PREDICTIVE HABITAT MODELLING

Three Bird Species in the Swan Region

Prepared for
WA Department of Environment and Conservation

May 2009

Predictive Habitat Modelling

Three Bird Species in the Swan Region

PREPARED FOR	WA Department of Environment and Conservation
PROJECT NO	0169–0007
DATE	May 2009

DOCUMENT TRACKING

ITEM	DETAIL
Project Name	Predictive Habitat Modelling: Three Bird Species in the Swan Region
Project Number	0169-0007
File location	G:\Synergy\Projects\0169\0169-0007 Gnangarra Modelling Phase 2\Reports\Draft Reports
Prepared by	EL
Approved by	NK
Status	Final
Version Number	2
Last saved on	20 May 2009

ACKNOWLEDGEMENTS

This document has been prepared by Eco Logical Australia Pty Ltd with support from the WA Department of Environment and Conservation.

Contents

List of Figures

List of Tables

1 Introduction

Eco Logical Australia Pty. Ltd. (ELA) was contracted by the Western Australian Department of Environment and Conservation (DEC) in July 2008 to undertake a pilot study of predictive habitat modelling for up to twelve flora and fauna species in the Swan Region, Western Australia. This region included the whole of the Swan Coastal Plain Bioregion, and parts of the Geraldton Sandplains, Jarrah Forest and Avon Wheatbelt Bioregions (Figure 1).

Predictive species habitat modelling can be used to provide an indication of the potential distribution and relative suitability of habitat for flora and fauna species within a given study area. This information has a wide range of applications from guiding further survey effort, to land management and planning, and applications with development control.

First and foremost, this pilot study aimed to provide the Department with exposure to predictive habitat modelling. This exposure aimed to provide the Department with knowledge and understanding of what is involved in undertaking such work. The use of the Department's spatial data was a further goal with regard to developing an understanding of the suitability and limitations of this data for such work. Another goal was for the Department to be able to examine the application of the resultant modelling products for use in conservation planning, development control and land management.

This project was designed for a staged approach, with four species to be modelled initially, followed by modelling for remaining species. Preliminary results for two of the four initial species to be modelled (Waxy-leaved Smokebush, *Conospermum undulatum*, and Graceful Sunmoth, *Synemon gratiosa*) were provided to the DEC in a previous report (ELA 2008). However, results for the remaining initial species to be modelled (Keighery's Macarthuria, *Macarthuria keigheryi*, and Native Bee, *Leioproctus douglasiellus*) were not provided as data on these species were limited, preventing the generation of meaningful predictive habitat models for these species.

In October 2008, ELA received climate data for the Swan Region from the DEC for use in subsequent predictive habitat modelling of the remaining species. In March 2009, ELA also received species location data for three bird species in the Swan Region which had been chosen by DEC for the second stage of the modelling pilot project. These species included:

- Splendid Fairy-wren (*Malurus splendens*);
- Scarlet Robin (*Petroica multicolor*); and
- Western Thornbill (*Acanthiza inornata*).

This report presents the results of the predictive habitat modelling for the above bird species. The report also outlines the methods used, identifies limitations of the project and suggests recommendations for future project work.

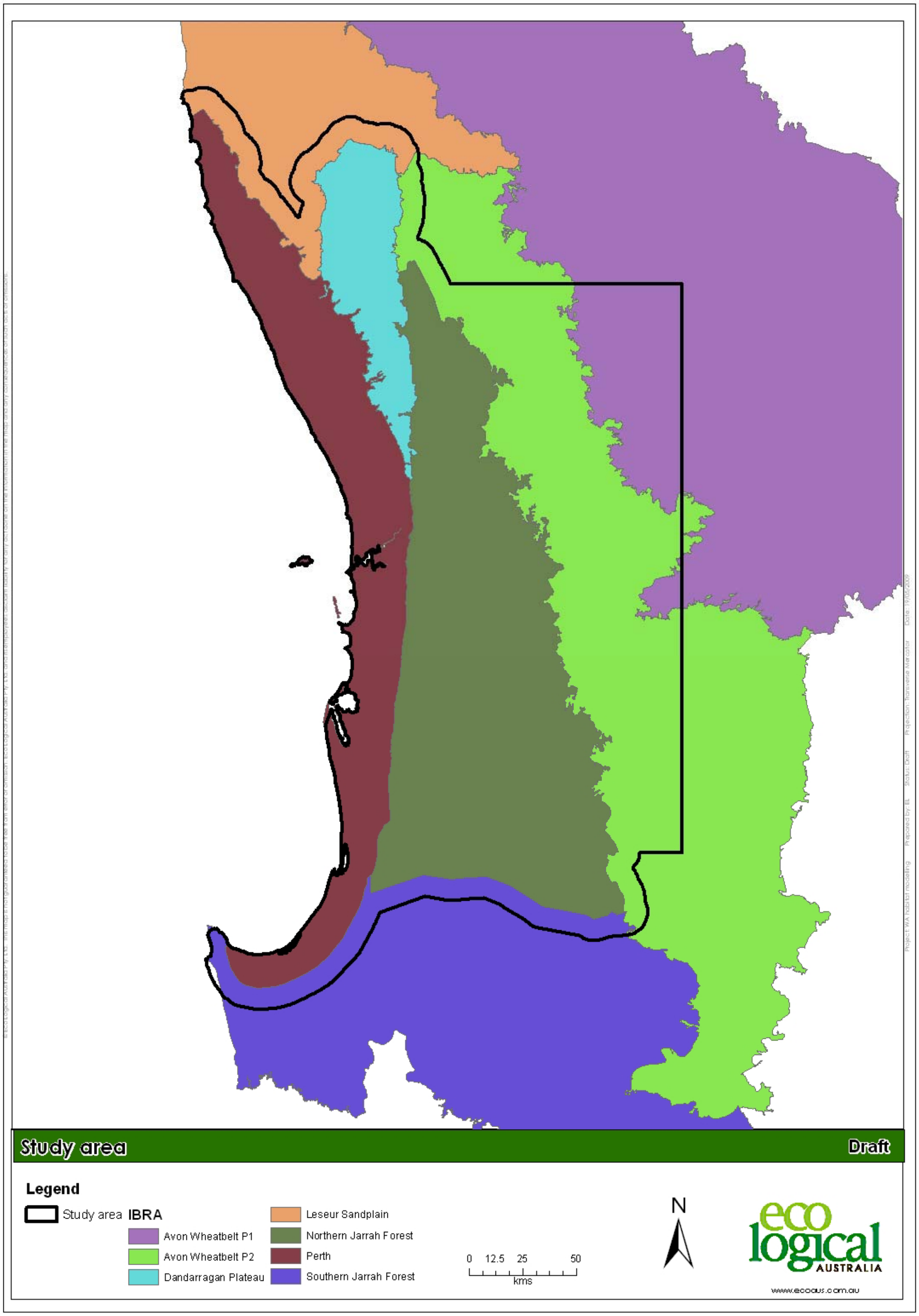

Figure 1: Modelling study area.

2 Methodology

2.1 GENERAL APPROACH

The approach taken in this predictive species habitat modelling project involved statistically analysing relationships between records of species observations with spatial datasets describing the environmental characteristics of the study area. This process relies on the input data used (records of species observations and datasets describing environmental character) and the ability to identify statistically important relationships between them. Statistical analysis software, a geographic information system (GIS), and purpose built software, was heavily utilised for this work. The major components of this work are described below.

Prior to conducting any habitat modelling, species record data and predictor layers were assessed and prepared. Data were provided by the DEC, and were prepared for modelling by ELA. Some species record data were removed from the species dataset due to accuracy concerns including:

- Records with a stated locational accuracy of greater than 100m;; and
- Records without datum information.

Records were only removed following consultation with the DEC. Preparation of the environmental data is described in Section 2.2.2.

Once species location and environmental data was prepared, predictive habitat modelling took a two-stage approach. In the first stage, relationships between species locations and environmental variables were determined to identify potential drivers for species distribution. In the second stage, areas of potential habitat were predicted in geographic space based on the identified drivers for species distribution.

Following the generation of models, all models were evaluated for model strength. Poor models were re-run as necessary until the model that performed the best against evaluation criteria was generated (see Section 2.3 for details of criteria used to evaluate model strength).

Purpose-built software developed by Lehmann *et al.* (2002; 2004) was used to determine statistical relationships between species presence records and a range of spatial predictors (environmental variables covering terrain and topographic, climate, drainage, soil and vegetation indices), and predict the distribution of potential habitat in geographic space. The software, known as “GRASP” (Generalised Regression Analysis and Spatial Prediction), was based around the use of Generalised Additive Models (GAM) in determining the relationships between species records and spatial predictors. It was used both as a module in the S-Plus Statistical Package, and as an extension in ArcView GIS.

Generally, a step-wise procedure was used to select significant environmental predictors, where a starting model including all continuous predictors smoothed with 4 degrees of freedom was first fitted. The significance of either dropping smooth terms or converting them to a linear form was then tested using an analysis of variance (F-test). The minimum contribution for predictors was set to 5% and the maximum correlation between predictors was set at 75%.

In addition, full models only including specific predictors thought to contribute to species distribution were used and then analysed. This was undertaken if the resultant models from the step-wise procedure did not accurately reflect recorded species observations, the variables chosen were not perceived to be relevant to predicting the distribution of the species, and strength statistics of the step-wise models were low.

A more detailed step-by-step breakdown of the modelling procedure is provided in Appendix A.

2.2 INPUT DATA

2.2.1 Species

Habitat modelling was undertaken for Splendid Fairy-wren, Scarlet Robin, and Western Thornbill. Of the three species, Splendid Fairy-wren had the highest number of presence records in the study area (1201 records), followed by Scarlet Robin (339 records) and Western Thornbill (185 records). The presence records for Splendid Fairy-wren, Scarlet Robin, and Western Thornbill are shown in Figure 2, **Error! Reference source not found.**, andFigure 4, respectively.

Records removed during data cleaning included 1101 records for Splendid Fairy-wren, 445 records for Scarlet Robin, and 264 records for Western Thornbill.

2.2.2 Environmental Variables

Table 1 shows the environmental variables compiled for the modelling procedure. Most were provided by the DEC apart from a number that were generated specifically for the project including slope, aspect, ruggedness and topographic position. The following steps outline the procedures used to prepare the environmental data for the project:

1. Conversion to raster dataset where source was polygon data:
    a. In these cases the values of the source polygon data were classified into a discrete number of categories prior to conversion to raster;
2. All raster data were standardised with a grid cell size of 20m;
    a. The climate data were provided as approximately 5km grid cell data. In order for this data to be useful to the modelling project it was resampled to a 20m grid cell size using the Cubic resampling technique within ESRI ArcGIS. This also provided some smoothing of the data which was also beneficial.
3. Clipped to the modelling extent (the study area as shown in Figure 1 above);
4. Converted to an integer:
    a. Various raster data was reformatted / transformed prior to conversion (particularly climate data) in order to retain a sufficient level of detail. For example, the maximum temperature variable was transformed by multiplying all values by 100 such that detail to 2 decimal places was retained prior to conversion to an integer.

Two environmental variables that were also prepared, but were not used in analyses were flow accumulation and distance to coast. These variables were not used in analyses due to their large file sizes; ArcView was not able to process any files larger than 2GB, and files containing these variables were greater than 2GB in size.

Table 1: Environmental variables used for modelling

Data Category	Data type	Layer Name	Explanation of Layer
Terrain and Topographic	Continuous	Dem	Digital elevation model
Terrain and Topographic	Continuous	Slope	Slope derived from DEM
Terrain and Topographic	Continuous	Aspect	Aspect
Terrain and Topographic	Continuous	Top100	Topographic position 100m
Terrain and Topographic	Continuous	Top250	Topographic position 250m
Terrain and Topographic	Continuous	Top500	Topographic position 500m
Terrain and Topographic	Continuous	Top1000	Topographic position 1000m
Terrain and Topographic	Continuous	Rug100	Ruggedness 100m
Terrain and Topographic	Continuous	Rug250	Ruggedness 250m
Terrain and Topographic	Continuous	Rug500	Ruggedness 500m
Terrain and Topographic	Continuous	Rug1000	Ruggedness 1000m
Climate	Continuous	Maxtemp	Average maximum temperature
Climate	Continuous	Mintemp	Average minimum temperature
Climate	Continuous	Rainfal	Average annual rainfall
Drainage	Continuous	Allstrm	Distance to waterway (all streams)
Drainage	Continuous	Mjstrm	Distance to waterway (major streams only)
Soil	Categorical	Acidity	Acidity
Soil	Categorical	Perm050	Permeability 0-50cm
Soil	Categorical	Perm150	Permeability 0-150cm
Soil	Categorical	Salinity	Salinity
Soil	Categorical	Wlog	Water logging
Soils	Categorical	Wstore	Water Storage
Vegetation	Categorical	Preveg	Pre-European Vegetation
Vegetation	Categorical	Remveg	Remnant Vegetation

2.3 MODEL EVALUATION

All models produced were evaluated by assessing 1) the predicted distribution of the species, 2) validation and cross-validation statistics, and 3) the contribution of the variables (or predictors) to final models. These are described below:

- Distribution – the mapped prediction surface for each model produced was reviewed with regards to its match to recorded species location, locations and habitats selected, and knowledge of the potential use of such locations and habitat types;
- Validation and Cross-validation statistics – the correlation between the actual values and values predicted by each model (ROC test used), and the graph shape for validation and cross-validation were reviewed to determine the strength of each model. A strong model was one which had a high ROC value (close to 1) and a near perpendicular shape.
- Variable contribution and utilisation – the variables used in each model were reviewed for their contribution to the model (alone and within the final model) and their importance to the model (variable cannot be compensated by other variables if dropped from the model). The best models were determined based on the variable's perceived relevance and importance to determining species distribution.

After undertaking a number of modelling attempts for each species with different input data and using different techniques, an evaluation of the best model was undertaken using the above criteria. The best model was chosen and these "final" models are those discussed in more detail in the following sections. The models are represented by both the habitat maps (showing predicted relative probability of occurrence) and statistics and graphs of the model in Appendices B, C and D.

A broad evaluation ranking of very high, high, moderate, low or poor has been given to each model to represent the perceived strength in the model.

3 Results

3.1 SPLENDID FAIRY-WREN

3.1.1 Environmental Space Occupied by Splendid Fairy-wren

Initial examination of the environmental space (the habitat preferences and spatial distribution of the species) occupied by Splendid Fairy-wren was undertaken using specially designed histograms (see Figures B1 and B2, Appendix B). This examination revealed that Splendid Fairy-wren is often found close to streams, at low elevations, in gentle terrain, in low topographic positions (ie. gullies) and in flat areas, and where average maximum temperatures were moderate (around 25ºC), average minimum temperatures were high (over 11ºC), and average annual rainfall was moderate (around 800mm). Further, Splendid Fairy-wren was found in areas where pH was slightly acid, in vegetation classified as woodland and forest, and where salinity was low to moderate.

3.1.2 Analysis of Models

The stepwise modelling procedure, using input variables thought to potentially have a relationship with Splendid Fairy-wren presence, yielded models that consistently selected dem, maximum temperature, distance to all streams and aspect as the predictors driving Splendid Fairy-wren distribution. The best model generated also selected these predictors, with dem unable to be compensated by other variables if dropped from the model (Figure B3, Appendix B). The generated predictive surface from the stepwise model was considered to be a moderate match to presence records, with the majority of the presence records showing a high probability of occurrence. Cross validation statistics were also moderate (Table 2; Figure B5, Appendix B).

Despite the moderate predictive model produced using the stepwise procedure, it was thought that other predictors not selected in the model were important in driving Splendid Fairy-wren distribution. As such, a full model only using a number of selected predictive variables was run (see Table 2). This model produced a predictive surface that was slightly better at matching presence records than the surface generated using the stepwise procedure (the majority of the presence records showed a high probability of occurrence, but some of the records had a higher probability of occurrence than the same records in the stepwise model). Cross validation statistics were moderate for the full model (Table 2; Figure B6, Appendix B). Like the model generated using the stepwise procedure, the predictor driving the full model was again dem (Figure B4, Appendix B).

As the full model matched Splendid Fairy-wren presence records better than the model generated using the stepwise procedure, the former was considered to be a better model overall. Further, the full model was considered to be a better reflection of Splendid Fairy-wren distribution as it included predictor variables that were perceived to be of the greatest importance to the species (eg. ruggedness, topographic position, vegetation, temperature and rainfall).

3.1.3 Final Model

The final predictive surface for Splendid Fairy-wren, generated using the full model, is shown in Figure 2. Partial response curves for Splendid Fairy-wren presence records against predictors in the full model are shown in Figure B7 in Appendix B.

Generally, the final predictive surface predicted areas closer to the coast better than areas further from the coast, with presence records showing a higher probability of occurrence within the predictive surface closer to the coast. This appeared to be an accurate reflection of Splendid Fairy-wren distribution. Even so, the final predictive surface showed areas with higher probabilities of occurrence in the southern coastal areas compared to in the northern coastal areas. This was not reflected in the distribution of Splendid Fairy-wren presence records, which showed a relatively even spread of presence records in the northern and southern coastal areas. As such, the final predictive surface may have somewhat under-predicted the potential habitat / distribution of Splendid Fairy-wren.

In relation to the Splendid Fairy-wren, the final predictive surface could be improved if more accurate information on vegetation structure were available. While a vegetation variable was included in the final model which was used to create the predictive surface, information in this variable was broad and could be improved with more detailed information on whether a shrub layer was present. Further, field validation could be conducted to ground truth the final model. More general improvements to generating a more accurate predictive surface are provided in Section 6.

Table 2: Summary of the final predictors in the best models (stepwise and full) produced for Splendid Fairy-wren, and model evaluation.

Model	Predictors in model	Strength of model (validation and cross validation statistics)	Predictive surface fit to presence records	Overall model strength
Stepwise model	dem, maxtemp, allstrm, aspect	0.842 and 0.846 (Moderate)	Moderate	Moderate
Full model	dem, rug100, top250, rainfal, maxtemp, allstrm, peveg, aspect	0.854 and 0.845 (Moderate)	Moderate - high	Moderate - High

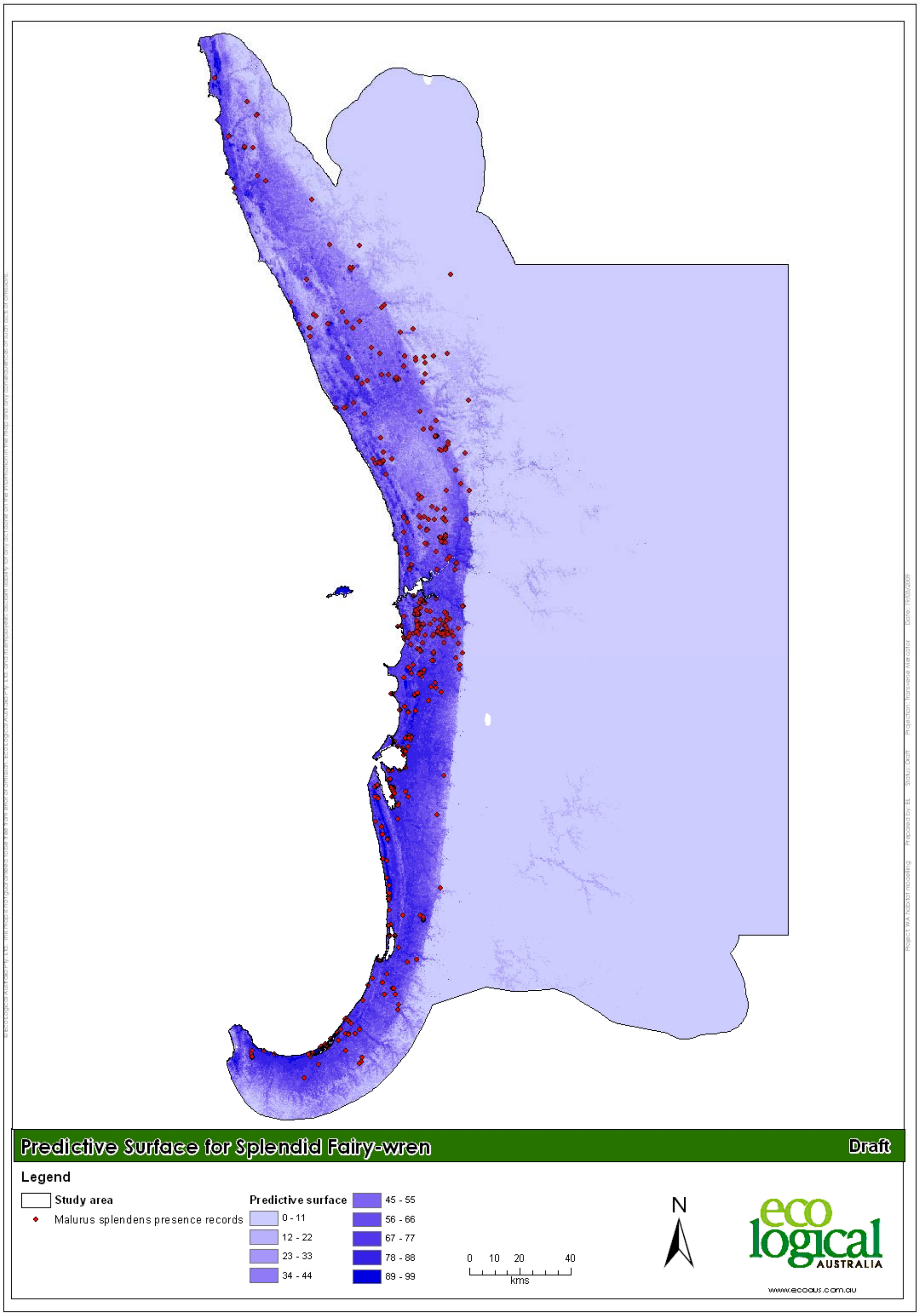

Figure 2: Habitat predictive map for Splendid Fairy-wren.

3.2 SCARLET ROBIN

3.2.1 Environmental Space Occupied by Scarlet Robin

Examination of the environmental space occupied by Scarlet Robin (see histograms in Figures C1 and C2, Appendix C) revealed that Scarlet Robin was found in very similar environmental space as Splendid Fairy-wren, that is, close to streams, at low elevations, in gentle terrain, in low topographic positions (ie. gullies) and in flat areas, where average maximum temperatures were moderate, average minimum temperatures were high, and average annual rainfall was moderate, in areas where pH was slightly acid, in vegetation classified as woodland and forest, and where salinity was low to moderate.

3.2.2 Analysis of Models

The best model generated for Scarlet Robin was produced using a stepwise procedure. The stepwise procedure, using input variables thought to potentially have a relationship with Scarlet Robin presence, yielded models that consistently selected dem, minimum temperature, maximum temperature, distance to all streams and aspect as predictors driving Scarlet Robin distribution. The best model generated predicted dem, distance to major streams, maximum temperature, distance to all streams, and aspect, as driving Scarlet Robin distribution, with dem unable to be compensated by other variables if dropped from the model (Figure C3, Appendix C).

The generated predictive surface was considered to be a good (high) match to presence records, with the majority of presence records showing a high probability of occurrence. Cross validation statistics were low (Table 3; Figure C4, Appendix C). As such, the overall model strength was assessed to be moderate - high.

While full models including predictors thought to be important in driving Scarlet Robin were run, the predictive surfaces generated were considered to be poorer matches for Scarlet Robin presence records, with areas without Scarlet Robin records showing a high probability of occurrence. In addition, validation and cross-validation statistics were low. Full models were assessed as having poor strength overall.

3.2.3 Final Model

The final predictive surface for Scarlet Robin is shown in Figure 3. Partial response curves for Scarlet Robin presence records against predictors in the final model are shown in Figure C5 in Appendix C.

The final predictive surface predicted that Scarlet Robin occurred in areas closer to the coast more so than in areas further from the coast. This appeared to be an accurate reflection of Scarlet Robin distribution. Even so, the final predictive surface showed areas along the coast with low probabilities of occurrence (near the north of the modelling area and in the far south of the modelling area). This may have been an accurate reflection of potential habitat / distribution for Scarlet Robin. However, dem, which was the main driver for the final predictive layer, was consistent in regards to elevation in this area. As such, the final predictive surface may have somewhat under-predicted the potential habitat / distribution of Scarlet Robin in these areas.

In relation to the Scarlet Robin, the final predictive surface could be improved if more accurate information on vegetation structure were available. No vegetation variables were included in the final model, but with more detailed information on whether a shrub layer was present, vegetation variables may have been included in models to produce a more accurate predictive layer. In addition to the inclusion of more detailed vegetation data, field validation could be conducted to ground truth the final model. More general improvements to generating a more accurate predictive surface are provided in Section 6.

Table 3: Summary of the final predictors in the model produced for Scarlet Robin, and model evaluation.

Model	Predictors in model	Strength of model (validation and cross validation statistics)	Predictive surface fit to presence records	Overall model strength
Stepwise model	dem, mjstrm, maxtemp, allstrm, aspect	0.76 and 0.737 Low	High	Moderate-High

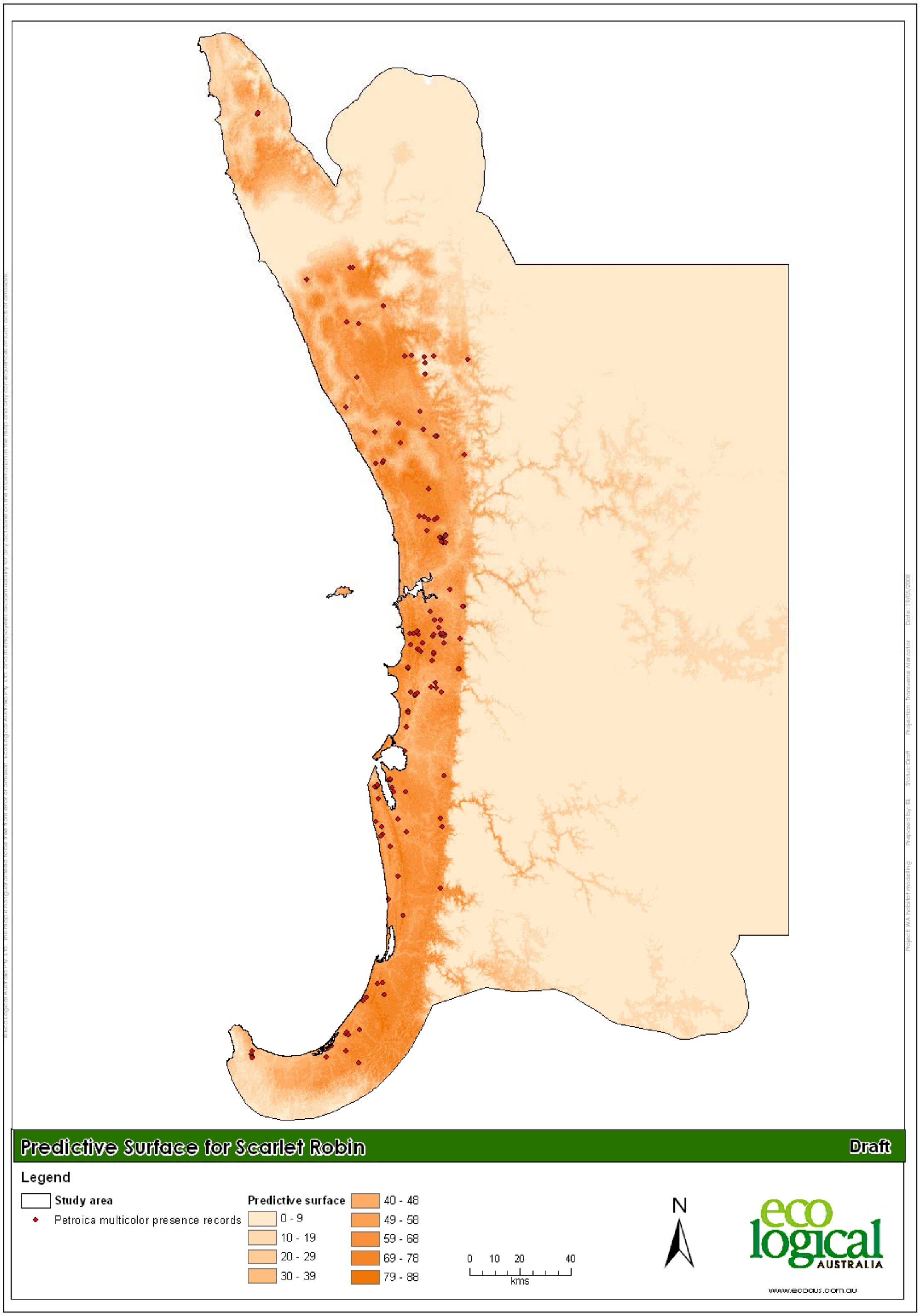

Figure 3: Habitat predictive map for Scarlet Robin.

3.3 WESTERN THORNBILL

3.3.1 Environmental Space Occupied by Western Thornbill

Examination of the environmental space occupied by Western Thornbill (see histograms in Figures D1 and D2, Appendix D) revealed that Western Thornbill was found in very similar environmental space as Splendid Fairy-wren and Scarlet Robin, that is, close to streams, at low elevations, in gentle terrain, in low topographic positions (ie. gullies) and in flat areas, where average maximum temperatures were moderate, average minimum temperatures were high, and average annual rainfall was moderate, in areas where pH was slightly acid, in vegetation classified as woodland and forest, and where salinity was low to moderate.

3.3.2 Analysis of Models

The best model generated for Western Thornbill was produced using a stepwise procedure with the minimum contribution of predictors set to 1%. Poorer models were generated when the minimum contribution of predictors was set to 5%, and with full models including predictors thought to have the potential for affecting Western Thornbill presence. The best stepwise model predicted distance to all streams, dem, minimum temperature, rainfall, distance to major streams, maximum temperature, and aspect, as driving Western Thornbill distribution, with dem, minimum temperature and rainfall unable to be compensated by other variables if dropped from the model (Figure D3, Appendix D).

The generated predictive surface was considered to be a good (high) match to presence records, with the presence records in the northern part of the study area showing a high probability of occurrence and records in the southern part of the study area showing a moderate probability of occurrence. Cross validation statistics were low (Table 4; Figure D4, Appendix D). As such, the overall model strength was assessed to be moderate - high.

While full models including predictors thought to be important in driving Western Thornbill were run, the predictive surfaces generated were poor matches for Western Thornbill presence records, with areas without Western Thornbill records showing a high probability of occurrence. In addition, validation and cross-validation statistics were low. Full models were assessed as having poor strength overall.

3.3.3 Final Model

The final predictive surface for Western Thornbill is shown in Figure 4. Partial response curves for Western Thornbill presence records against predictors in the final model are shown in Figure D5 in Appendix D.

Like models for Splendid Fairy Wren and Scarlet Robin, the final predictive surface showed that Western Thornbill occurred in areas closer to the coast more so than in areas further from the coast. In addition, the final predictive surface showed that southern parts of the coast had a lower probability of Western Thornbill occurrence than northern parts of the coast. This was reflected in the distribution of Western Thornbill presence records, which showed more presence records to the northern coastal area than the south. This appeared to be an accurate reflection of Western Thornbill distribution.

In relation to Western Thornbill, the final predictive surface could be improved if more accurate information on vegetation structure were available. No vegetation variables were included in the final model, but with more detailed information on whether a shrub layer was present, vegetation variables may have been included in models to produce a more accurate predictive layer. In addition to the inclusion of more detailed vegetation data, field validation could be conducted to ground truth the final model. More general improvements to generating a more accurate predictive surface are provided in Section 6.

Table 4: Summary of the final predictors in the model produced for Western Thornbill, and model evaluation.

Model	Predictors in model	Strength of model (validation and cross validation statistics)	Predictive surface fit to presence records	Overall model strength
Stepwise model	allstrm, dem, mintemp, rainfall, mjstrm, maxtemp, aspect	0.77 and 0.71 Low	High	Moderate - High

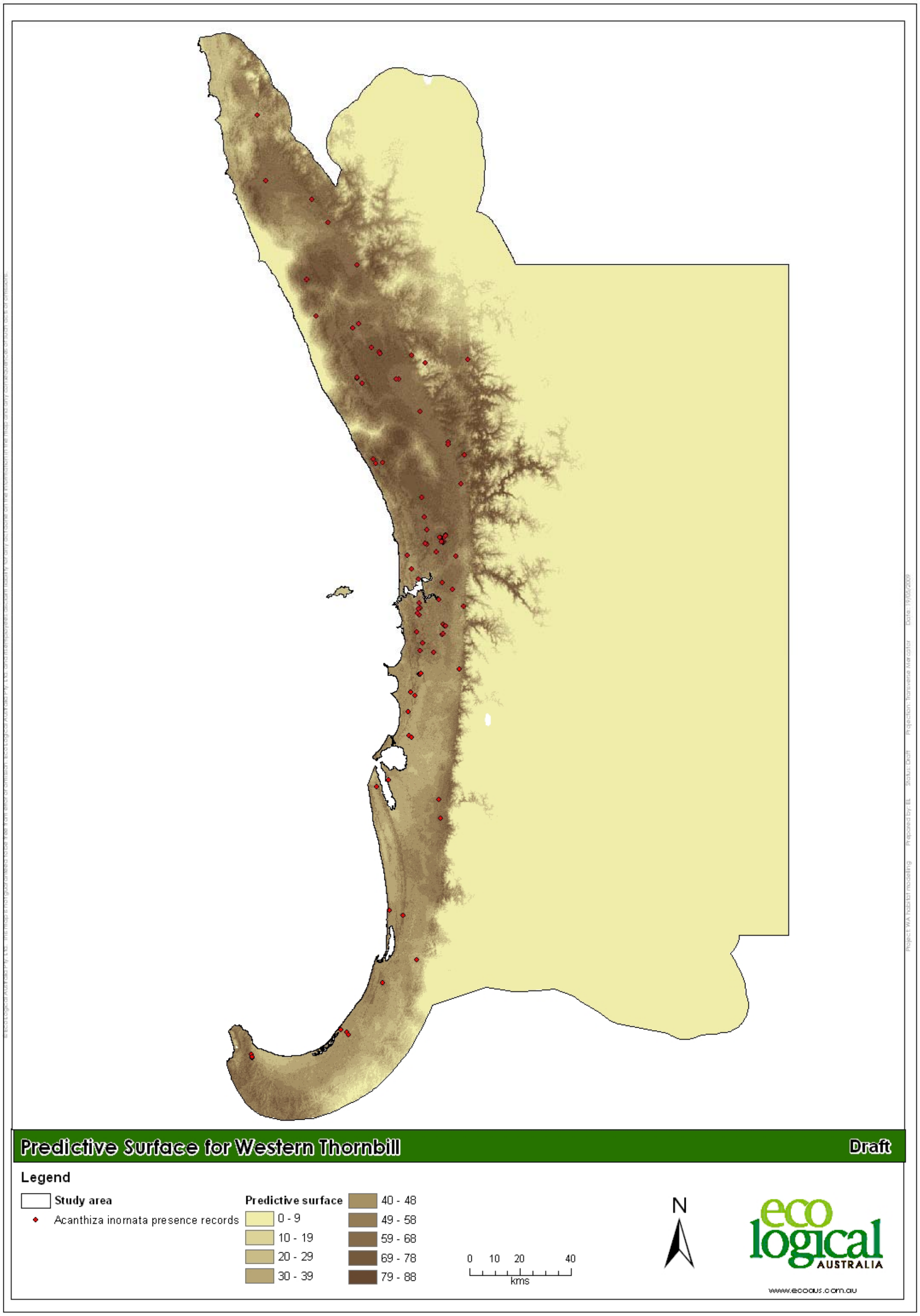

Figure 4: Habitat predictive map for Western Thornbill.

4 Model Limitations

The models produced show areas of the landscape that contain potential habitat for Splendid Fairy-wren, Scarlet Robin and Western Thornbill. However, as for all models, the models generated were based on the data that were input into the system and these have inherent limitations.

- Survey effort and therefore the records of species observations may have been biased to particular areas, which could have caused the model to only predict areas with the same environmental information in these areas as potential habitat;
- No data was available for known species absence. To overcome the lack of this type of data, pseudo-absence data was used in the modelling which is inherently limited and cannot inform the modelling to the same degree as the inclusion of actual absence data; and
- The modelling is also limited by the type, accuracy/scale and other characteristics of the available environmental data, specifically with regard to how well the data can explain the distribution of potential species habitat. For example, it is noted that fine scale data describing vegetation type, condition, structure and other characteristics was not available.

In addition to the models being limited by available data, evaluation of models was constrained by the limited information available on the specific habitat preferences of the species modelled. A large proportion of model evaluation is based on knowledge of species habitat preferences. Available information on the habitat preferences of the three bird species was limited to information on the broad vegetation types and floristics preferred by the three species, as well as foraging and nesting habitat preferences (in relation to vegetation structure eg. Robinson 1992; Tibbets and Pruett-Jones 1999). Further, there was even less information on the habitat preferences of Western Thornbill. Arnold (1988) was unable to find a relationship between Western Thornbill and vegetation structure. No information was found on how the three bird species responded to other environmental characteristics (such as those represented by variables for ruggedness, topographic position, distance to waterways, climate variables, and soil-derived variables), although it is understood that these predictor variables often have a significant influence on the characteristics of vegetation and thus the presence or absence of habitat for the bird species. Without increased knowledge of species habitat preferences, analysis of model strength must by necessity, be based less on the perceived relevance of predictors and more on the matching of presence records and validation and cross-validation statistics.

Overall, the models should be treated as indicative only, highlighting those parts of the landscape where there is potentially a higher probability of the species being present (or at least utilised during some part of its life cycle).

5 Conclusions

ELA was contracted by the DEC to undertake a predictive habitat modelling pilot study for up to twelve flora and fauna species in the Swan Region, Western Australia. Habitat modelling occurred in two stages: in the first stage, the potential habitats of two species were modelled (with an additional 2 species unable to be modelled due to lack of species record data), while in the second stage, the potential habitats of three bird species were modelled.

Models that were previously generated for the stage 1 species, Waxy-leaved Smokebush and the Graceful Sunmoth, were both evaluated as high.

For the three bird species in the Swan region, modelled in stage 2 (this report), climate data was available for use as an additional environmental variable. Indeed, at least one of the climate variables introduced into stage 2 was selected in the final model for each species. Further, higher numbers of presence records for the bird species were available in which to model relationships with predictor variables, which no doubt assisted the modelling process. The models produced for the three bird species were evaluated as moderate - high.

5.1 RECOMMENDATIONS

Given the experience and results of the modelling completed for stage 1 and stage 2, a number of recommendations can be made for the purposes of improving future modelling exercises (for these or additional species) as well as for utilising the current models to their full value. Recommendations include:

1. Undertaking selective field validation of the models to ground truth them and also to assist with assigning a cut off or classification to the relative probability (discussed below);

2. Utilising the models to guide further survey effort, particularly in those areas modelled as higher probability habitat but where the species has not yet been recorded from. Further survey data collected should aim to:

    a. Collect presence and absence data as there is value in understanding where the species is not found;
    b. Ensure proper data collection and management protocols are established and maintained to ensure data collected has a high degree of accuracy and usability for different projects. As an example, many of the species records provided by the DEC for stage 2 were excluded from the models due to uncertainty of their accuracy (48% of Splendid Fairy-wren records, 57% of Scarlet Robin records, and 59% of Western Thornbill records). This could have been avoided had locational accuracy and datum information been properly recorded;

3. Involve where possible a species expert to advise the modelling process and also to provide review and feedback on the output models;

4. Seek to acquire or develop/produce additional environmental datasets. In analysis of the models it was suggested that fine scale vegetation related datasets may have greatly assisted the modelling process, particularly for the chosen species. As such, it is proposed that the DEC

(perhaps in collaboration with other agencies) take a lead role in the development of such key datasets for natural resource management;

5. Undertake modelling exercises for different species types;

6. Model to the most relevant study area for the species chosen. The modelling study area selected (by DEC) for this pilot study covered a much larger geographic range than the species selected for modelling. Given the difficultly in adjusting the modelling area it is recommended that future modelling exercises consider all the species wishing to be modelled prior to selecting the modelling study area (such that the most appropriate area can be selected);

7. Apply treatments to the prediction surface of the final models (GIS raster data) to ensure they are used appropriately:

    a. Determine a cut off point in the relative probability and exclude all values beneath this level; and/or
    b. Classify the relative probability values into broad groups (i.e. high to low); and/or
    c. Mask out areas to exclude, such as:
        i. Unvegetated areas;
        ii. Area (i.e. certain portions of the study area); or
        iii. Based on the values of one or more contributing predictor variables.

8. Integrate the modelling into DEC core business, including (but not limited to):

    a. Strategic conservation planning;
    b. Development control;
    c. Survey effort;
    d. Species conservation and management;
    e. Reserve acquisition; and
    f. Reserve management.

References

Arnold, G. W. (1988). The effects of habitat structure and floristics on the densities of bird species in Wandoo Woodland. *Australian Wildlife Research*. **15:** 499-510.

Eco Logical Australia (2008). *Predictive Habitat Modelling for Flora and Fauna in the Swan Region*. Report prepared for the WA Department of Environment and Conservation.

Lehman A., Leathwick J.R. and Overton J.McC. (2004) GRASP v3.0 User's Manual. Swiss Centre for Faunal Cartography, Switzerland.

Lehmann A., Overton J.McC. & Leathwick, J.R. (2002) GRASP: Generalised regression analysis and spatial predictions, Ecological Modelling, 157: 189-207.

Robinson, D. (1992). Habitat use and foraging behaviour of the scarlet robin and the flame robin at a site of breeding-season sympatry. *Wildlife Research*. **19:** 377 – 395.

Tibbet, E and Pruett-Jones, S. (1999). Habitat and nest-site partitioning in splendid and variegated fairy-wrens (Aves : Maluridae). *Australian Journal of Zoology.* **47:** 317 – 326.

Appendix A

STEP-BY-STEP BREAKDOWN OF MODELLING PROCEDURE

The modelling procedure used was as follows:

1. Define the study area;
2. Generate environmental variables to cover the study area;
3. Collate records of species locations within the study area;
4. Undertake a data cleaning exercise on the species records in order to remove records that may bias the modelling procedure such as those with:
    a. A low recorded positional accuracy;
    b. Duplicate records;
5. Randomly generate pseudo absence records;
    - Pseudo absence sites were generated and joined to the presence sites (recorded species locations) for each species;
6. Analyse each environmental variable for all site locations:
    - Create a matrix showing a row for each species and a column for each environmental variable;
7. Analyse the entire environmental space of the study area by recording the full extent of each environmental variable;
    - Note the maximum and minimum value of each variable within the study area;
8. Initialise textual data into GRASP (running under S-Plus) and analyse;
    - Generally, a “Step-wise” GAM and an “F Test” was the model type used for modelling species’ responses to predictors. Minimum contribution was set to 5% and max correlation between predictors set at 75%;
    - Full GAM models were used to model species’ responses to predictors if interpretation of models generated using a step-wise procedure revealed the model strength to be poor or to investigate forcing the inclusion of other predictor variables;
9. Interpret and validate each model;
    - Review the outputs of the modelling procedure (graphs showing contributing variables selected, validation and cross-validation etc);
10. Output lookup tables from GRASP for selected models, which are then used in ArcView to generate prediction surfaces for each species (using an extension under ArcView);
11. Generate predictive surfaces in ArcView and evaluate models;
12. Adjust and rerun models where necessary.
    - Modelling was undertaken using an iterative approach for each species. That is, broad relationships were identified between species locations and habitat variables using a stepwise procedure (an automated procedure within the software that seeks to identify the strongest relationships). Initial “draft” models were further refined through the incorporation of additional variables and/or removal of others until what was considered the best model possible was generated.

Appendix B

ENVIRONMENTAL SPACE OCCUPIED BY SPLENDID FAIRY-WREN

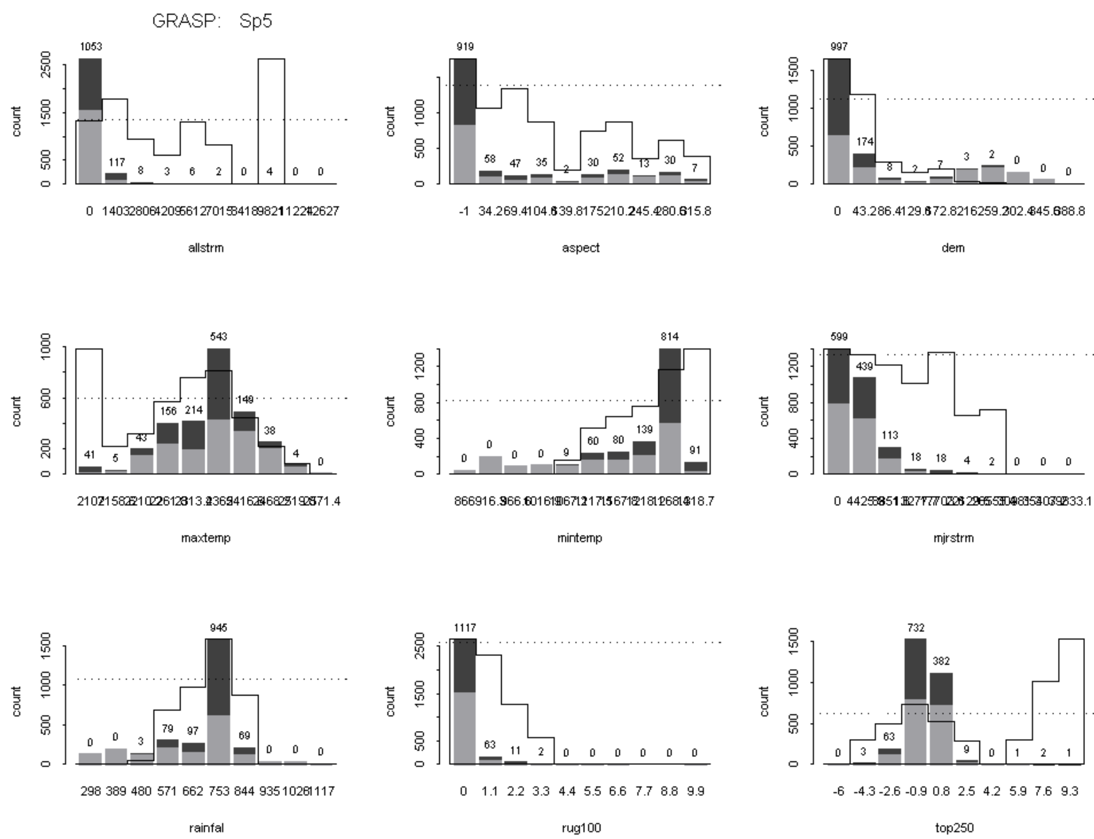

Figure B1: Distribution of Splendid Fairy-wren presence records in environmental space (only uncorrelated continuous variables shown). Entire histogram bars represent the distribution of presence and pseudo-absence records and, where darker areas represent the number of presence records for the species in each bar. This number is also written on top of each bar. The plain line is the ratio between presence and pseudo-absence records and the dashed line corresponds to the overall mean proportion of presence records.

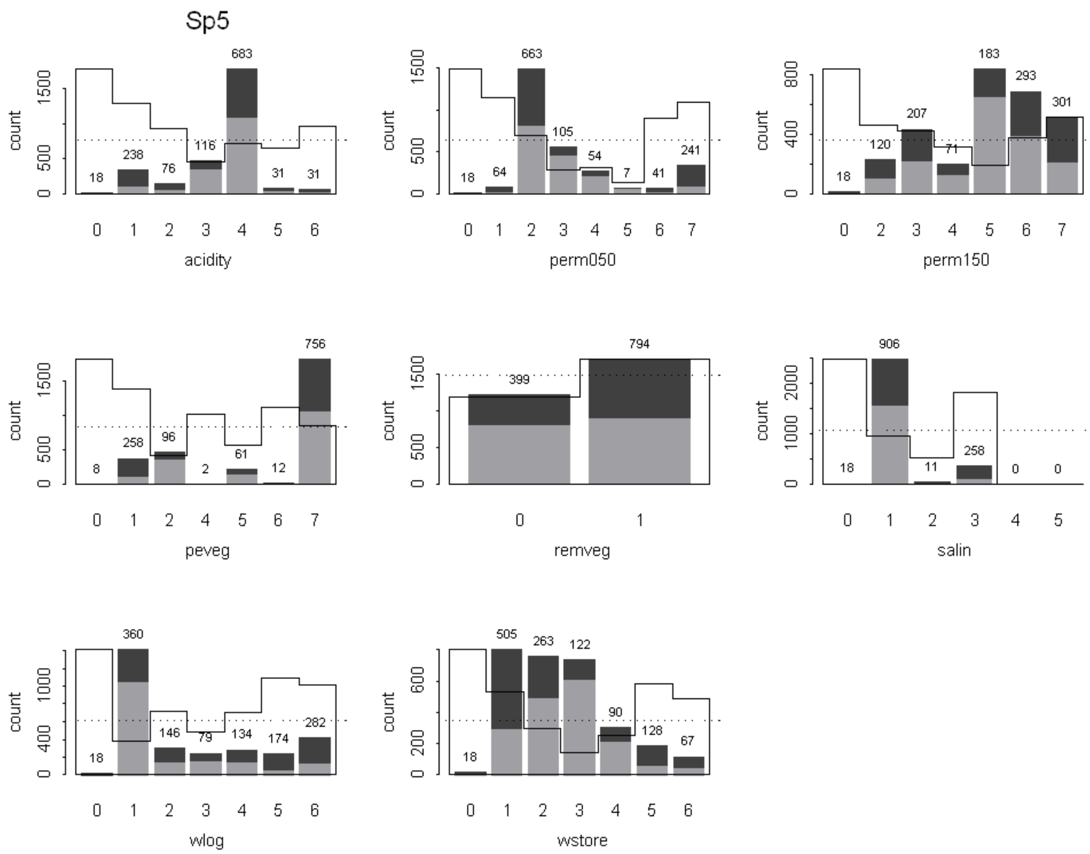

Figure B2: Distribution of Splendid Fairy-wren presence records in environmental space (categorical variables only). Entire histogram bars represent the distribution of presence and pseudo-absence records and, where darker areas represent the number of presence records for the species in each bar. This number is also written on top of each bar. The plain line is the ratio between presence and pseudo-absence records and the dashed line corresponds to the overall mean proportion of presence records.

CONTRIBUTION OF SELECTED PREDICTORS IN MODELLING SPLENDID FAIRY-WREN DISTRIBUTION

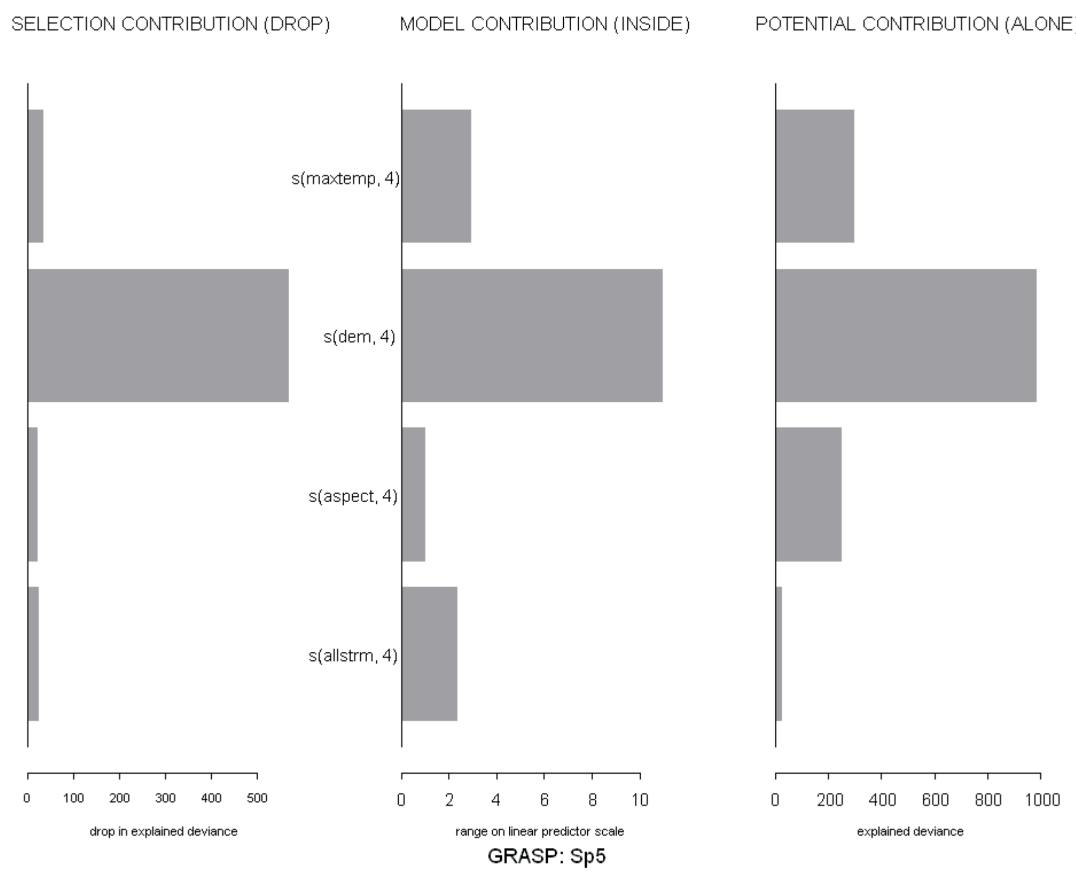

Figure B3: Contribution of selected predictors in modelling Splendid Fairy-wren distribution when dropped from the model, within the model, and on their own (alone). The model was generated using the stepwise procedure.

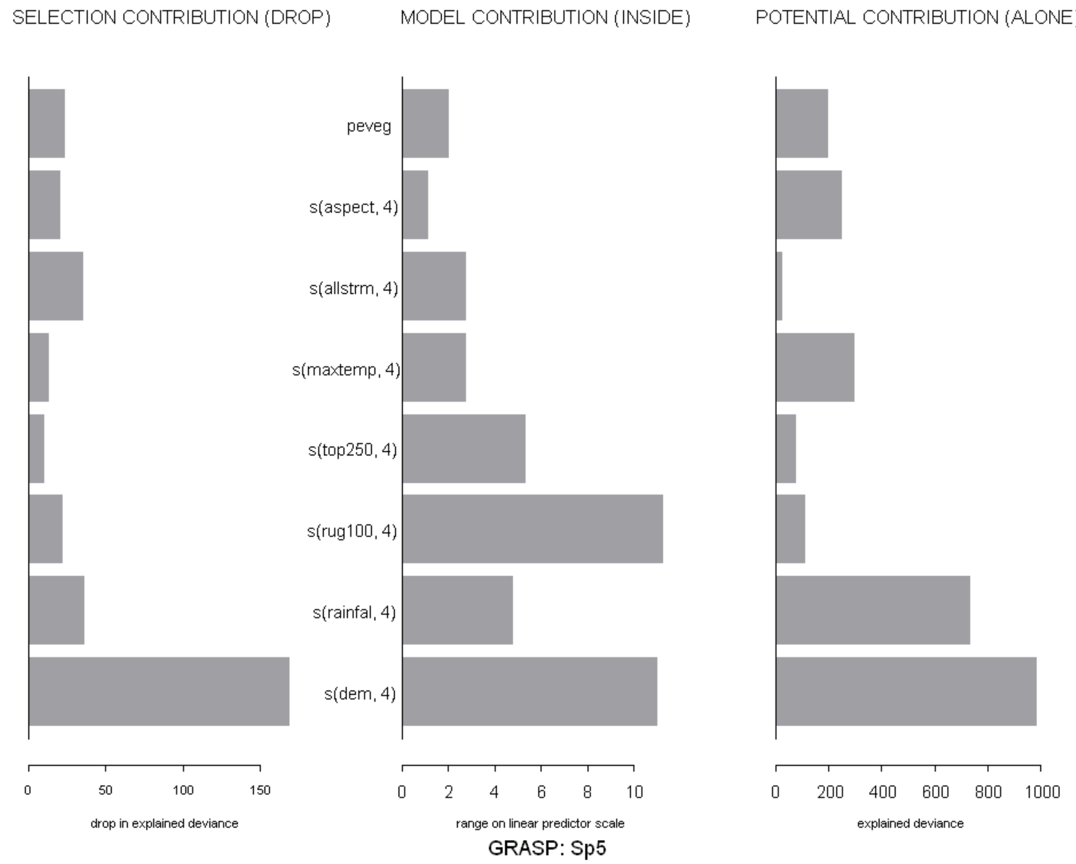

Figure B4: Contribution of selected predictors in modelling Splendid Fairy-wren distribution when dropped from the final model, within the final model, and on their own (alone). The final model was generated using a full model using all predictors shown (the stepwise procedure was not used).

VALIDATION AND CROSS-VALIDATION STATISTICS AND GRAPH SHAPES FOR SPLENDID FAIR-WREN MODELS

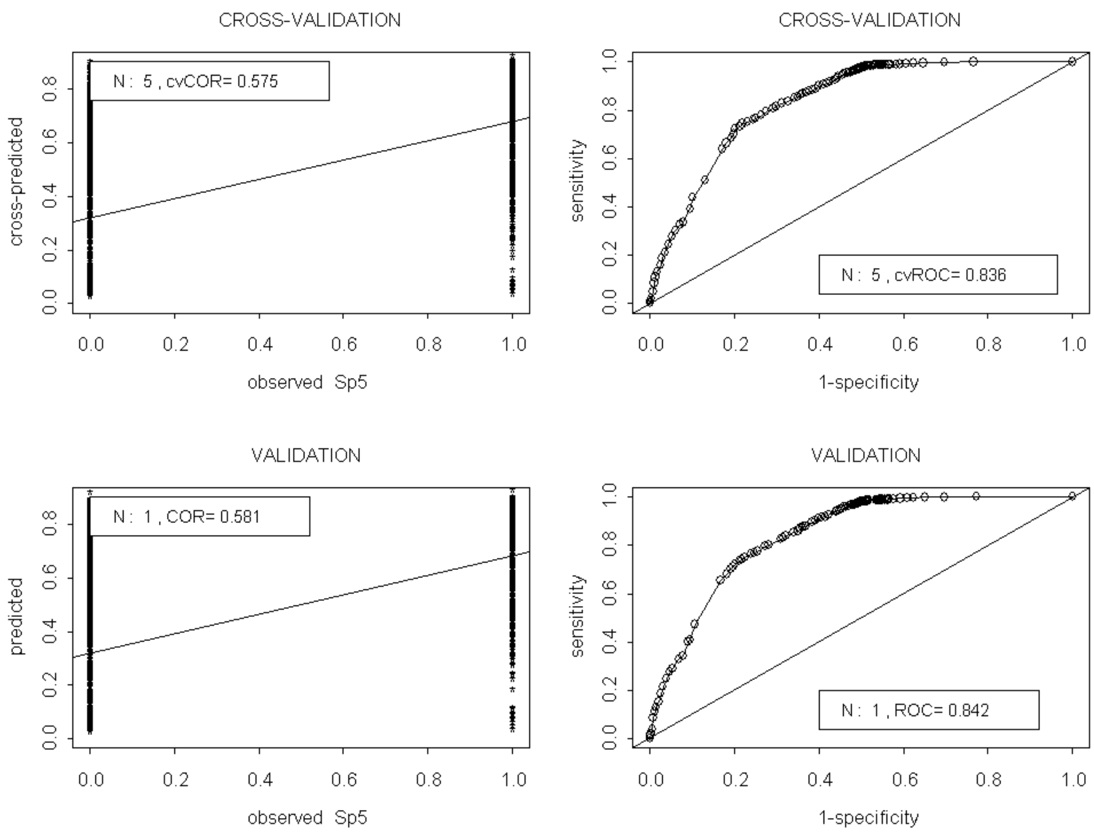

Figure B5: Validation and Cross-validation statistics and graph shapes for Splendid Fairy-wren model, generated using the stepwise procedure.

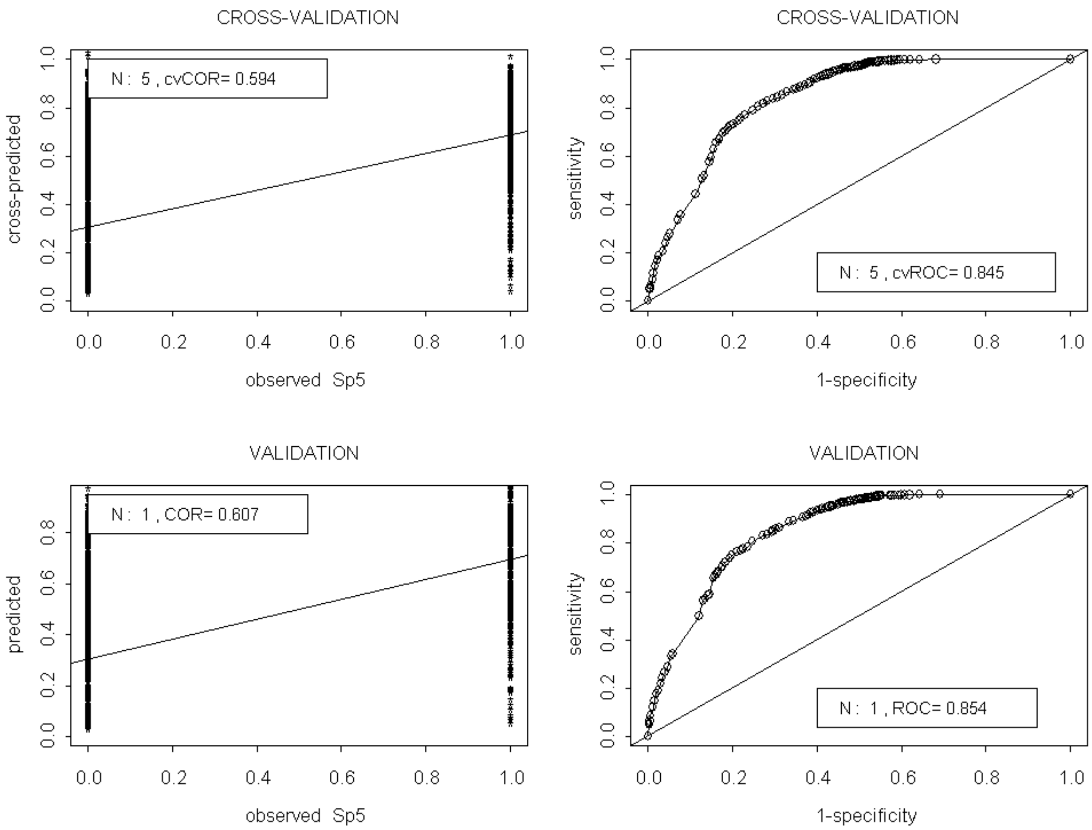

Figure B6: Validation and Cross-validation statistics and graph shapes for Splendid Fairy-wren model, generated from the full model.

PARTIAL RESPONSE CURVES FOR SPLENDID FAIRY-WREN PRESENCE RECORDS AGAINST PREDICTORS IN THE FULL MODEL

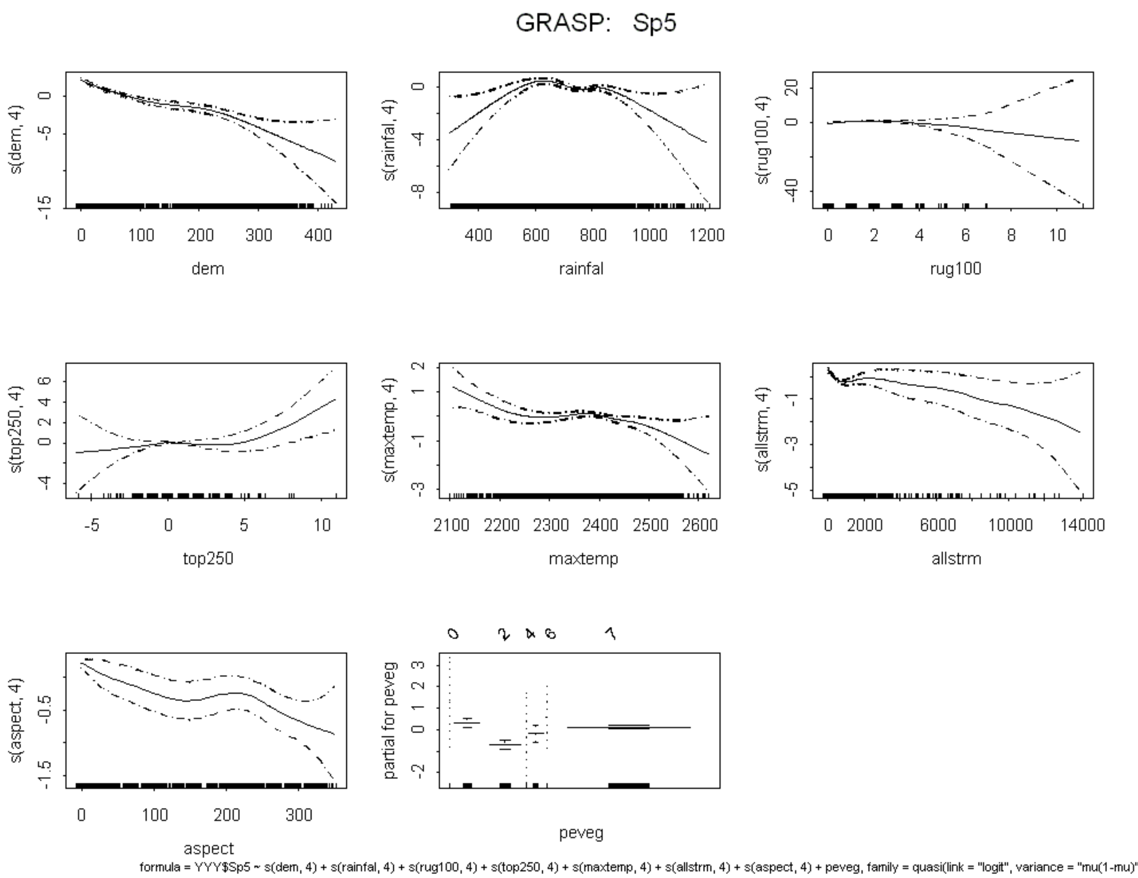

Figure B7: Partial response curves for Splendid Fairy-wren presence records against predictors in the full model.

Appendix C

ENVIRONMENTAL SPACE OCCUPIED BY SCARLET ROBIN

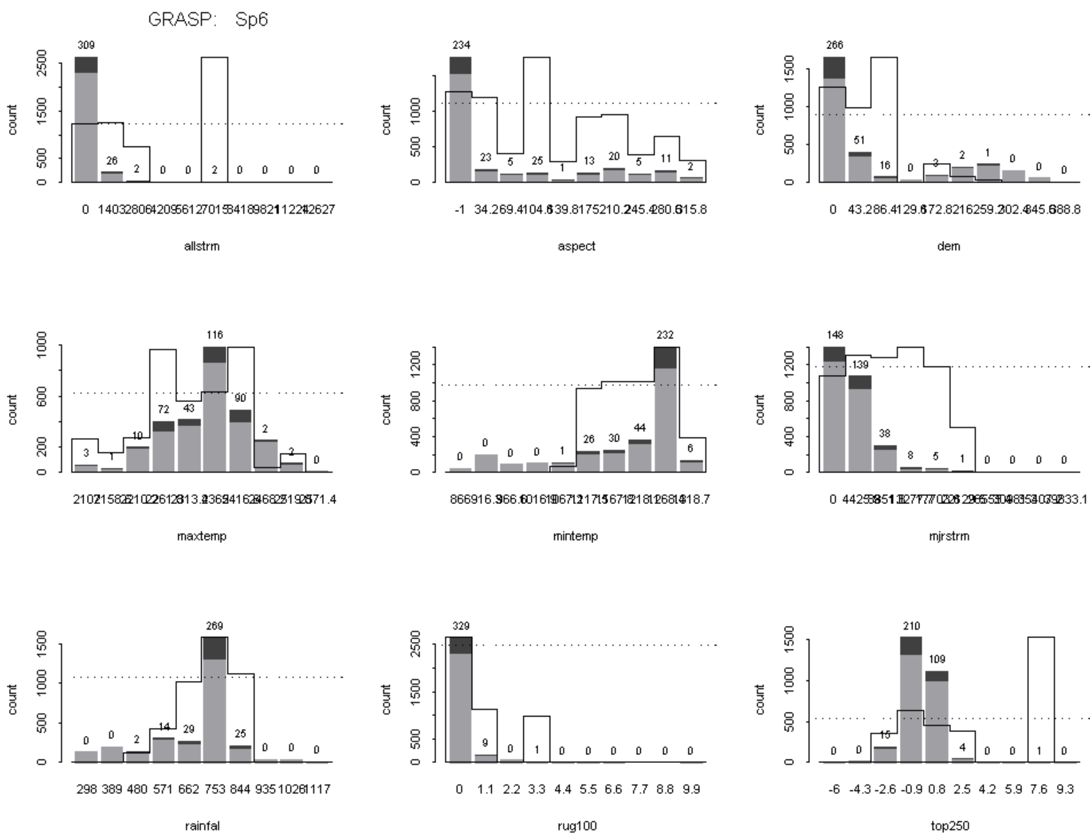

Figure C1: Distribution of Scarlet Robin presence records in environmental space (only uncorrelated continuous variables shown). Entire histogram bars represent the distribution of presence and pseudo-absence records and, where darker areas represent the number of presence records for the species in each bar. This number is also written on top of each bar. The plain line is the ratio between presence and pseudo-absence records and the dashed line corresponds to the overall mean proportion of presence records.

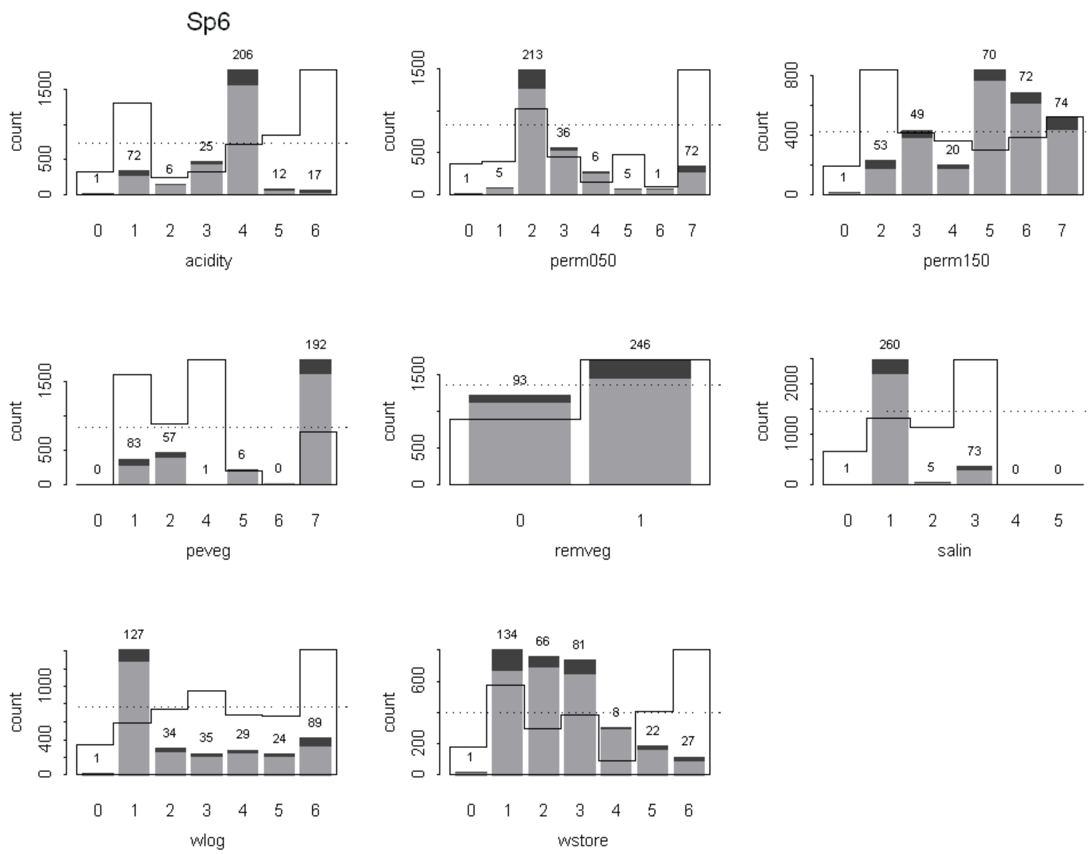

Figure C2: Distribution of Scarlet Robin presence records in environmental space (categorical variables only). Entire histogram bars represent the distribution of presence and pseudo-absence records and, where darker areas represent the number of presence records for the species in each bar. This number is also written on top of each bar. The plain line is the ratio between presence and pseudo-absence records and the dashed line corresponds to the overall mean proportion of presence records.

CONTRIBUTION OF SELECTED PREDICTORS IN MODELLING SCARLET ROBIN DISTRIBUTION

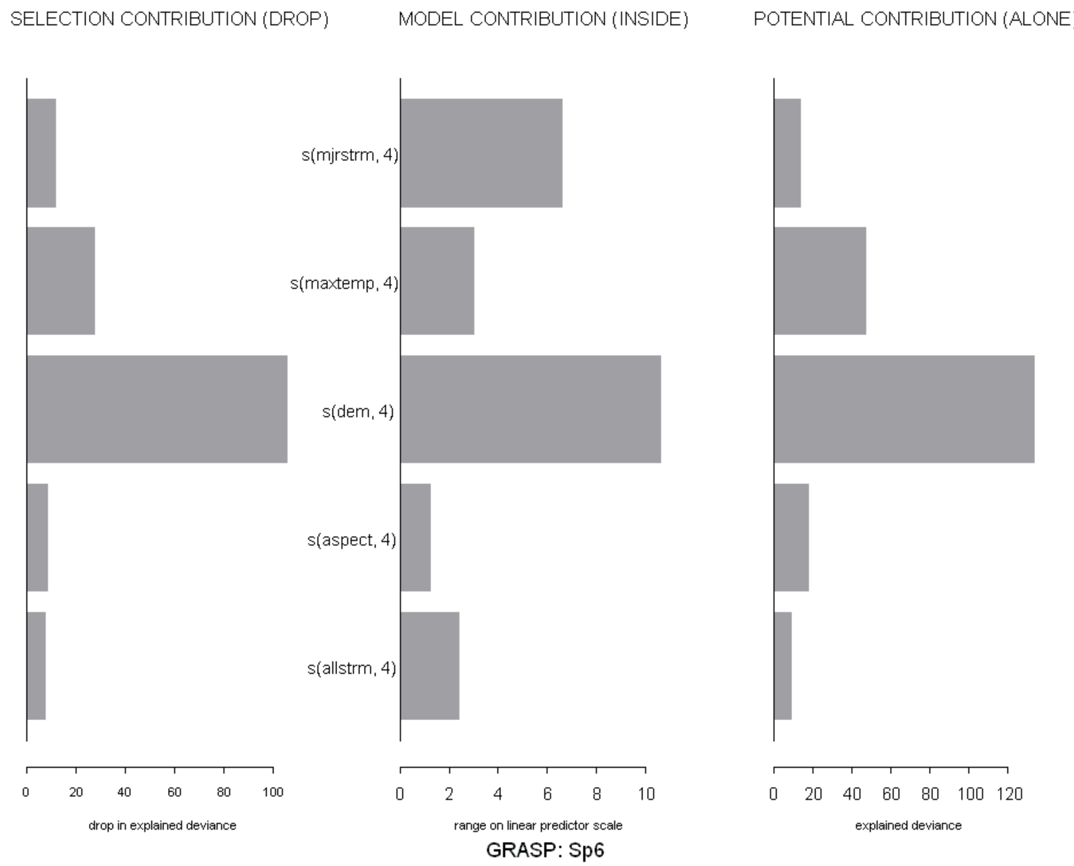

Figure C3: Contribution of predictors in modelling Scarlet Robin distribution when dropped from the final model, within the final model, and on their own (alone). The final model was generated using the stepwise procedure.

VALIDATION AND CROSS-VALIDATION STATISTICS AND GRAPH SHAPES FOR SCARLET ROBIN

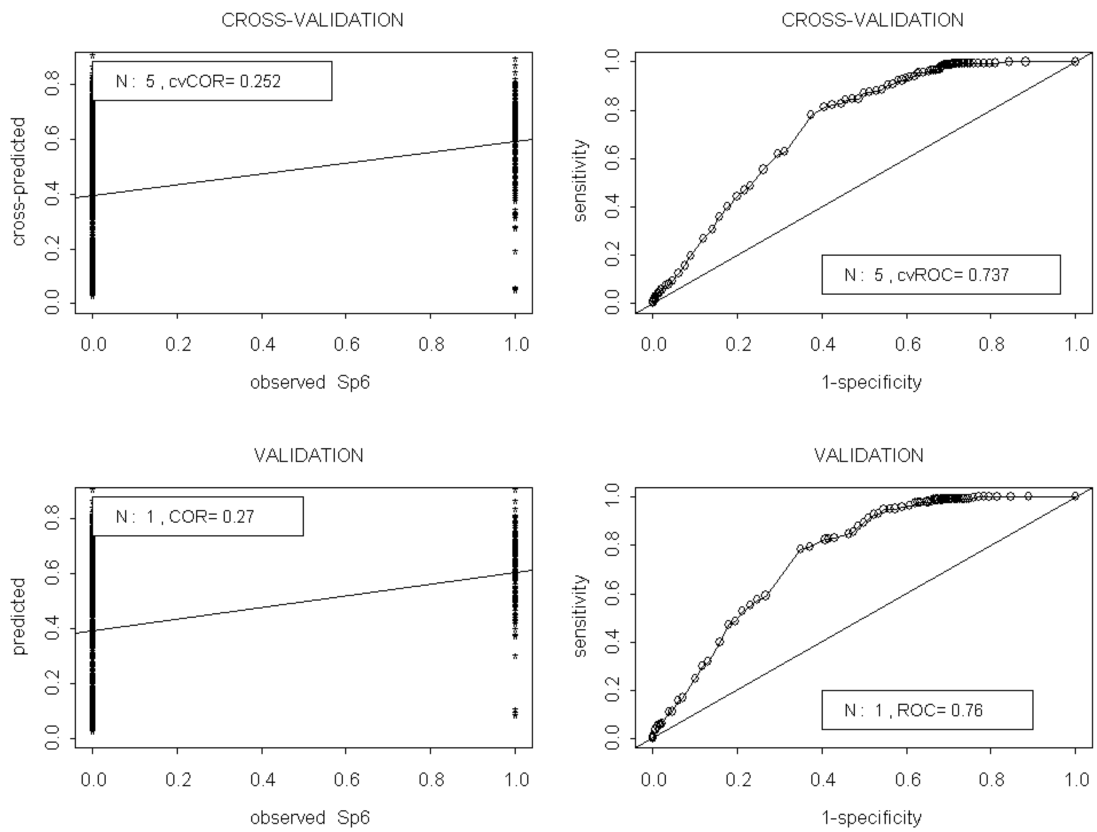

Figure C4: Validation and Cross-validation statistics and graph shapes for the final Scarlet Robin model.

PARTIAL RESPONSE CURVES FOR SCARLET ROBIN PRESENCE RECORDS AGAINST PREDICTORS IN THE FINAL MODEL

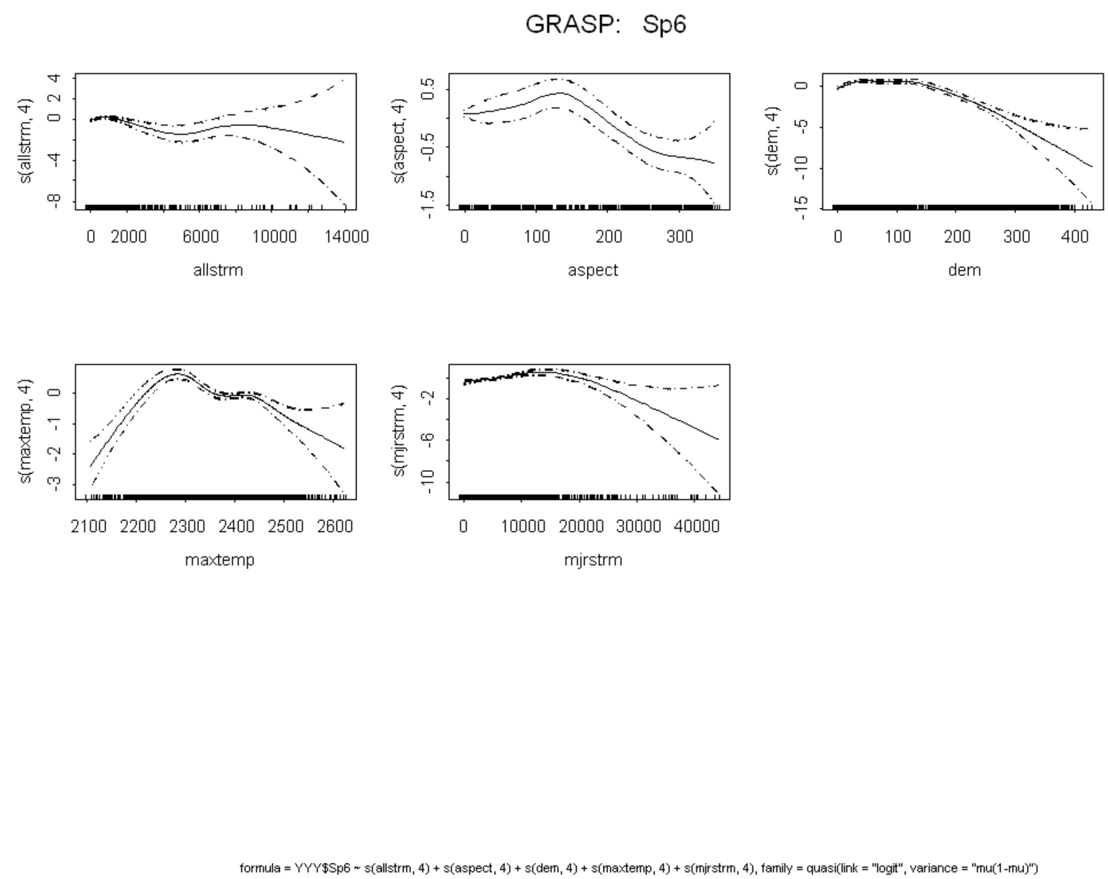

Figure C5: Partial response curves for Scarlet Robin presence records against predictors in the final model.

Appendix D

ENVIRONMENTAL SPACE OCCUPIED BY WESTERN THORNBILL

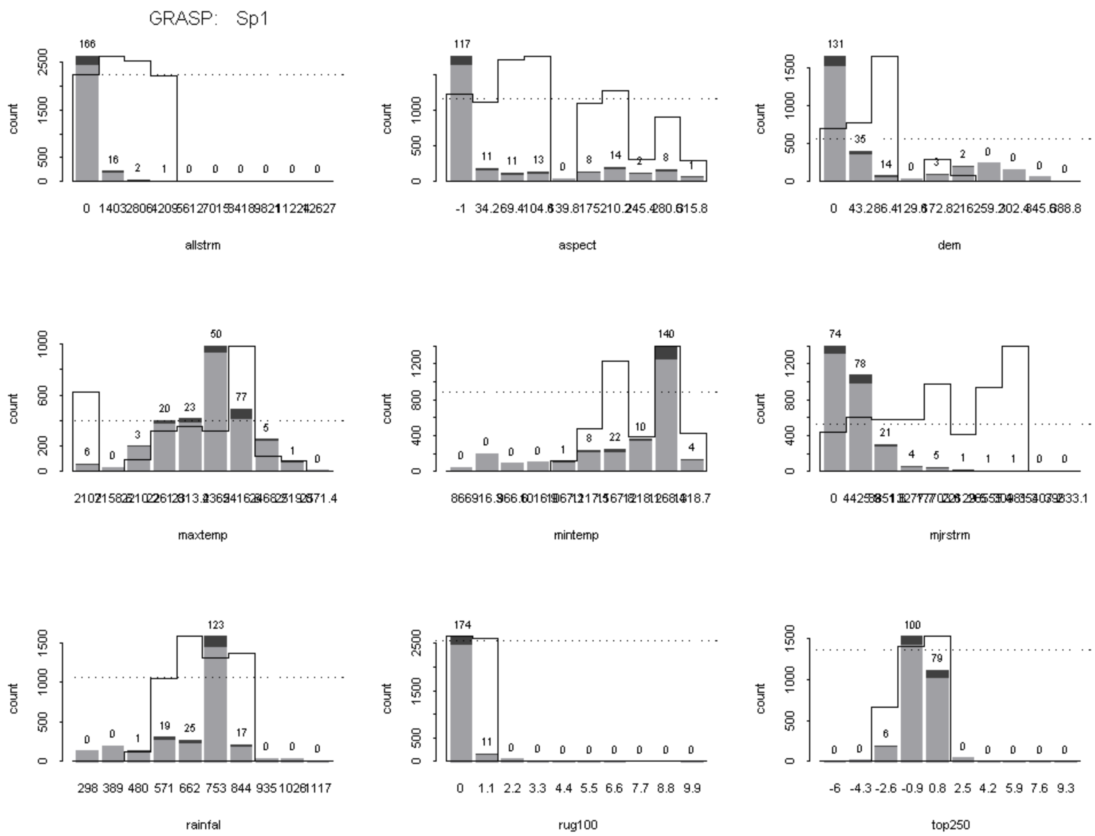

Figure D1: Distribution of Western Thornbill presence records in environmental space (only uncorrelated continuous variables shown). Entire histogram bars represent the distribution of presence and pseudo-absence records and, where darker areas represent the number of presence records for the species in each bar. This number is also written on top of each bar. The plain line is the ratio between presence and pseudo-absence records and the dashed line corresponds to the overall mean proportion of presence records.

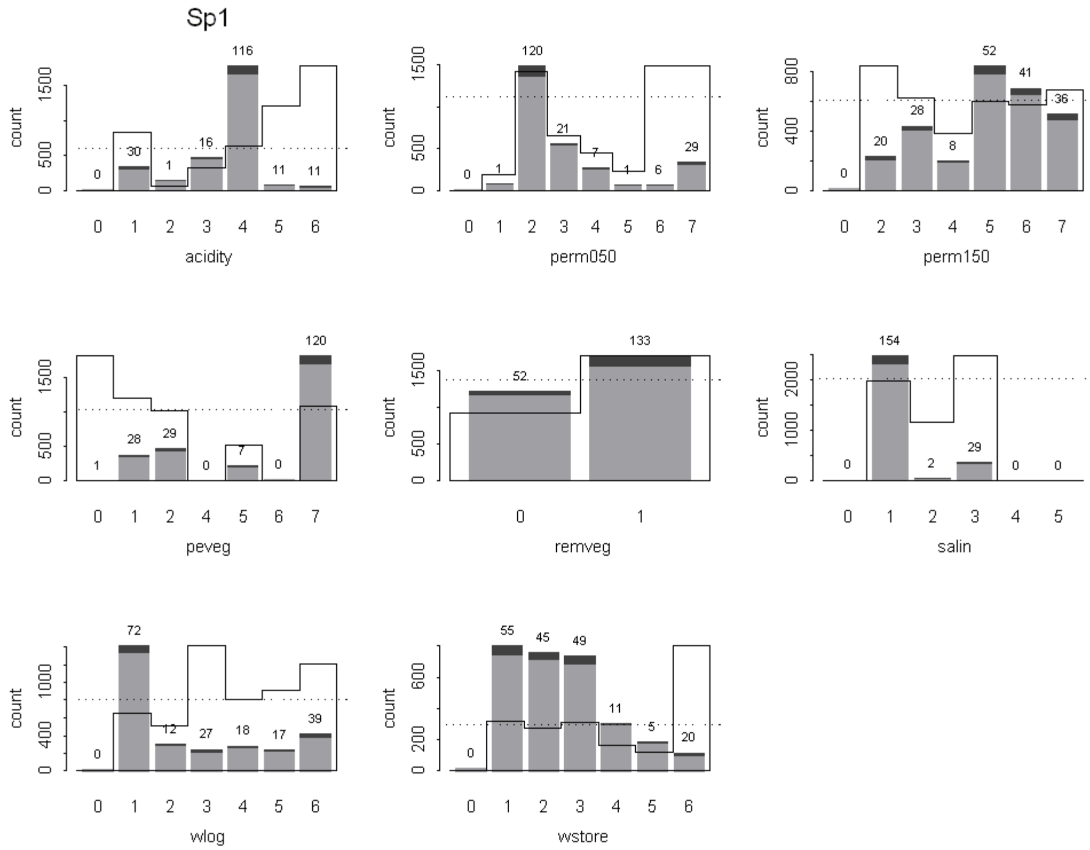

Figure D2: Distribution of Western Thornbill presence records in environmental space (categorical variables only). Entire histogram bars represent the distribution of presence and pseudo-absence records and, where darker areas represent the number of presence records for the species in each bar. This number is also written on top of each bar. The plain line is the ratio between presence and pseudo-absence records and the dashed line corresponds to the overall mean proportion of presence records.

CONTRIBUTION OF SELECTED PREDICTORS IN MODELLING WESTERN THORNBILL DISTRIBUTION

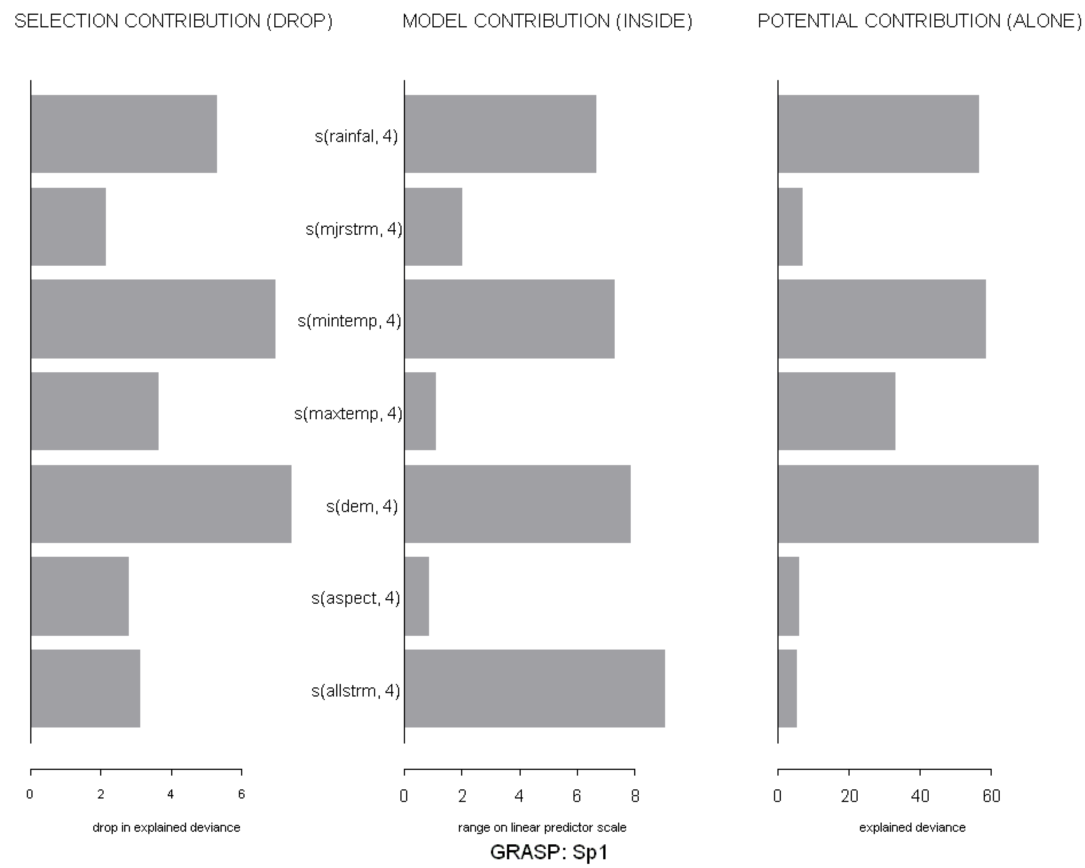

Figure D3: Contribution of predictors in modelling Western Thornbill distribution when dropped from the final model, within the final model, and on their own (alone). The final model was generated using the stepwise procedure.

VALIDATION AND CROSS-VALIDATION STATISTICS AND GRAPH SHAPES FOR WESTERN THORNBILL

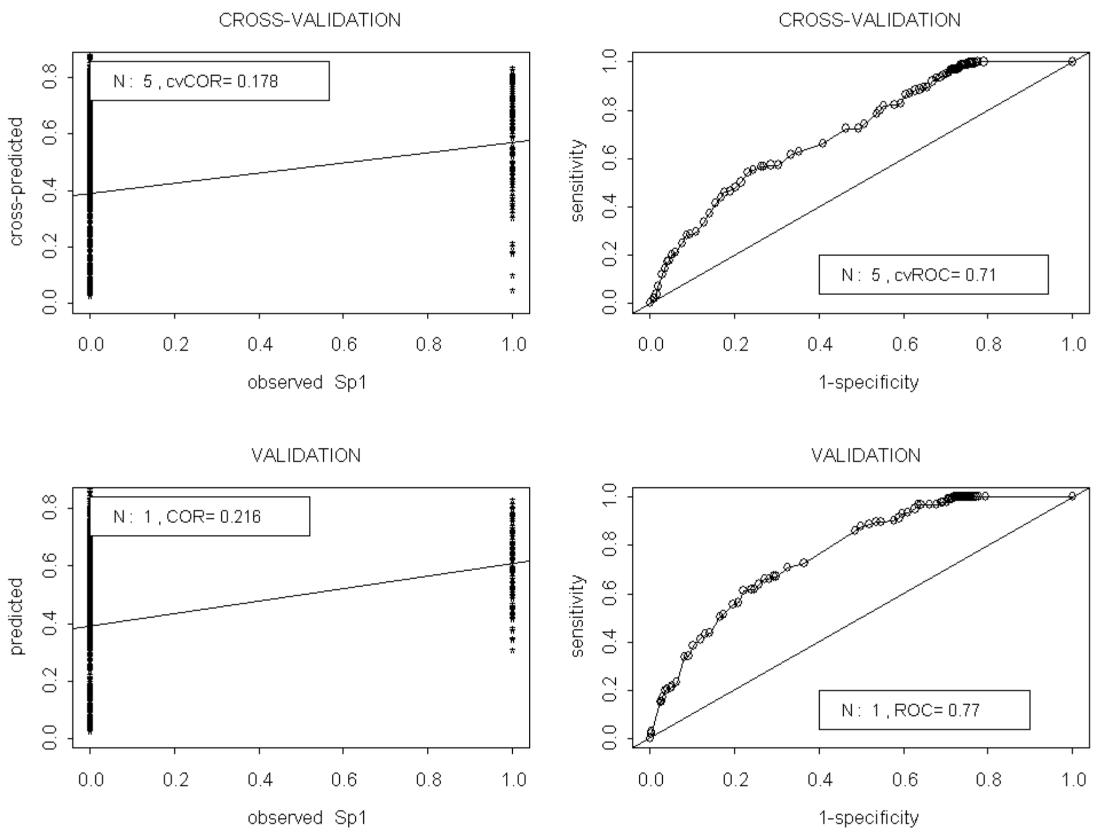

Figure D4: Validation and Cross-validation statistics and graph shapes for the final Western Thornbill model.

PARTIAL RESPONSE CURVES FOR WESTERN THORNBILL PRESENCE RECORDS AGAINST PREDICTORS IN THE FINAL MODEL

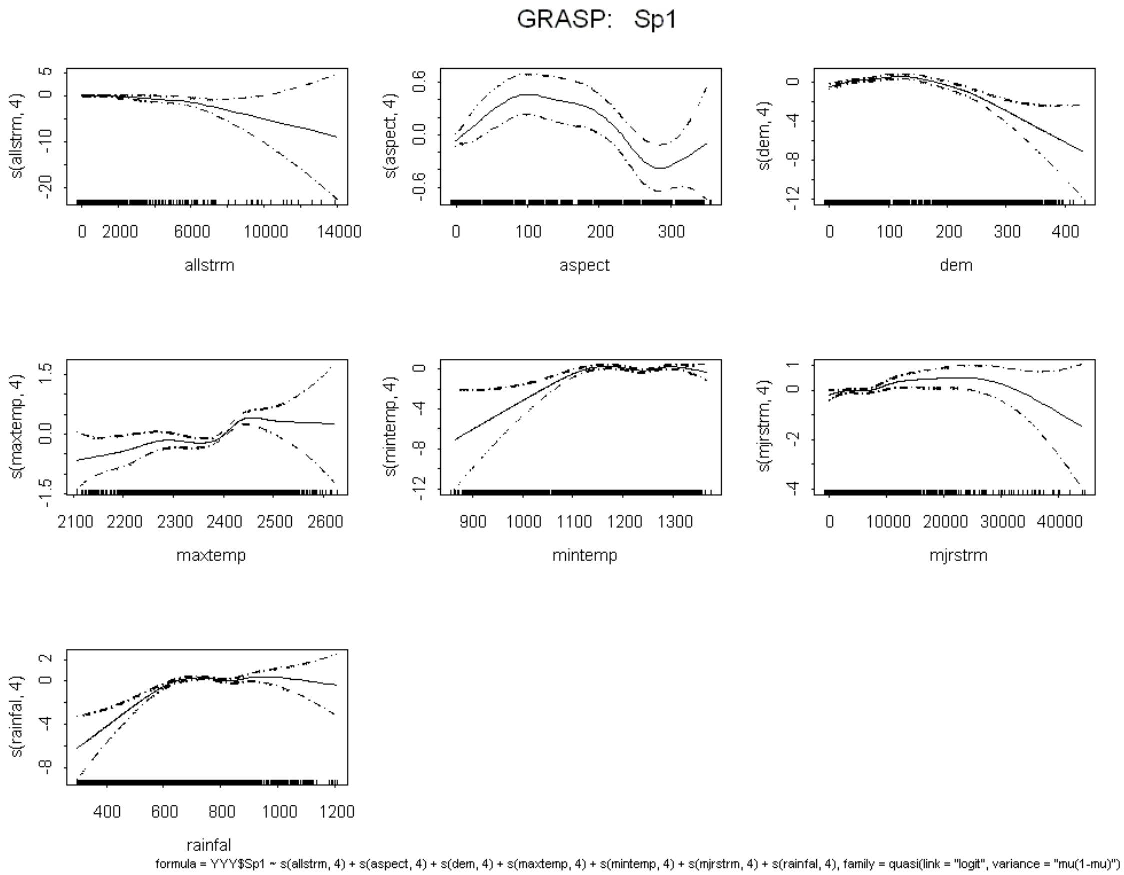

Figure D5: Partial response curves for Western Thornbill presence records against predictors in the final model.

HEAD OFFICE
Suite 4, Level 1
2-4 Merton Street
Sutherland NSW
T 02 8536 8600
F 02 9542 5622

SYDNEY
Suite 604, Level 6
267 Castlereagh Street
Sydney NSW 2000
T 02 9993 0566
F 02 9993 0573

ST GEORGES BASIN
8/128 Island Point Road
St Georges Basin NSW 2540
T 02 4443 5555
F 02 4443 6655

CANBERRA
Level 4
11 London Circuit
Canberra ACT 2601
T 02 6103 0145
F 02 6103 0148

HUNTER
Suite 17, Level 4
19 Bolton Street
Newcastle NSW 2300
T 02 4910 0125
F 02 4910 0126

NAROOMA
5/20 Canty Street
Narooma NSW 2546
T 02 4476 1151
F 02 4476 1161

COFFS HARBOUR
35 Orlando Street
Coffs Harbour Jetty NSW 2450
T 02 6651 5484
F 02 6651 6890

ARMIDALE
92 Taylor Street
Armidale NSW 2350
T 02 8081 2681
F 02 6772 1279

BRISBANE
93 Boundary St
West End QLD 4101
T 0429 494 886

WESTERN AUSTRALIA
108 Stirling Street
Perth WA 6000
T 08 9227 1070
F 08 9227 1078