peterson et al. 2000, 2001, 2002, Peterson and Cohoon 1999, Peterson and Vieglais 2001, Feria and Peterson 2002). GARP has different output options; we used the probability surface output option to enable comparisons with the probability surfaces generated by the other two envelope models.

Model output was evaluated (Oreskes et al. 1994, Guisan and Zimmerman 2000) by comparing predicted distributions with known records. However, model output consists of a probability of occurrence value and to compare this continuous variable with a binary presence-absence value presents a methodological problem. Although the predicted probability output can be dichotomised to predicted presence-absence by applying a threshold probability, this procedure results in a significant loss of information (Deleo 1993, Fielding and Bell 1997). For example, if a threshold probability of 0.5 is chosen, this means that species having a predicted probability of occurrence of 0.51 will be regarded as present together with species that have a much higher predicted probability of occurrence. However, receiver operating characteristic (ROC) analysis enables the use of the actual predicted probabilities, and is threshold-independent. ROC analysis has its origins in engineering in the context of measuring the ability of a detector to detect a particular signal. It has found wide application in the evaluation of clinical medical tests (Fielding and Bell 1997) and here it is used to measure the performance of a predictive species distribution model. ROC analysis primarily concerns the calculation of specificity and sensitivity values. Sensitivity is defined as the number of true occurrences of a species (true positive predictions, TP) divided by the total number of positive predictions, whether true (TP) or not (false negatives predictions,

FN). Similarly, specificity is defined as the number of true negative predictions (TN) divided by the total number of negative predictions, whether TN or false positive (FP) predictions (Fielding and Bell 1997). The area-under-curve (AUC) of a plot of 1-specificity against sensitivity at every given probability of occurrence value is used as a test statistic. An AUC value of 0.5 corresponds to a random test with no discriminatory ability. AccuROC® software (Vida 1993) was used to determine the statistical significance of the difference between any two or three AUC values according to the nonparametric method of Delong et al. (1988). The confidence interval for the AUC was calculated using the asymptotic method (Obuchowski and Lieber 1998).

The chosen model evaluation criteria should be prevalence (frequency of occurrence) independent across models (Manel et al. 2001). Following suggestions from previous model comparisons (Manel et al. 1999, 2001, Guisan and Zimmerman 2000), we tested ROC analysis (Fielding and Bell 1997, Deleo 1993, Obuchowski and Lieber 1998) for prevalence independence by performing a linear least squares regression of AUC values on prevalence.

Kraemer's (1988) critique of the use of ROC in biology is limited to medical and behavioural contexts where a human observer applies a procedure under evaluation to detect the presence of a signal. This signal is typically the presence of a disease or a particular animal behaviour. Kraemer (1988) regarded such a human observer as subjective and prone to be influenced by preconceived ideas. In the classical engineering application of ROC analysis, such a bias does not exist for a mechanical or electronic detector (Kraemer 1988). Likewise, in predictive distribution modelling, the predicted probability of occurrence is a value determined by a statistical model that is not subject to observer bias. Indeed, the output value of this model is analogous to the reading an electronic detector would record for a particular electronic signal. Although the model can have biases in terms of predictor variables and modelling assumptions, these are explicit, constant and objective. Erroneous species identification may also lead to a false evaluation of model performance, but the bird data set has been subjected to extensive expert vouching procedures and as such presents as close an approximation of objective observations and accurate identifications as can realistically be expected (Harrison et al. 1997, Van Jaarsveld et al. 1998a). In summary, the specific application of ROC analysis to predictive distribution modelling evaluation is much closer to the original engineering application than Kraemer's (1988) "biobehavioural context", and therefore we regard classical ROC analyses as suitable for our purposes of model comparison.

Model comparisons proceeded in three phases. In the first phase, we compared the output of each model for each species to a random model with no discriminatory ability. Phase two consisted of obtaining a mean AUC value for all 34 species, for each model, and comparing these three values using a nonparametric Kruskal-Wallis test. Phase three compared models on a species-by-species basis and as such provides more detail on which models disagree for which species. The procedure followed in phase three is summarised in Figure 1, and consists of the following steps:

*Step 1*: Three species-by-species pairwise between-model comparisons were undertaken (e.g. Jeffree-BIOCLIM (J-B), Jeffree-GARP (J-G), GARP-BIOCLIM (G-B)), as well as a species-by-species three-way between-model comparison (Jeffree-GARP-BIOCLIM (JGB)), resulting in a total of four model comparisons. Each of these four comparisons was done for every species, summarised in the table shown for step 1 in Figure 1. Each comparison resulted in a list of species for which that particular comparison showed significant differences between the models that were compared.
Steps 2 to 4 were subsequently repeated for each of these species lists.

*Step 2*: The original presence-absence distribution data for every species on the lists (i.e. those species for which model outputs differed) generated in Step 1 were drawn from the database, and all individual presence-absence maps for species were overlaid in a GIS.

*Step 3:* Every grid cell was queried for the presence or absence of a species, and the number of species from the lists that occurred in each of the grid cells was recorded.
The end result of this process was a map where each grid cell represented the number of species for which the models in question gave significantly different model outputs ($p < 0.05$). Grid cells that contained a large number of species for which model outputs disagreed, were referred to as cells rich in model disagreement species.

*Step 4:* Homogenous spatial clusters of grid cells that contain similar numbers of model disagreement species are identified by calculating Moran's I values as a local indicator of spatial association (LISA) (Anselin 1995). These values were calculated using the SPLUS® for ArcView extension to ArcView® GIS. The module calculates generalised LISA values
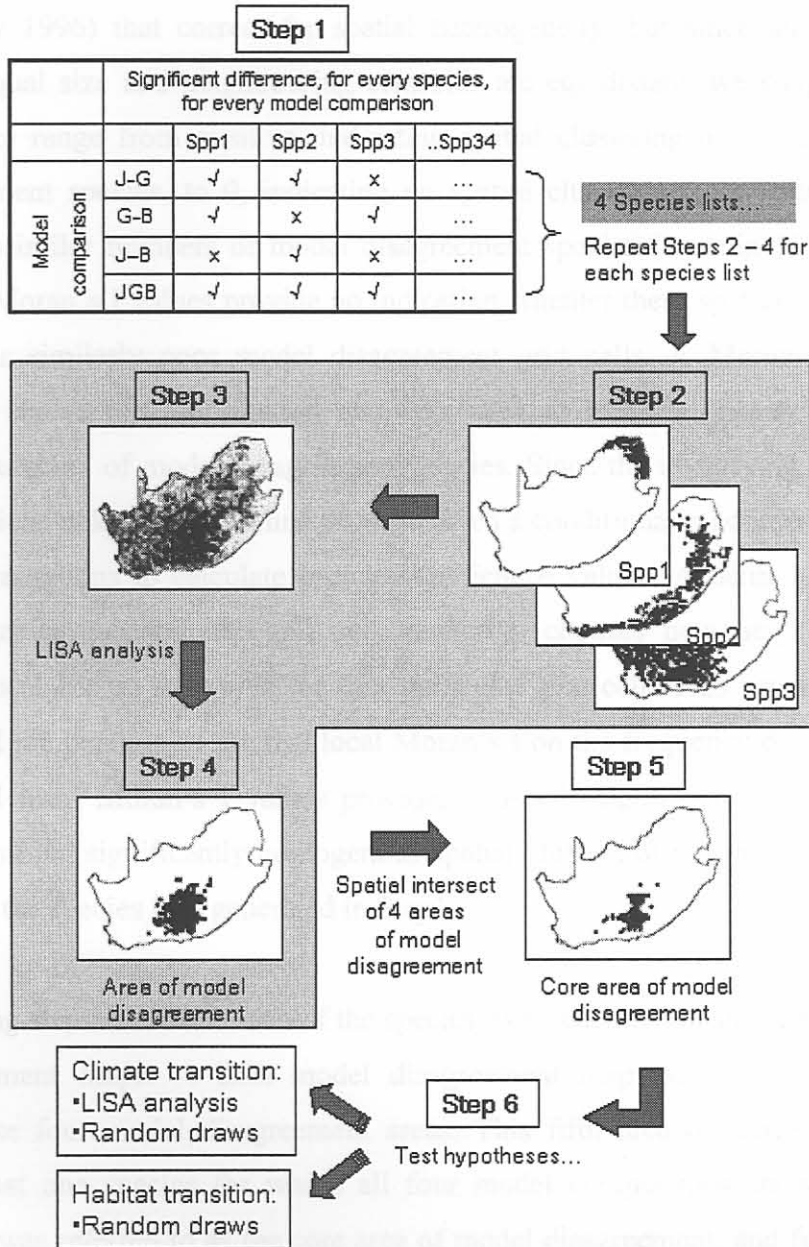
142

**Fig 1:** Flow chart of the methods followed. Significant disagreement between the predictive distribution models for a particular species is shown in the table under step 1. For example, species 1 and 2, but not 3, would be members of the species list for model comparison J-G in figure 1. These four generated lists contained the species for which model outputs differed significantly. Each of these lists are used as input to first generate species richness maps, where richness refers to the number of model disagreement species occurring in each grid cell (steps 2 and 3). Step 4 calculates significantly homogenous areas of model disagreement by LISA analysis, and then a spatial intersect between the 4 LISA maps from step 4 creates the map of the core area of model disagreement in step 5. Two hypotheses about this core area of model disagreement are tested in step 6.

(Bao and Henry 1996) that correct for spatial heterogeneity, but since all grid cells are effectively of equal size and neighbouring centroids are equidistant, we simply use LISA. Moran's I values range from positive, indicating spatial clustering of similar numbers of model disagreement species, to 0, indicating no spatial clustering, to negative, indicating clustering of dissimilar numbers of model disagreement species (Anselin 1995). However, highly positive Moran's I values provide no indication whether these spatial clusters contain similarly rich or similarly poor model disagreement grid cells. A Moran's I scatterplot (Anselin 1998) shows this information and was used to identify clusters of cells with similarly high numbers of model disagreement species. Since the underlying distribution of this LISA statistic is unknown (Anselin 1995), we used a conditional randomisation approach with 10000 permutations to calculate pseudo-significance values (Anselin 1995). In short, this approach selects a particular grid cell, randomly chooses new neighbours and re-calculates the local Moran's I value for that particular grid cell. This process is repeated 10000 times and the position of the real local Moran's I on the frequency distribution of the 10000 permuted local Moran's I values provides a pseudo-significance value. Using this method, we identified significantly homogenous spatial clusters of rich model disagreement cells for each of the species lists generated in Step1.

**Step 5:** Repeating steps 2 to 4 for each of the species lists generated in step 1 resulted in four model disagreement maps. A fifth model disagreement map was created by spatially intersecting these four model disagreement areas. This fifth area of model disagreement contained at least one species for which all four model comparisons showed significant differences and was referred to as the core area of model disagreement, and further analyses were focussed on this particular area.

**Step 6:** We tested two hypotheses relating to this core area of model disagreement: first, that it is an area of climate transition, and secondly, that it is an area of habitat transition.
For the climate transition hypothesis, spatial variation of the climatic variables across the entire study area as well as in the core area of model disagreement was investigated. Mean annual precipitation is a measure of the expected quantity of water available to a catchment in a decadal to centennial time scale (Schulze et al. 2001). Potential evapotranspiration provides a measure of how easily water could potentially be lost to the atmosphere through evaporation and transpiration (Currie 1991, Hulme 1996). Therefore, the relationship between precipitation and potential evapotranspiration gives a broad indication of whether a

system is run-off or evaporation dominated, and as such summarises the water balance in a system (Knapp and Smith 2001). We used this single climate variable, the precipitation/evapotranspiration ratio (PPT/PET), as a descriptor of the general environmental constraints on water availability, and indirectly, species distributions. Calculating LISA values and corresponding pseudo-significance values for the PPT/PET ratio, we identified areas with high and low local spatial autocorrelation of their PPT/PET ratios. The geographic positions of homogenous spatial clusters of PPT/PET ratios indicate areas that are subject to similar water balance conditions. The boundaries between these homogeneous clusters provide an indication of areas that can be interpreted as transitional zones between low and high PPT/PET ratios.

The climate transition hypothesis was further tested by determining whether the core area of model disagreement showed greater values of the PPT/PET ratio and local Moran's I of the PPT/PET ratio than expected by chance. The mean values for the PPT/PET ratio and local Moran's I of the PPT/PET ratio in the core area of model disagreement were compared to the mean PPT/PET ratio and mean local Moran's I of the PPT/PET ratio for 10 000 sets of randomly selected grid cells from across the study region. The number of randomly selected grid cells was equivalent to the number of grid cells in the core area of model disagreement.

The habitat transition zone hypothesis was tested using 10 000 random draws in a similar fashion, using biome heterogeneity as a variable. Low and Rebelo's (1996) classification of the study area into seven biome types was used to calculate the percentage of each biome type per quarter-degree cell. These were the forest, thicket, savanna, grassland, Nama karoo, succulent karoo, and fynbos biomes. Each of these is characterised by several vegetation types, giving a total of 68 different vegetation types for the study area (see Low and Rebelo, 1996 for further information). Following Gaston et al. (2001), biome heterogeneity was obtained using Simpson's index of diversity (Krebs 1989):

$$1 - \text{Sum } (p_i)^2$$

where $p_i$ is the fraction of the grid cell's area occupied by biome $i$. This index ranges from zero (only one biome present in a cell) to 0.86 (all seven biomes present in the same proportions) and is expected to reach high values in areas of transition between biomes. The values recorded ranged between 0 and 0.76.

**Results**

In contrast to Manel et al.'s (2001) study, we found a significant (p<0.001), but weak negative relationship for the linear regressions of AUC on prevalence for each of the three models that we employed. (Jeffree model: adjusted $R^2$=0.28587591, $F_{(1,32)}$=14.210 p<0.00067; GARP: adjusted $R^2$=0.28261882, $F_{(1,32)}$=14.001 p<0.00072; BIOCLIM: adjusted $R^2$=0.17839935, $F_{(1,32)}$=8.1655 p<0.00745) This relationship between prevalence and AUC values was constant across all three models – there was no significant difference between the intercepts or slopes of each of the regression lines of AUC on prevalence across the predictive models (ANCOVA, Zar 1986) (Intercept: $F_{(2,98)}$=1.06; p<.3509, slope: $F_{(2,96)}$ = 0.007; p<.9932).

The first phase of model comparison showed that all three models performed significantly better than a random model. In the second phase of model comparison, the mean AUC value across species for each of the three models was not significantly different from the mean AUC value for any of the other models (AUC values Jeffree: 0.843± 0.084, GARP: 0.834±0.082, BIOCLIM: 0.816±0.105).

Results from the third phase of model comparison, which was conducted on a species-by-species basis, showed that there were several significant differences in model outputs (p < 0.05) for a number of species. In total , 28 (82%) of the species produced modelled outputs that differed significantly across all three models. Fourteen (41%) species were common to all four lists of species for which model outputs differed significantly. From this it appears that there are more species from this sample for which the comparative model outputs disagree than ones about which they agree.

The core area of model disagreement derived in step 5 of Figure 1 is presented as the hatched area in Figure 2. This core area of model disagreement showed a proportional overlap (Prendergast et al. 1993) with the other four areas of model disagreement ranging from 44% to 59%.

Testing the climate transition zone hypothesis, LISA analyses of PPT/PET ratios across the study area showed two highly significant (p<0.001) spatial clusters: a cluster of low PPT/PET
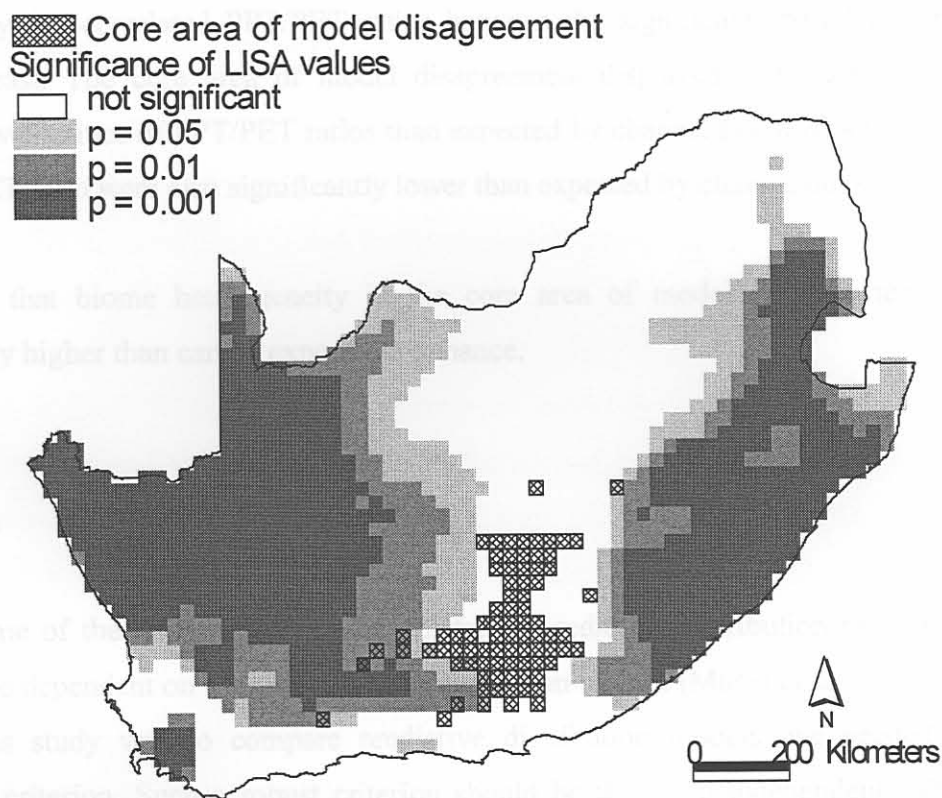
146

**Fig 2:** Significant LISA values for the PPT/PET ratio over the study area. Hatched cells represent the core area of model disagreement.

values in the west and a cluster of high PPT/PET values in the east (Fig. 2). Figure 2 also shows the geographical position of the core area of model disagreement relative to these two PPT/PET spatial clusters. The core area of model disagreement is spatially congruent (proportional overlap value of 83% - Prendergast et al. 1993) with a narrow strip of non-significantly autocorrelated PPT/PET ratios between the significant spatial clusters in the west and east. The core area of model disagreement displayed significantly (two-tailed $p<0.05$) lower values of PPT/PET ratios than expected by chance. Local Moran's I values of the PPT/PET ratio were also significantly lower than expected by chance, but still positive.

We found that biome heterogeneity in the core area of model disagreement was also significantly higher than can be expected by chance.

## Discussion

The outcome of the comparison between different predictive distribution models has been shown to be dependent on the chosen model evaluation criteria (Manel et al. 2001). Since the aim of this study was to compare predictive distribution models, we needed a robust evaluation criterion. Such a robust criterion should be threshold independent, which ROC analysis is by definition, and prevalence independent (Manel et al. 2001). The literature shows two approaches to assess prevalence independence: a regression of the AUC values from ROC analysis on the prevalences for each species in the dataset, so that the number of data points is equal to the number of species in the data set (the approach followed here, and by Manel et al. 2001); or using random subsets of different sizes of the data for an individual species and regressing those AUC values on the prevalences determined by the size of the subsets.

The AUC values in the first approach are derived from different known species distributions with different climate envelopes. These known species distributions have different prevalences and different spatial configurations. Therefore, differences in AUC values derived in this fashion reflect not only the effects of prevalence, but also the model's ability to predict species distributions for these different climate envelopes. This approach is well-suited to evaluate model performance over a range of different types of distributions and get an overall view of model performance. Manel et al. (2001) used ROC analysis to identify

optimal probability thresholds, and therefore this approach was well suited to that study. No regression diagnostics were provided with which to compare the regression results from the current study and we can only conclude that for our study, the relationship between AUC and prevalence is weak when using data from all species. However, the main aim of the current study was not to provide an overall view of model performance, but to compare model outputs for individual species. Such comparisons were done with a fixed set of predictor variables applied by different models to the distribution of a specific species. Therefore, even though there is a relationship between prevalence and AUC values, this relationship is weak and constant for all three models and therefore its influence, if any, will be constant across model comparisons.

Further support for ROC analysis as an appropriate model evaluation criterion for this study comes from other studies that have followed the second approach to assess the prevalence independence of AUC values. Cumming (2000) used different sized subsets of known data for a tick species to derive an AUC-prevalence graph and found that the AUC values were prevalence independent, all be it without regression diagnostics. Since our study focuses on model comparisons for individual species, rather than overall model performance for an assemblage of species, it is closer in scope to Cumming (2000) than Manel et al. (2001). Cumming (2000) also investigated the effects of hypothetically small prevalences on AUC values and found that AUC values are more prevalence dependent at low prevalences. Using the same data as the current study, Erasmus et al. (2002) investigated model performance for individual species at low prevalences and found that, for the majority of species studied, there was no significant difference between models based on a 20% random subsample of the known distribution and models based on 100% of the known distribution. In short, we maintain that ROC analysis is an appropriate model evaluation criterion for our purposes of comparing different predictive distribution models, not only for the reasons provided above, but also because of its broad support in the literature (Fielding and Bell 1997).

The first two phases of model comparison showed little differences between models, irrespective of whether the comparison was with a random model or with another model. Using other model evaluation criteria, Manel et al. (1999) compared a discriminant analysis, a neural network and a logistic regression model and also found little differences in model performance. These similar results for studies using different models with different predictor variables, suggest that, despite the differences in the underlying assumptions of the various

models, all of the models have an ability to extract at least partial environmental dependencies of species distributions and facilitate meaningful predictions. Therefore, these results suggest that model performance is more strongly influenced by predictor variable selection than by model selection.

The core area of model disagreement identified in phase three of the model comparison process showed the geographic location of an area with a high number of model disagreement species, i.e. the three model outputs, although reasonably accurate by themselves, differed from each other for a high number of species. The fact that this area was a discrete homogenous spatial cluster, and not a large extensive area, raises a question about the cause of such model disagreement in an area that comprises only 4% of the country's total surface area. Although the different models use predictor variables in different ways to arrive at a predicted distribution, the common factor between all models remains their use of climatic predictor variables. If model disagreement shows a common spatial clustering, it stands to reason that some climate-related process drives this pattern.

LISA analysis of PPT/PET values identified a transition zone of environmental water surplus, i.e. an area with no significant local spatial autocorrelation in PPT/PET ratios. This narrow strip where significant spatial autocorrelation of the PPT/PET ratio is absent, represents a transition zone from areas where evaporation plays a more important role, such as the drier western regions of the country, to areas where run-off is more prevalent, e.g. the moister eastern regions. The interpretation of PPT/PET ratio is relative. Knapp and Smith (2001) showed grasslands to have PPT/PET ratios of around unity, with deserts closer to zero and forests larger than unity in North America. For South Africa, some 93% of grid cells experience a PPT/PET ratio of smaller than 0.5, therefore South Africa generally experiences an evaporation-dominated water balance, and that the transition zone identified by the LISA analysis is a relative one with run-off only dominating evapotranspiration for short and strictly seasonal periods across the eastern half of the country. Note that it is only where there is a steep transition gradient that model predictions disagree; the core area of model disagreement does not extend to the north where the PPT/PET ratio transition is diluted over a larger area. Therefore, by virtue of the spatial congruence between the core area of model disagreement and the steep PPT/PET transition zone, this study suggests that the distribution prediction models implemented here differed in their predictions of species that are found in this climate transition zone. Figure 3 shows the PPT/PET values for the
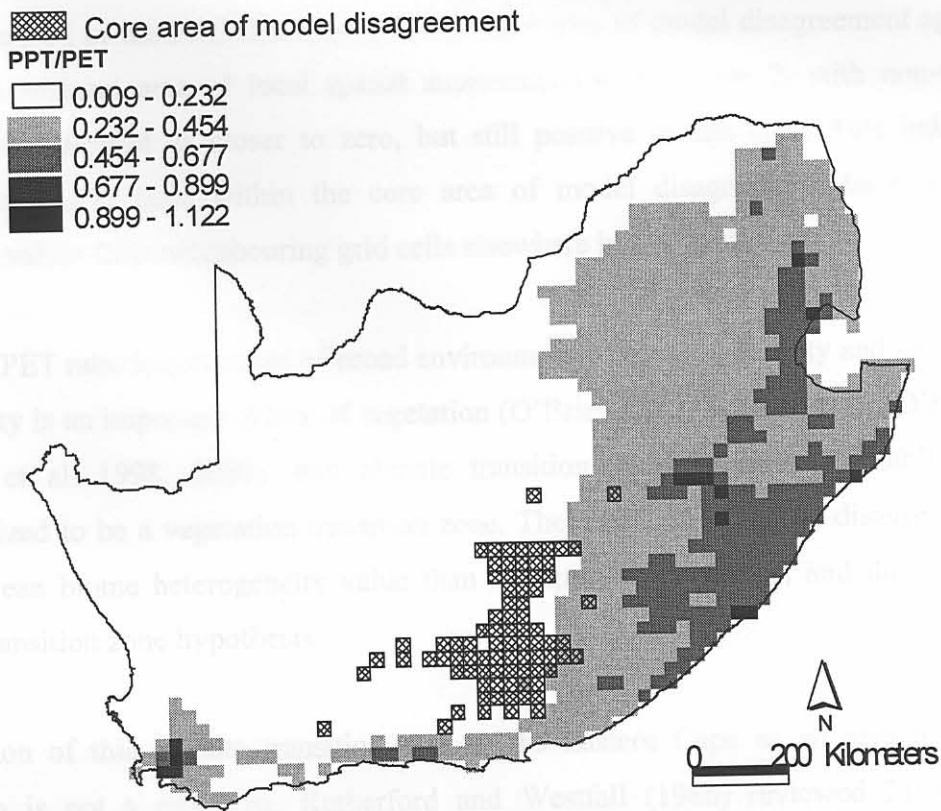
**Fig 3:** Map of PPT/PET values for South Africa. Legend classes are of equal interval, so areas with similar shading indicate the frequency of values in that class interval.

study area with equal interval classes. The core area of disagreement coincided almost totally with an area that has low values of PPT/PET. This area was also the largest (white area in Figure 3) and therefore it is to be expected that PPT/PET ratios will be lower than can be expected by chance in the core area of model disagreement. The significantly lower mean local Moran's I of the PPT/PET values for the core area of model disagreement agreed with the non-significant area of local spatial autocorrelation in Figure 2, with non-significant values being defined as closer to zero, but still positive in this case. This indicates that neighbouring grid cells within the core area of model disagreement have less similar PPT/PET values than neighbouring grid cells elsewhere in the country.

The PPT/PET ratio is a measure of broad environmental water availability and since moisture availability is an important driver of vegetation (O'Brien 1993, Schulze 1997, O'Brien 1998, O'Brien et al. 1998, 2000), this climate transition zone of water availability can be hypothesized to be a vegetation transition zone. The core area of model disagreement has a higher mean biome heterogeneity value than is expected by chance, and this supports the habitat transition zone hypothesis.

The notion of this climate transition area in the Eastern Cape as an area of vegetation transition is not a new one. Rutherford and Westfall (1986) reviewed 21 studies that identified "major natural biotic divisions" (Rutherford and Westfall 1986) between 1936 and 1986. These classifications were by no means transparent and used combinations of edaphic factors, plant life forms, climate and expert opinion to do the classification. In spite of these differences in methods, the area that we identified as the core area of model disagreement coincides with boundaries between four biomes in almost half of these old classifications. An earlier review of phytogeographical studies (Werger 1978) found a similar result of agreement between 9 studies dating from between 1886 and 1971. A more rigorous definition (Rutherford and Westfall 1986) of South African biomes, primarily based on climate and plant life forms, lends further support for the core area of model disagreement as a vegetation, and habitat, transition zone. The climate component of this classification was based on a Summer Aridity Index (SAI), calculated from precipitation for the four hottest months of the year, and winter concentration of precipitation, calculated from winter precipitation as a proportion of mean annual precipitation. SAI gives an indication of moisture stress during growing periods of peak physiological water demand (Rutherford and Westfall 1986) and winter precipitation concentration provide an indication of the importance

of seasonal rainfall. Together, these precipitation-derived indices give a similar, but more detailed picture of environmental water availability than the PPT/PET ratio used in our study. On a scatterplot of winter precipitation concentration against SAI, Rutherford and Westfall (1986) identified biomes and biome boundaries. Using Acocks' veldtypes, (Acocks 1975), we found that four veld types dominated in the core area of model disagreement, accounting for 74% of this area. These veld types were located along the common boundaries of the nama karoo, grassland, fynbos, forest and savanna biomes on the scatterplot of SAI against winter precipitation concentration. These principles that aided in the identification of Rutherford and Westfall's (1986) biomes provide further evidence that the core area of model disagreement is an area of climate, and subsequently, habitat transition.

An important feature of Rutherford and Westfall's (1986) classification is the role the inclusion of a measure of rainfall seasonality plays in delineating biomes. Low and Rebelo (1996) identified a thicket biome as intermediary between forest and savanna biomes (Vlok and Euston-Brown 2002). This thicket biome, which is present in 28% of the core area of model disagreement's grid cells, is maintained by the balance between winter and summer rainfall. Winter rainfall is more dominant in the west, and here the thicket becomes fragmented and is replaced by fynbos, whereas more summer-dominated rainfall fragments the thicket in the north and east. It is clear then, that the vegetation transition in the core area of model disagreement is not only influenced by the amount of precipitation but also the seasonality.  This balance between winter and summer rainfall also explains to some extent the more gradual vegetation transition that takes place north of the core area of model disagreement. Winter rainfall is limited to the coastal area, and therefore the climate gradient further north is only a gradient of the amount of precipitation, and not seasonality too, as is the case along the coastal area.  The complex vegetation patterns in this core area of model disagreement also have some roots in the distant past. Van Zinderen Bakker (1978) found that Quaternary glacial cycles enabled repeated intrusion and retreat of fynbos into temperate grassland and forest areas, resulting in a mosaic of fynbos remnants in altitudinal refugia during unfavourable periods.

So far we have shown that the areas in which the models disagreed were areas of climate and habitat transition. However, it has been demonstrated elsewhere that habitat heterogeneity is an important driver for bird species richness not only in sub-Saharan Africa (Jetz and Rahbek 2002), but also for South Africa (Van Rensburg et al. 2002). Spatial congruence between

avian species range limits and areas with high vegetation heterogeneity have also been found for other regions (e.g. see Boone and Krohn 2000b). More importantly, in the same study area, Gaston et al. (2001) have shown that biome edges experience a significantly higher degree of avian beta diversity than expected by chance. Therefore it is clear that the areas in which the three predictive distribution models disagree are not only areas of climate and habitat transition, but also very likely form areas of avian community transition. At broader scales, using 1437 Afrotropical endemic bird species, De Klerk et al. (2002) showed that 4 biogeographical districts share boundaries in the area that we have identified as the core area of model disagreement.

Due to the unique biodiversity elements, i.e. the thicket biome, in this ecological transition zone, considerable effort has been directed at integrated conservation planning in this region (Cowling et al. 2003, Kerley et al. 1999). An integral part of this conservation plan is the integration of process that drive biodiversity patterns in this area; macroclimatic gradients has been explicitly identified as important drivers for maintaining biodiversity in the thicket biome (Cowling et al. 2003).

In summary, we have shown that there is little difference in model output for the three different predictive distribution models employed here. We have also shown that these models fail in the same ecological transition zone by having different predicted distributions for species occurring here. Although this disagreement seems to point to differences between models, the fact that there is agreement about where the disagreement occurs is noteworthy and supports the usefulness of a modelling approach to species distributions. We suggest that such general ecological transition zones should be sought in environmentally complex landscapes before blanket analyses (i.e. Erasmus et al. 2002, Van Jaarsveld and Chown 2001) are applied. Understanding the potential effects of these areas, and how to deal with them explicitly is likely to be important in efforts to improve model performance. Given the rate of environmental change and the need for up-to-date conservation planning, the use of predictive distribution models is likely to increase. It is in the interests of climate-change integrated conservation planning that such modelling exercises are as accurate and transparent as possible and therefore, procedures that identify sources of model error are important. Further research is needed not only into how existing conventional equilibrium (Guisan and Zimmerman 2000) models treat such transition zones, but also how these transition zones should be treated to ensure geographically homogenous model performance.

Such a model performance might only be attained if the model allows for predictor variables to make a spatially variable contribution to prediction accuracy. Models that have a spatially explicit component (i.e. spatial regression, cellular automata) might perform better in areas where predictor variables undergo transitions.

## Acknowledgements

## References

Acocks, J.P.H. 1975. Veld types of South Africa. Memoirs of the Botanical Survey of South Africa 54.

Andrewartha, H.G., and Birch, L.C. 1954. The distribution and abundance of animals. University of Chicago Press, Chicago, 782 p.

Andrewartha, H.G. and Birch, L.C. 1986. The ecological web: more on the distribution and abundance of animals. University of Chicago Press, Chicago, 506 p.

Anselin, L. 1998. Exploratory spatial data analysis in a geocomputational environment. In: Longley, P. A., Brooks, S., MacMillan, B., McDonnel, R. (Eds). Geocomputation: a primer. John Wiley, New York, pp. 77-94.

Anselin, L. 1995. Local indicators of spatial association - LISA. Geographical Analysis 27, 93-115

Bao, S., Henry, M. 1996. Heterogeneity isues in local measurements of spatial association. Geographical Systems 3: 1-13.

Beaumont, L. J., Hughes, L. 2002. Potential changes in the distributions of latitudinally restricted Australian butterfly species in response to climate change. Global Change Biology 8, 954-971.

Boone, R. B., Krohn, W. B. 2000a. Partitioning sources of variation in vertebrate species richness. Journal of Biogeography 27, 457-470.

Boone, R. B., Krohn, W. B. 2000b. Relationship between avian range limits and plant transition zones in Maine. Journal of Biogeography 27, 471-482.

Brito, J. C., Crespo, E. G., Paulo, O. S. 1999. Modelling wildlife distributions: logistic multiple regression vs overlap analysis. Ecography 22, 251-260.

Busby, J. R. 1986. A biogeoclimatic analysis of Nothofagus cunninghamii (Hook.) Oerst. in southeast Australia. Australian Journal of Ecology 11, 1-7.

Busby, J. R. 1988. Potential impacts of climate change on Australia's flora and fauna. In: Pearman, G. I. (Ed). Greenhouse: planning for climate change. CSIRO, Melbourne, pp. 387-396.

Busby, J. R. 1991. BIOCLIM - a bioclimate analysis and prediction system. In: Margules, C. R.,Austin, M. P. (Eds). Nature conservation: cost effective biological surveys and data analysis. CSIRO, Australia pp. 64-68.

Bylund, H. 1999. Climate and the population dynamics of two insect outbreak species in the north. Ecological Bulletins 47, 54-62.

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

Cowling, R. M., Lombard, A.T. 2002. Heterogeneity, speciation/extinction history and climate: explaining regional plant diversity patterns in the Cape Floristic Region. Diversity and Distributions 8, 163-179.

Cowling, R.M., Lombard, A.T., Rouget, M., Kerley, G.I.H., Wolf, T., Simms-Castley, R., Knight, A., Vlok, J.H.J., Pierce, S.M., Boshoff, A.F. Wilson, S.L. 2003. A conservation assessment of the subtropical thicket biome. TERU Report 43, Terrestrial Ecology Research Unit, University of Port Elizabeth, pp. 112. http://zoo.upe.ac.za/step

Cumming, G.S. 2002. Using between model-comparisons to fine-tune linear models of species' ranges. 27, 441-455.

Currie, D. J. 1991. Energy and large-scale patterns of animal- and plant-species richness. American Naturalist 137,27-49.

Currie, D. J., Paquin, V. 1987. Large-scale biogeographical patterns of species richness of trees. Nature 329, 326-327.

Daszak, P., Cunningham, A. A., Hyatt, A. D. 2000. Emerging infectious diseases of wildlife - threats to biodiversity and human health. Science 287, 443-449.

Davis, A. J., Lawton, J. H., Shorrocks, B.,  Jenkinson, L. S. 1998. Individualistic species responses invalidate simple physiological models of community dynamics under global environmental change. Journal of Animal Ecology 67: 600-612.

Dean, W. R. J. 2000. Factors affecting bird diversity patterns in the Karoo. South Africa. South African Journal of Science 96, 609-616.

De Klerk, H.M., Crowe, T.M., Fjeldså, J.,Burgess, N.D. 2002. Biogeographical patterns of endemic terrestrial Afrotropical birds. Diversity and Distributions 8,147-162.

Deleo, J. M.1993. Receiver operating characteristic laboratory (ROCLAB): software for developing decision strategies that account for uncertainty. Proceedings of the Second International Symposium on Uncertainty Modelling and Analysis, IEEE Computer Society Press, Ottawa.

Delong, E. R., Delong, D. M., Clarke-Pearson, D. L. 1988. Comparing areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44, 837-845.

Enquist, C.A.F. 2002. Predicted regional impacts of climate change on the geographical distribution and diversity of tropical forests in Costa Rica. Journal of Biogeography 29, 519-534.

Erasmus, B. F. N., Kshatriya, M., Mansell, M. W., Chown, S. L., Van Jaarsveld, A. S. 2000. A modelling approach to antlion (Neuroptera: Myrmeleontidae) distribution patterns. African Entomology 8, 157-168.

Erasmus, B. F. N., Van Jaarsveld, A. S., Chown, S. L., Kshatriya, M.,Wessels, K. J. 2002. Vulnerability of South African animal taxa to climate change. Global Change Biology 8, 678-693.

Feria A., T. P., Peterson, A. T. 2002. Prediction of bird community composition based on point-occurrence data and inferential algorithms: a valuable tool in biodiversity assessments. Diversity and Distributions 8, 49-56.

Fielding, A. H., Bell, J. F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24, 38-49.

Gaston, K.J. 2003. The Structure and Dynamics of Geographic Ranges. Oxford University Press, Oxford.

Gaston, K. J., Rodrigues, A. S. L., van Rensburg, B. J., Koleff, P., Chown, S. L. 2001 Complementary representation and zones of ecological transition. Ecology Letters 4, 4-9.

Guegan J-F., Lek, S.,Oberdorff, T. 1998. Energy availability and habitat heterogeneity predict global riverine fish diversity. Nature 391, 382-384.

Guisan, A., Zimmerman, N. E. 2000. Predictive habitat distribution models in ecology. Ecological Modelling 135, 147-186.

Hannah, L., Midgley, G. F., Lovejoy, T., Bond, W. J., Bush, M., Lovett, J. C., Scott, D., Woodward, F. I. 2002. Conservation of biodiversity in a changing climate. Conservation Biology 16, 264-268.

Harrison, J.A., Allan, D.G., Underhill, L.G., Herremans, M., Tree, A.J., Parker, V., Brown, C.J. 1997. The atlas of southern African birds. Birdlife South Africa, Johannesburg, vol I 732 pp., vol II 785 pp.

Harvell, C. D., Mitchell, C. E., Ward, J. R., Altizar, S., Dobson, A. P., Ostfeld, R. S., Samuel, M. D. 2002. Climate warming and disease risks for terrestrial and marine biota. Science 296, 2158-2162.

Hulme, M., Conway, D., Joyce, A., Mulenga, H. 1996. A 1961-90 climatology for Africa south of the Equator and a composition of potential evapotranspiration estimates. South African Journal of Science 92, 334-343.

Jeffree, C. E., Jeffree, E. P. 1996. Redistribution of the potential geographical ranges of Mistletoe and Colorado beetle in Europe in response to the temperature component of climate

change. Functional Ecology 10, 562-577.

Jeffree, E. P., Jeffree, C. E. 1994. Temperature and the biogeographical distributions of species. Functional Ecology 8, 640-650.

Jetz, W., Rahbek, C. 2002. Geographic range size and determinants of avian species richness. Science 297, 1548-1551.

Kerley, G.I.H., Boshoff, A.F., Knight, M.H. 1999. Ecosystem integrity and sustainable land-use in the thicket biome South Africa. Ecosystem Health, 5(2):104-109.

Knapp, A. K., Smith, M. D. 2001. Variation among biomes in temporal dynamics of aboveground primary production. Science 291, 481-484.

Kobler, A., Adamic, M. 2000. Identifying brown bear habitat by a combined GIS and machine learning method. Ecological Modelling 135, 291-300.

Koch, S.O. *et al.* 2000. Conservation strategies for poorly surveyed taxa: a dung beetle (Coleoptera, Scarabaeidae) case study from southern Africa. Journal of Insect Conservation, 4, 45-56.

Kraemer, H. C. 1988. Assessment of 2x2 associations: generalization of signal-detection methodology. The American Statistician 42, 37-49.

Krebs, C. J. 1989. Ecological methodology, 2nd edition. Harper Collins, New York, pp. 654.

Leathwick, J. R., Burns, B. R.,Clarkson, B. D. 1998. Environmental correlates of tree alpha-diversity in New Zealand primary forests. Ecography 21, 235-246.

Lindenmayer, D. B., Nix, H. A., McMahon, J. P., Hutchinson, M. F., Tanton, M. T. 1991. The conservation of Leadbeater's possum, *Gymnobelideus leadbeateri* (McCoy): a case study of the use of bio climatic modelling. Journal of Biogeography 18, 371-383.

Low, A.B., Rebelo, A.G. 1996. Vegetation of South Africa, Lesotho, and Swaziland. Department of Environmental Affairs and Tourism, Pretoria, South Africa.

Manel, S., Dias, J-M., Ormerod, S. 1999. Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with a Himalayan

river bird. Ecological Modelling 120, 337-347.

Manel, S., Williams, H. C., Ormerod, S. J. 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. Journal of Applied Ecology 38, 921-931.

McLaughlin, J. F., Hellman, J. J., Boggs, C. L.,Ehrlich, P. R. 2002. Climate change hastens population extinctions. Proceedings of the National Academy of Sciences of the United States of America 99, 6070-3074.

Nix, H. 1986. A biogeographic analysis of Australian elapid snakes. In: Longmore, R. (Ed). Atlas of elapid snakes of Australia. Bureau of Flora and Fauna, Canberra. Pp. 4-15.

O'Brien, E. M. 1993. Climate gradients in woody plant species richness: towards an explanation based on an analysis of southern Africa's woody flora. Journal of Biogeography 20,181-198.

O'Brien, E. 1998. Water-energy dynamics, climate, and prediction of woody plant species richness: an interim general model. Journal of Biogeography 25,379-398.

O'Brien, E. M.,Whittaker, R. J.,Field, R. 1998. Climate and woody plant diversity in southern Africa: relationships at species, genus and family levels. Ecography 21, 495-509.

O'Brien, E. M., Field, R.,Whittaker, R.J. 2000. Climatic gradients in woody plant (tree and shrub) diversity: water-energy dynamics, residual variation, and topography. Oikos 89, 588-600.

Obuchowski, N. A., Lieber, M. L. 1998. Confidence intervals for the receiver operating characteristic area in studies with small samples. Academic Radiology 5, 561-571.

Olwoch, J.M, Rautenbach, C.J.de W., Erasmus, B.F.N., Engelbrecht, F.A., Van Jaarsveld, A.S. In press. Simulating ticks distributions over sub-Saharan Africa: the use of observed and simulated climate surfaces. Journal of Biogeography.

Oreskes, N., Shrader-Frechette, K.,Belitz, K. 1994. Verification, validation and confirmation of numerical models in the earth sciences. Science 263, 641-644.

Petchey, O.L., McPhearson, P.T., Casey, T.M., Morin, P. J. 1999. Environmental warming

alters food-web structure and ecosystem function. Nature 402, 69-72.

Peterson, A.T.,Cohoon, K. P. 1999. Sensitivity of distributional prediction algorithms to geographic data completeness. Ecological Modelling 117, 159-164.

Peterson, A.T., Egbert, S. L., Sánchez-Cordero, V., and Price, K. P. 2000. Geographic analysis of conservation priority: endemic birds and mammals in Veracruz, Mexico. Biological Conservation 93, 85-94.

Peterson, A.T., Sanchez-Cordero, V., Soberon, J., Bartley, J., Buddemeier, R.H., Navarro-Siguenza, A. G. 2001. Effects of global climate change on geographic distributions of Mexican Cracidae. Ecological Modelling 144, 21-30.

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

Peterson, A.T., Ortega-Huerta, M. A., Bartley, J., Sanchez-Cordero, V., Soberon, J., Buddemeier, R. H., Stockwell, D. R. B. 2002. Future projections for Mexican faunas under global climate change scenarios. Nature 416, 626-628.

Pounds, J. A., Fogden, M. P., Savage, J. M., and Gorman, G. C. 1997. Tests of null models for amphibian declines on a tropical mountain. Conservation Biology 11, 1307-1322.

Prendergast, J.R., Quinn, R.M., Lawton, J.H., Eversham, B.C., Gibbons, D.W. 1993. Rare species, the coincidence of diversity hotspots and conservation strategies. Nature 356, 335-337.

Robertson, M. P., Caithness, N.,Villet, M. H. 2001. A PCA-based modellingtechnique for predicting environmental suitability for organisms from presence records. Diversity and Distributions 7, 15-27.

Rogers, D. J.,Randolph, S. E. 2000. The global spread of malaria in a future warmer, warmer world. Science 289, 1763-1766.

Rutherford, M. C., Powrie, L. W., Schulze, R. E. 1999. Climate change in conservation areas of South Africa and its potential impact on floristic composition: a first assessment. Diversity and Distributions 5, 253-262.

Rutherford, M.C., Westfall, R.H. 1986. Biomes of southern Africa – an objective categorization. Memoirs of the Botanical Survey of South Africa 54.

Samways, M.J. 2003. Critical response from Professor Michael J. Samways. Journal of Biogeography, 30:817.

Scholtz, C.H., Chown, S.L. 1995. Insects in southern Africa: How many species are there? South African Journal of Science, 91, 124-126.

Schulze, R. 1997. Climate. In: Cowling, R.M., Richardson, D.M.  Pierce, S.M. (eds.) Vegetation of Southern Africa, Cambridge University Press, Cambridge, pp613.

Schulze, R., Maharaj, M. 1997. South African Atlas of Agrohydrology and Climatology.Water Research Commission report TT82/96, Pretoria, 276 pp.

Schulze, R., Meigh, J., Horan, M. 2001. Present and potential future vulnerability of eastern and southern Africa's hydrology and water resources. South African Journal of Science 97, 150-160.

Stockwell, D.,Peters, D. 1999. The GARP modelling system: problems and solutions to automated spatial prediction. International Journal of Geographical Information Science 13,143-158.

Sutherst, R.W. 2003. Prediction of species geographical ranges. Journal of Biogeography, 30: 805–816

Sutherst, R. W., Maywald, G. F. 1985. A computerised system for matching climates in ecology. Agriculture, Ecosystems and Environment 13, 281-299.

Todd, M. C., Washington, R., Cheke, R. A., Kniveton, D. 2002. Brown locust outbreaks and climate variability in southern Africa. Journal of Applied Ecology 39, 31-42.

Van Jaarsveld, A. S., Gaston, K. J., Chown, S. L., Freitag, S. 1998a. Throwing biodiversity out with the binary data? South African Journal of Science 94, 210-214.

Van Jaarsveld, A. S., Freitag, S., Chown, S. L., Muller, C., Koch, S., Hull, H., Bellamy, C., Kruger, M., Endrody-Younga, S., Mansell, M. W.,Scholtz, C. H. 1998b. Biodiversity assessment and conservation strategies. Science 279, 2106-2108.

Van Jaarsveld, A.S., Chown, S.L. 2001. Climate change and its impacts in South Africa. Trends in Ecology and Evolution 16, 13-14

Van Rensburg, B. J., Chown, S. L.,Gaston, K. J. 2002. Species richness, environmental correlates, and spatial scale: a test using South African birds. The American Naturalist 159(5), 566-577.

Van Staden, V., Erasmus, B.F.N., Roux, J., Wingfield, M.J., Van Jaarsveld, A.S. In press. Modelling the spatial distribution of two important South African plantation forestry pathogens. Forest Ecology and Management.

Van Zinderen Bakker Sr, E.M. 1978. Quaternary vegetation changes in southern Africa. In: Werger, M.J.A. (ed) Biogeography and Ecology of Southern Africa. W.Junk Publishers, The Hague, pp 131 – 143.

Vida, S. 1993. A computer program for non-parametric receiver operating characteristic analysis. Computer Methods and Programs in Biomedicine 40, 95-101.

Vlok, J.H., Euston-Brown, D. 2002. Subtropical thicket ecosystem planning: biological survey report. Unpublished report. Terrestrial Ecology Research Unit, University of Port Elizabeth, pp. 165. (http://zoo.upe.ac.za/step)

Walker, P. A., Cocks, K. D. 1991. HABITAT: a procedure for modelling a disjoint environmental envelope for a plant or animal species. Global Ecology and Biogeography Letters 1, 108-118.

Watt, A. D., McFarlane, A. M. 2002. Will climate change have a different impact on different trophic levels? Phenological development of winter moth *Operophtera brumata* and its host plants. Ecological Entomology 27, 254-256.

Werger, M. J. A. 1978. Biogeographical division of southern Africa. In: M. J. A. Werger (ed) Biogeography and Ecology of Southern Africa, W. Junk, The Hague, pp. 145-170.

Zar J.H. (1996). Biostatistical analysis, 3$^{rd}$ edition. Prentice-Hall, Upper Saddle River, NJ, pp. 121.