

Species distribution modelling for conservation planning in Victoria of Australia

C. Liu^a, M. White^a, G. Newell^a and P. Griffioen^b

^aArthur Rylah Institute for Environmental Research, Department of Sustainability and Environment, 123 Brown Street, Heidelberg, Victoria 3084, Australia; ^bPeter Griffioen Consulting, 10 Cleve Grove, Heidelberg, Victoria 3084, Australia
Email: canran.liu@dse.vic.gov.au

Abstract: Biodiversity and natural ecosystems are under pressure from a large range of drivers. In south eastern Australia these include periods of drought and occasional floods, increasing urbanisation pressures, changing land use patterns, altered fire frequency and regime, habitat loss and degradation, impacts of invasive species, etc. Detailed and reliable information about the spatial distribution of species provides critical information for effective conservation planning. In Victoria, Australia, this responsibility is held by the Department of Sustainability and Environment, which maintains and curates a database of site records for species: the Victorian Biodiversity Atlas (VBA). Although this database is comprehensive and contains many records for numerous species, the information is provided in point form. This point data can be useful for some planning and management tasks, but does not provide an adequate view of distribution for all management purposes. By integrating known occurrences of species with environmental GIS data layers using statistical or machine learning algorithms, it is possible to build models and to predict the likelihood that the species may occur at particular locations across large landscapes. We have used the VBA to build species distribution models (SDMs) for 523 vertebrate fauna species across the whole state, providing predictive ‘maps’ that are available for use in various conservation planning activities. In this paper, we introduce and discuss the methods we developed and implemented for producing these models.

Modelling requires a number of steps: gathering species occurrence data, vetting or mining this data for useful records and winnowing out potentially misleading records, collating existing and / or developing new relevant environmental variables, selecting an appropriate modelling algorithm, training and evaluating models, mapping predictions to the whole state scale, and selecting a threshold to transform the continuous model result into binary (i.e. presence/absence) mapped outputs.

All the occurrence data for the 523 fauna species were derived from the VBA. This database contains presence data for fauna species, but for the majority of survey techniques information on absences cannot be known or inferred. Twenty four explanatory variables were selected for use in developing models including 12 bioclimatic, terrain and soil radiometric (BTSR) variables, two coordinates (latitude and longitude), and 10 remotely sensed variables primarily derived from a time series of Landsat imagery. For each species, three versions of models (V1, V2 and V3 models) reflecting different constraint were built. V1 models used 12 BTSR variables; V2 models used 12 BTSR variables and two coordinates; and V3 models used all the 24 variables.

Since true absence data was not available, pseudo-absences were created by filtering random points by a profile model based on Mahalanobis distance method. Using both pseudo-absences and the true presences data SDMs were developed with a machine learning algorithm, random forest, which was implemented in the statistical computing environment R. The models were evaluated using six accuracy measures, which have been demonstrated suitable for evaluating presence-only models. These metrics include *AUL* (area under the lift curve), *MVDI* (the maximum vertical distance from lift curve to the diagonal line), *MPA* (minimal predicted area with pre-specified true positive rate), *VDI* (vertical distance from lift curve to the diagonal line), *Lift* and *Se* (sensitivity). These measures can be used for both within-species and among-species model comparisons. A threshold was set for each model with a method based on maximising the sum of sensitivity and specificity to provide binary outputs or maps of species’ distributions.

Keywords: species distribution modelling, prediction, evaluation, random forest, presence-only

1. INTRODUCTION

Victoria's biodiversity provides critical ecosystem services (BES, 2010), such as water purification, carbon storage for climate regulation, nutrient cycling, pollination of food crops and so on. However, Victoria's biodiversity and ecosystems are under pressure from a number of factors including droughts, floods, altered and fluctuating water regimes, increasing human population growth and urbanisation, land use change, altered fire regimes, invasive plants and animal species, and other drivers (BES, 2010; DSE, 2010). Climate change will exacerbate these pressures, resulting in far reaching changes to species distribution and abundance, and ecosystem function (BES, 2010).

Information on the spatial distributions of species is critical for both protecting the biodiversity and setting appropriate and achievable conservation priorities. Data on the distribution of plants and animals have been collected over many decades by public and Government agencies, and also by interested and keen private citizens. These records have been incorporated into the Victorian Biodiversity Atlas (VBA) and various databases that precede this current form. These datasets only hold the original point data, but do not provide information on locations where the species *do not* occur. Therefore, these data can be difficult to interpret directly for conservation planning. However, these data can be used together with environmental data to build distributional models, which can be extrapolated across the state to provide an estimate of the likelihood of suitable habitat. We have developed and tested several modelling strategies through many modelling experiments. We also used the best of these approaches to construct distributional models for 523 of Victoria's vertebrate fauna species. In this paper, we describe the strategy we have adopted to construct these models.

2. SPECIES DATA

Ideally data on species occurrences would all be derived from surveys that have been based upon rigorous statistical sampling design. However, this is the exception rather than the rule. The majority of data within the databases come from small scale studies and from opportunistic samples. Our data are from VBA, which includes records from surveys carried out by the researchers in Victoria Department of Sustainability and Environment, as well as data from university studies, external contractors, and members of the general public. As might be expected, there are some limitations when data from various sources are pooled (Graham *et al.*, 2004). Sources of error and bias include (Graham *et al.*, 2004): (1) Taxonomic identification error: This can be from incorrect identification as well as an incomplete or outdated taxonomic knowledge. To avoid such errors, a thorough knowledge of the taxonomic and systematic history of the group under study is needed and physical examination of the specimens is required. (2) Spatial error: This can be derived from many sources, but can often be detected by plotting the record points in the geographical and environmental spaces to detect outliers. This kind of error can be corrected by checking the specimens and the filed notes. The suspect records should be eliminated or down-weighted. (3) Collection bias in geographical space: We have more records near roads, towns and cities as well as accessible areas on public land. We have relatively very few records from freehold land, or from locations with steep or difficult terrain, very dense vegetation or remote areas. For species distribution modelling, spatial bias in the records may not be a problem if the data are not environmentally biased (Newbold, 2010). Some studies have shown that museum records did not completely capture the environmental conditions inhabited by the target species (e.g. Hortal *et al.*, 2008). Other studies have shown that spatially biased museum record data are unrelated to major biases in environmental space (e.g. Kadmon *et al.*, 2004; Newbold *et al.*, 2009). (4) Presence-only data: As indicated above, biodiversity databases regularly contain a mixture of data from a variety of sources ranging from systematic surveys to incidental observations. It is very difficult to infer absence data. We therefore considered our data as presence-only, and pseudo-absence data (Graham *et al.*, 2004) need to be created for some modelling techniques (including the one used in our work) requiring both presence and absence data.

We addressed these four issues in our work. The first two were dealt with by carefully checking the records and eliminating misleading data. The last two issues were dealt with by selecting pseudo-absences using targeted strategies which will be discussed later in this paper. The presence data for each species was split into two parts: training data (70%) and test data (30%). When the model is successfully evaluated with the test data, another model using all of the presence data is built to provide the predictions over the whole state.

3. ENVIRONMENTAL VARIABLES

The distribution of species in space and time is influenced by many drivers, including their environmental context. Theoretically a large number of environmental variables are potentially related to species distributions. While some variables may have direct effect on species distributions, others may only exert

indirect effects. Understanding the main drivers shaping species geographic distributions is particularly important for species distribution modelling (Pulliam, 2000). Spatial representations of environmental data have become easily accessible, however caution is required since it is important to only include variables that are ecologically relevant to the target species (Elith and Leathwick, 2009). Primary drivers are related to environmental regimes of heat, moisture and nutrients (Franklin, 2009), and these variables may have variable effects on species at different scales.

Though geographical coordinates (i.e. latitude and longitude) are not environmental variables directly, they can be used to account for some aspects of 1) historic accident, 2) founder effects, 3) species interactions, and 4) hidden or latent variables that may be to some degree spatially autocorrelated. For these reasons we have incorporated coordinates in SDM development (as in Ohse *et al.*, 2009; Yen *et al.*, 2011).

In addition, remotely sensed data can also provide important information relevant to (current) species distributions. Some studies have demonstrated that including remotely sensed variables as predictors in SDMs improves their performance (e.g. Zimmermann *et al.*, 2007). This kind of data is especially related to current environment for the species.

For our work 24 explanatory variables were selected to develop models (Table 1), and these were all represented at 30m×30m resolution. Three versions of spatial models (V1, V2 and V3 models) were developed for a range of anticipated uses. The V1 models only use the 12 bioclimatic, terrain and soil radiometric (BTSR) variables. The intention is to produce a model that extrapolates beyond the range of known geographic occurrence and predicts unoccupied but suitable habitat and/or potentially occupied but poorly surveyed habitat. Such a model could be used for designing new surveys for particular species or identifying potential translocation sites. V2 models used the same 12 BTSR variables however the x and y coordinates have also been supplied as additional independent variables. These models have been developed to predict the biophysically defined niche of the species, within its known geographic range and prior to habitat removal. Such models can be used to identify and select areas for habitat restoration and in conjunction with V3 models determine rough levels of habitat depletion. V3 models used all 24 variables, as predictors. The intention is to produce an accurate and contemporary distribution of the species. It is anticipated that V3 models would be used for systematic strategic and conservation planning and in support of the regulation of land-use and development.

4. MODELLING TECHNIQUE

There are many methods available to construct species distribution models, including statistical methods as well as machine-learning approaches, and both approaches are used widely for these tasks (see Elith *et al.*, 2006). Through comparison of 16 modelling techniques, these authors concluded that boosted regression trees (BRT) and Maxent are among the best methods. An additional machine-learning method developed by Breiman (2001), random forest (RF), has also been used in species distribution modelling (e.g. Cutler *et al.*, 2007; Liu *et al.*, 2009; Williams *et al.*, 2009; Stohlgren *et al.*, 2010). Through modelling and evaluation experiments we identified that models built with RF and BRT have similar levels of accuracy. We subsequently chose to use RF approach for computation and spatial application purposes.

RF is an ensemble learner, which consists of many (*ntree*) classification trees. A different bootstrap sample of the data is used to construct each tree. For each tree each node is split using the best among a subset (*mtry*) of predictors randomly chosen at that node, and the tree is grown to the largest extent possible without pruning (Liaw and Wiener, 2002). When classifying a new point, the point is run through each of the trees of the forest, and each tree provides a classification. The forest chooses the classification having the most votes (across all trees in the forest) (Breiman and Cutler, Random forests, URL http://oz.berkeley.edu/users/breiman/RandomForests/cc_home.htm). RF consequently has two main parameters [the number of variables in the random subset at each node (*mtry*), and the number of trees in the forest (*ntree*)]. As a general rule, *ntree* should not be set to too small a number to ensure that every training point gets predicted at least a few times (Liaw and Wiener, 2002). The software package provides default settings for these parameters (e.g. *ntree* = 500, and *mtry* = square root of the number of variables). Models are regularly robust to these values (Liaw and Wiener, 2002), and our experiments confirm these findings.

5. BUILDING MODELS

For most fauna species databases provide no information about absences or when a species was not detected at a site. RF models require absence data, and we therefore created pseudo-absences for modeling purposes. There are several approaches to developing and selecting pseudo-absences. Two of these include (1) choosing points randomly from all the state; or (2) choosing points randomly from all the state but excluding

the sites with known records. A third approach that we developed for our work was to use a model to filter the random points. This process built a profile model with only the true presence data to develop an environmental envelope for the species, and then selected a cut-off value to ensure 95% of the true presences are correctly predicted. Finally this model was applied to random points, and points predicted as presences were excluded while the remaining points were retained as pseudo-absences. For this project we used Mahalanobis distance (MD) to develop profile models. We compared the above three pseudo-absence selection methods with modelling experimentation and determined that the third approach outlined above performed much better than the other two processes.

Table 1. List of predictors used in developing each of the three different versions of SDMs.

Code	Variable	Model version
1	Mean annual rainfall	1,2,3
2	Mean annual rainfall less mean annual evaporation	1,2,3
3	The highest temperature of any weekly maximum temperature	1,2,3
4	The lowest temperature (°C) of any weekly minimum temperature	1,2,3
5	Log of altitude	1,2,3
6	Flow accumulation	1,2,3
7	Terrain Complexity Index	1,2,3
8	Shade relief at summer	1,2,3
9	Shade relief at winter	1,2,3
10	Ratio of the inverse radioelement count of Pottasium and the radioelement count of Thorium	1,2,3
11	Ratio of the radioelement count of Thorium and the radioelement count of Potassium	1,2,3
12	Ratio of the inverse radioelement count of Thorium and the radioelement count of Potassium	1,2,3
13	Longitude	2,3
14	Latitude	2,3
15	Median Normalised Difference Wetness Index derived from LANDSAT 7 TM of years 1989 – 2005	3
16	Median Tasseled Cap Brightness derived from LANDSAT 7 TM of years 1989 – 2005	3
17	Median Tasseled Cap Greenness derived from LANDSAT 7 TM of years 1989 – 2005	3
18	Median Tasseled Cap Wetness derived from LANDSAT 7 TM of years 1989 – 2005	3
19	Median Tasseled Cap fourth (haze transformation1) Wetness derived from LANDSAT 7 TM of years 1989 – 2006	3
20	Median Tasseled Cap fifth (haze transformation2) Wetness derived from LANDSAT 7 TM of years 1989 – 2007	3
21	Median Tasseled Cap sixth (haze transformation3) Wetness derived from LANDSAT 7 TM of years 1989 – 2008	3
22	Median Soil-adjusted Total Vegetation Index derived from LANDSAT 7 TM of years 1989 – 2009	3
23	Median Normalised Difference rock index derived from LANDSAT 7 TM of years 1989 – 2010	3
24	Median Enhanced Vegetation Index derived from LANDSAT 7 TM of years 1989 – 2009	3

For V1 model, we first used the 12 BTR variables to build a MD model, selected pseudo-absences from the environmental space which were outside of the 95% profile defined by this MD model, and built the V1 model using random forest with these pseudo-absences and presence data from the VBA.

For V2 model, we selected pseudo-absences from geographical space which were outside the 95% profile defined by the V1 model, and built the V2 model with these pseudo-absences and the presences.

For V3 model, we used the 10 remotely sensed variables to build a MD model, selected pseudo-absences from the geographical space which were within the 95% profile defined by this MD model and outside the 95% profile defined by the V2 model.

Many approaches to forming SDMs use a large number of pseudo-absences in their modelling. The results from our modelling experiments clearly identified that this may not be a good strategy when using the random forest approach. We found that using the number of pseudo-absences ($n0$) a few times larger than the number of training presences ($n1$) provided better model performance. Specifically, we set $n0 = k * n1$ with $k = 5$ if $n1 < 20$, and $k = 2$ otherwise.

6. MODEL EVALUATION

Model evaluation describes the processes required to justify the acceptance of a model for its intended purpose, and the choice of the evaluation strategy needs to be explicitly related to the subject and goals of modelling (Araújo and Guisan, 2006). Our primary purpose for building species distribution models was to predict the potential distributions of species, and therefore model generality or transferability (the ability of a model to predict independent events) must be evaluated.

If both presence and absence data are available, many accuracy measures can be used for model evaluation (see Liu *et al.*, 2011 for a review). Some researchers believe that the existing accuracy measures developed for presence/absence data (PA measures) cannot be used legitimately when absence data are not available (e.g. Boyce *et al.*, 2002; Pearson *et al.*, 2007; Peterson *et al.*, 2008). However others have considered that (at least) *AUC* [area under the receiver's operating characteristic (ROC) curve] can be used when absence data are replaced with random points (e.g. Phillips *et al.*, 2009). We proved theoretically and demonstrated through simulations that five existing accuracy measures including *NPV* (negative predictive value), *PPV* (positive predictive value), *TSS* (true skill statistic), *MVDr* (the maximum vertical distance from ROC curve to the diagonal line) and *AUC* can be used for within-species and but not among-species model comparisons.

We also investigated the behaviours of eight accuracy measures based on presence-only data (PO measures) and found that six including *AUL* (area under the lift curve), *MVDI* (the maximum vertical distance from lift curve to the diagonal line), *MPA* (minimal predicted area with pre-specified true positive rate), *VDI* (vertical distance from lift curve to the diagonal line), *Lift* and *Se* (sensitivity) can be used for both within-species and among-species model comparisons. However, it should be noted that different elements are being compared in these two different types of model comparisons. While the former compares the abilities of the models to distinguish true presences from true absences, the latter only compares the abilities of the models to distinguish true presences from random points. We also theoretically proved that within-species model comparisons with *AUL*, *MVDI*, *VDI* and *Lift* using PO data are equivalent to those with *AUC*, *MVDr*, *VDr* and *PPV* using PA data. Since PO measures can be used for model comparisons for both within-species and among-species, they (especially *AUL*, *MVDI*, *VDI*, *Lift* and *Se*) are preferred to PA measures when only presence data is available. Consequently we used these measures in our model evaluation.

7. THRESHOLD SELECTION

The species distribution modelling procedure used here produces continuous suitability predictions. These products need to be transformed into binary outputs using a threshold for both some model evaluation and for some particular applications. When both presence and absence data are available more than a dozen of threshold selection methods are available. Liu *et al.* (2005) compared 12 methods and recommended four including the method based on maximising the sum of sensitivity and specificity (*max SSS*). When absence data are not available, we also proved theoretically and demonstrated using simulations that *max SSS* is still a suitable method; however specificity is now apparent rather than true. *max SSS* was used to threshold all models developed in this study.

8. STATE-WIDE PREDICTION

When the model is considered adequate after evaluation with the test data, all the presence data is used to build another MD profile model and a set of pseudo-absences is selected for use with this model. The number of pseudo-absences is determined using the same rule as previously mentioned. With these presences and pseudo-absences another RF model is built and is used to make the predictions over the whole state (i.e. at each of the 30m pixels). As this dataset is very large, customised parallel computing was developed to use a number of personal computers simultaneously to facilitate the production of spatial outputs. As an example, the three versions of models as well as the training data for the Brush-tailed Phascogale (*Phascogale tapoatafa*) are shown in Figure 1.

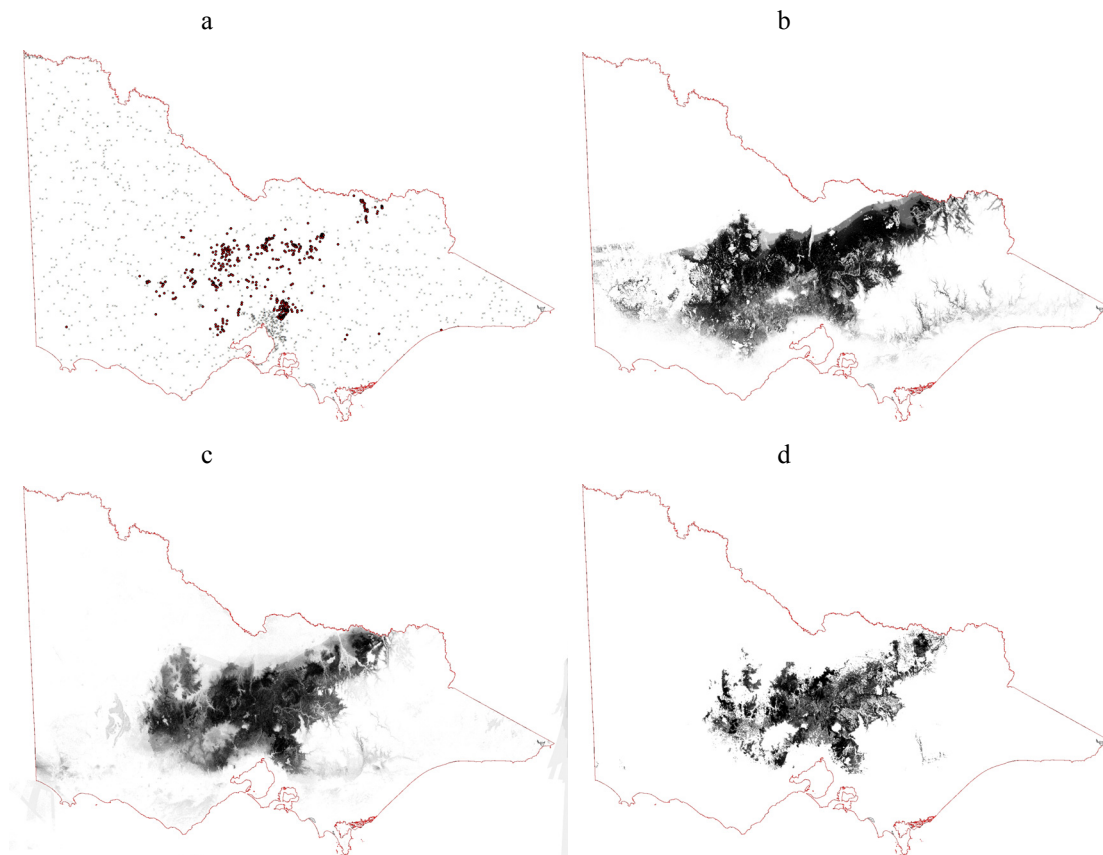


Figure 1. Data and three versions of species distribution models for the Brush-tailed Phascogale (*Phascogale tapoatafa*). (a) Presence data obtained from Victorian Biodiversity Atlas (large dots), along with pseudo-absence data (small crosses); (b) V1 model developed from 12 bioclimatic, terrain and soil radiometric (BTSR) variables; (c) V2 model developed from longitude and latitude in addition to the 12 BTSR variables; (d) V3 model developed from all the 24 variables.

9. POTENTIAL APPLICATIONS

The product of our work, i.e. the predicted distributions for each species, can be used in various applications, which may include the following:

- Advice to various levels of government and associated agencies on the likelihood of the presence of habitat. This has direct applicability for land use planning and investment (high degree of spatial accuracy and credibility required). Anticipated contexts might be choosing between bids in environmental auctions, reserve design problems, trading and finding suitable offset sites for species.
- Predicting habitat outside the known range of a species for threatened species survey, land-use planning, translocations and scenario planning.
- Analyses of historic distributions for levels of habitat depletion, etc.
- Investigation, exploration and determination of specific concepts such as critical habitat, persistent habitat and thresholds / criteria for conservation actions. SDMs would be a primary input into such analyses.
- Modelling invasive species and species that are not at equilibrium.

REFERENCES

- Ara'ujo, M. B., and A. Guisan (2006). Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, 33, 1677–1688.
- Biodiversity and Ecosystem Services (BES) (2010). *Business Plan 2010-11*. Victorian Department of Sustainability and Environment, Melbourne, Australia.
- Victorian Department of Sustainability and Environment (DSE) (2010). *Victoria's biodiversity strategy 2010-2015*. Victorian Department of Sustainability and Environment, Melbourne, Australia.
- Boyce, M. S., P. R. Vernier, S. E. Nielsen, and F. K. A. Schmiegelow (2002). Evaluating resource selection functions. *Ecological Modelling*, 157, 281–300.

- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32.
- Cutler, D. R., T. C. Edwards, Jr., K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler (2007). Random forests for classification in ecology. *Ecology*, 88, 2783–2792.
- Elith, J., C.H. Graham, R.P. Anderson, M. Dudík, S. Ferrier, A. Guisan, R.J. Hijmans, F. Huettmann, J.R. Leathwick, A. Lehmann, J. Li, L.G. Lohmann, B.A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J.M. Overton, A.T. Peterson, S.J. Phillips, K. Richardson, R. Scachetti-Pereira, R.E. Schapire, J. Soberón, S. Williams, M.S. Wisz, and N.E. Zimmermann (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29, 129–151.
- Elith, J., and J. Leathwick (2009). Species distribution models: Ecological explanation and prediction across space and time. *Annual Reviews of Ecology, Evolution and Systematics*, 40, 677–697.
- Franklin, J. (2009). *Mapping species distributions: Spatial Inference and Prediction*. Cambridge University Press, Cambridge.
- Graham, C. H., S. Ferrier, F. Huettman, C. Moritz, and A. T. Peterson (2004). New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology and Evolution*, 19, 497–503.
- Hortal, J., A. Jiménez-Valverde, J. F. Gómez, J. M. Lobo, and A. Baselga (2008). Historical bias in biodiversity inventories affects the observed environmental niche of the species. *Oikos*, 117, 847–858.
- Kadmon, R., O. Farber, and A. Danin (2004). Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecological Applications*, 14, 401–413.
- Liaw, A. and M. Wiener (2002). Classification and Regression by randomForest. *R News*, 2, 18–22.
- Liu, C., P. M. Berry, T. P. Dawson, and R. G. Pearson (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, 28, 385–393.
- Liu, C., M. White, and G. Newell (2009). Assessing the accuracy of species distribution models more thoroughly. In Anderssen, R.S., R.D. Braddock and L.T.H. Newham (eds) 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, July 2009, pp. 4234–4240.
- Liu, C., M. White, and G. Newell (2011). Measuring and comparing the accuracy of species distribution models with presence-absence data. *Ecography*, 34, 232–243.
- Newbold, T., F. Gilbert, S. Zalut, A. El-Gabbas, and T. Reader (2009). Climate-based models of spatial patterns of species richness in Egypt's butterfly and mammal fauna. *Journal of Biogeography*, 36, 2085–2095.
- Newbold, T. (2010). Applications and limitations of museum data for conservation and ecology, with particular attention to species distribution models. *Progress in Physical Geography*, 34, 3–22.
- Ohse, B., F. Huettmann, S. M. Ickert-Bond, and G. P. Juday (2009). Modelling the distribution of white spruce (*Picea glauca*) for Alaska with high accuracy: an open access role-model for predicting tree species in last remaining wilderness areas. *Polar Biology*, 32, 1717–1729.
- Pearson, R. G., C. J. Raxworthy, M. Nakamura, and A. T. Peterson (2007). Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, 34, 102–117.
- Peterson, A. T., M. Papes, and J. Soberón (2008). Rethinking receiver operating characteristic analysis applications in ecological niche modelling. *Ecological Modelling*, 213, 63–72.
- Phillips, S. J., M. Dudík, J. Elith, C. H. Graham, A. Lehmann, J. Leathwick, and S. Ferrier (2009). Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications*, 19, 181–197.
- Pulliam, H.R. (2000). On the relationship between niche and distribution. *Ecology Letters*, 3, 349–361.
- Thomas J. Stohlgren, T. J., P. Ma, S. Kumar, M. Rocca, J. T. Morisette, C. S. Jarnevich, and N. Benson (2010). Ensemble Habitat Mapping of Invasive Plant Species. *Risk Analysis*, 30, 224–235.
- Williams, J. N., C. Seo, J. Thorne, J. K. Nelson, S. Erwin, J. M. O'Brien and M. W. Schwartz (2009). Using species distribution models to predict new occurrences for rare plants. *Diversity and Distributions*, 15, 565–576.
- Yen, J. D. L., J. R. Thomson, P. A. Vesk, and R. Mac Nally (2011). To what are woodland birds responding? Inference on relative importance of in-site habitat variables using several ensemble habitat modelling techniques. *Ecography*, doi: 10.1111/j.1600-0587.2011.06651.x.
- Zimmermann, N. E., T. C. Edwards, G. G. Moisen, T. S. Frescino, and J. A. Blackard (2007). Remote sensing-based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah. *Journal of Applied Ecology*, 44, 1057–1067.