

School of Molecular and Life Sciences

**Modelling refugia for improved conservation outcomes in the northern
jarrah forest, southwestern Australia**

Sarah Jayne Luxton

**This thesis is submitted for the Degree of
Doctor of Philosophy
of
Curtin University**

March 2020

Author's Declaration

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Sarah Luxton

04 March 2020

Conference Titles

Luxton S., Wardell-Johnson G., Robinson T.P., Sparrow A., Trotter L., Grigg A. (2019, November). Classifying the continuous – testing the utility of community types for decision-making in the biodiverse forests of southwestern Australia. In *Vegetation Science for Decision Making*, Ecological Society of Australia, Launceston, Tasmania.

Luxton S., Wardell-Johnson G., Robinson T.P., Sparrow A., Trotter L., Grigg A. (2018, August). Using topographic data and streamflow elasticity to quantify hydrologic refugia at fine scales in the northern jarrah forest, south-western Australia. *Extreme events, ecosystem resilience, and human wellbeing*. Ecological Society of America, New Orleans, USA.

Luxton S., Wardell-Johnson G., Robinson T.P., Sparrow A., Trotter L., Grigg A. (2018, July). Quantifying hydrologic refugia at fine scales in the northern jarrah forest southwestern Australia. Vegetation modelling. *Natural Ecosystems as Benchmarks for Vegetation Science*. International Association for Vegetation Science, Bozeman, USA.

Luxton S., Wardell-Johnson G., Robinson T.P., Sparrow A., Trotter L., Grigg A. (2018, July). Using topographic data & streamflow elasticity to quantify floristic refugia at fine scales under climate change. H. Possingham (Chair). *Society for Conservation Biology 5th Oceania Congress*, Wellington, New Zealand.

Luxton S. (2018, June). Vegetation Classification – Lessons Learnt from Analysing a 30,000-Plot Database. *Flora and Vegetation Surveys in Environmental Impact Assessment Symposium*. Environmental Institute of Australia and New Zealand, Perth, Western Australia.

Luxton S., Wardell-Johnson G., Robinson T.P., Sparrow A., Trotter L., Grigg A. (2017, December). Automating vegetation types: a step towards identifying refugia in the Mediterranean Forests of south-western Australia. Community Ecology, J. Morgan (Chair). *EcoTas*. Ecological Society of Australia and New Zealand, Hunter Valley, NSW, Australia.

Luxton S., Wardell-Johnson G., Robinson T.P., Grigg A., Trotter L., Sparrow A. (2017, February). Using big-data to improve conservation outcomes in the northern jarrah forest. Current species conservation challenges. Human driven scenarios for evolutionary and ecological changes. *XIV MEDECOS & XIII AET Meeting*, Seville, Spain.

Luxton S., Wardell-Johnson G., Robinson T.P., Sparrow A., Trotter L., Grigg A. (2016, December). Repurposing survey data for systematic conservation planning -northern jarrah forest, SW Australia. *Ecological Society of Australia Annual Conference*, Fremantle, Western Australia.

Acknowledgements

I'd like to acknowledge the Traditional Owners of the land that I live and work on – the Wadjuk, Balardung and Wiilman people, and their elder's past, present and emerging.

Grant, what can I say, if I hadn't wandered into your office that day and heard you say "*I've got this dataset, if you want to work on that*" I would probably be a bit less scarred (aRgh!) and far less rich for it. Thanks, well, for everything... (aka reading and providing feedback on all of my work, awesome field trips, the lunch time chats, sharing your knowledge, botanical ID skills, fielding the oh-so questions, your cheeriness and pizazz).

Ashley, you started me off on this eco-philosophy path back in early 2000. It has been a winding road since then, via Reno, NV. I appreciate your patience, detailed feedback, willingness to talk through the finer points of things and generosity with your time. Not to mention introducing some of the critical elements that have stuck this thesis together, including the streamflow elasticity equation, data-interpretation, code pointers and all things vegetation theory. It wouldn't be the same without you.

Todd, thanks for being so calm and always responding when I sent my data and GIS queries your way, and for your support over the past five years. I know that the work is of a higher caliber because you had a hand in it. Lewis! Thanks for all your data-query help during those first few years and the encouragement. I felt lucky to have someone who was on the same track (and an offsider for those supervisor meetings). Cam, we didn't work together a lot, but I just wanted to say that it's been a pleasure and thanks so much for the help that you have given me. I hope that DBH model finishes up A++.

I would like to thank Andrew Grigg and Alcoa of Australia for giving me the opportunity to work with their database, which forms the foundation of this thesis. Also, for a very much appreciated industry scholarship during my first three years. I would also like to thank the 80-plus botanists that collected the data over the years. I would like to acknowledge the financial support provided by the Australian Government under the Research Training Scheme, which supported the majority of this work. A big thank you to Bush Heritage Australia for their support while I whiled away part-time and for the opportunities that have come with working for such an inspiring organisation.

Kaylene, Megan and Amethyst for assistance in the field. The help team at the Curtin Institute for Computation and Pawsey, for much needed and gratefully received code and computing support. The Department of Biodiversity, Conservation and Attractions for permission to print maps of the southwestern WA forests conservation reserve system and forest ecosystem units.

To Leanda, Mieke and Sheree. Bruh, where would I be without you?! (Don't answer that). Mieke and Sheree, I've loved our camps, bushwalks and chats. Natalia, thanks for putting up with me. I think I would've gone nuts if I didn't have you, our chats and diving.

Finally, to Jan and Graeme – I couldn't have done it without your kindness, generosity and never-ending encouragement and support. Thank you so much. Words can't describe how much it's meant to me to have been with you during this time.

Dedication

I dedicate this thesis to the northern jarrah forest – Wadjuk, Balardung and Wiilman boodjar and all of the plants and animals that call it home. I've now spent countless hours walking your wooded halls and I hope this work helps us, in some small way, to manage you well now and for the generations to come...

Sarah Luxton, March 2020



The northern jarrah forest is in close proximity to Western Australia's capital city, Perth. Land use includes aboriginal values, water catchment, recreation, mining, forestry, firewood collection and conservation. Photo: mixed jarrah (*Eucalyptus marginata*) and wandoo (*E. wandoo*) woodland, Darling scarp, 27 km east of Perth, credit: S. Luxton.

Abstract

In the Mediterranean forests of southwestern Australia, a regional decrease in rainfall has caused significant streamflow decline. How these changes are affecting riparian plant communities is unknown. To provide insight into potential change, and as a basis for climate-ready conservation decision making, this thesis sought to model and map areas of potential hydrologic refugia under future climate scenarios.

In **Chapter 1**, I introduce why climate-informed conservation is urgent in Mediterranean ecosystems; the limitations of the current 'systematic conservation planning' based reserve system; and how the concept of refugia can contribute. I then outline the research aims of the thesis.

In **Chapter 2**, I introduce the dataset that forms the basis of this thesis. Mining company Alcoa of Australia has released a large amount of floristic plot and mapping data for an area of their tenement. The dataset is the largest, most detailed survey of the northern jarrah forest region of SWA, with 30,000 plots collected over 436 km² on a 120 m² grid. The aim of Chapter 2 was to ensure a good understanding of the data's characteristics and that subsequent refugia modelling applications were appropriate. Limitations were defined using regional online herbaria data and vegetation maps. Specifically, floristic traits (taxonomic, habit, form, mode of perennation and nutrition, and naturalised status), vegetation complexes mapped at 1:250,000 scale and conservation status were used. Results show the database to be restricted floristically and spatially. Only shrub and tree taxa provide presence-absence data, while annual species were presence-only. Geographically, only five of 40 vegetation types were well represented. Results suggest that modelling should focus on shrub and tree species, with predictions restricted to the study area.

Modelling vegetation dynamics at the community-level has benefits over species modelling, by amalgamating individual responses into general patterns and providing easily recognisable units for management. In **Chapter 3**, I sought to determine the best groups for community-level refugia modelling for the study area. The floristic database parameterised in Chapter 2 was classified using a non-hierarchical clustering algorithm, and the partition compared to previous community-level vegetation mapping units for the forest. Clustering results showed indistinct types, with dominant species overlapping between groups and uncommon species classified to both core and outlying types. Results from the analysis did not correlate well with previously mapped community groups. As previous mapping has been ground-truthed and provides clearer groups than the current analysis, these community groups were used for further modelling.

Building on the results of Chapter 3, the aim of **Chapter 4** was to model potential hydrologic refugia at the community-level, to determine priority areas for conservation under climate change. Terrain variables were generated from NASA's SRTM 30 m digital elevation model and types automated using multinomial logistic regression in R. In lieu of climate data at an appropriate resolution, a non-parametric streamflow elasticity equation was used to correlate rainfall and streamflow, and project results to future climate scenarios. Changes in extent varied from 99% contraction in a riparian type found on sandy loams, to 300% expansion in the driest type, found on shallow soils over granite.

Patterns of expansion and contraction were combined to indicate areas of potential hydrologic refugia and mapped for the study area. Perennial stream-zones and swampy valley floors provide the basis for hydrologic refugia, with vulnerable floristic communities contracting tightly into these areas.

Defining hydrological refugia using multiple approaches will strengthen confidence in model results. In **Chapter 5**, final maps of potential hydrologic refugia, including zones of stable and transient hydrologic refugia, were produced using species distribution modelling. Indicator species for the riparian communities predicted to contract in Chapter 4 were modelled using 2 m LiDAR data and maxent in R. Topographic wetness and convexity, with aspect (eastness and northness) contributed significantly to models, while the accuracy ranged from 0.5 to 0.8. Overall, species results support community modelling. However, species modelling further highlights the importance of micro-topography in defining hydrological refugia in this landscape and provides an opportunity to identify areas of stable vs transient refugia.

Chapter 6 summarises the value and limitations of a hydrologic refugia approach and how results apply to conservation decision-making and management. This includes the use of refugia maps as a basis for further experimental and monitoring work and the importance of fire and weed management. Finally, a case is made for the inclusion of refugia in the current reserve system, as a step towards climate-ready conservation in the forest.

Table of Contents

| | |
|---|-----|
| Author's Declaration | ii |
| Conference Titles..... | iii |
| Acknowledgements | iv |
| Dedication | 1 |
| Abstract | 2 |
| Table of Contents..... | 4 |
| List of Figures..... | 11 |
| List of Tables | 17 |
| List of Abbreviations | 19 |
| Chapter 1 General Introduction..... | 20 |
| 1.1 A drying Mediterranean forest | 20 |
| 1.2 Conservation planning and climate change | 22 |
| 1.3 Refugia | 23 |
| 1.4 Thesis aims | 24 |
| Chapter 2 A practical approach to ensuring a large vegetation database is 'fit-for-purpose'..... | 25 |
| 2.1 Abstract | 25 |
| 2.2 Introduction | 25 |
| 2.3 Methods | 27 |
| 2.3.1 Data collection..... | 27 |
| 2.3.2 Data cleaning..... | 28 |
| 2.3.3 Floristic context | 28 |
| 2.3.4 Geographic context..... | 29 |
| 2.3.5 Conservation context..... | 29 |
| 2.4 Results | 29 |
| 2.4.1 Floristic context..... | 29 |
| 2.4.2 Geographic context..... | 32 |

| | | |
|-----------|---|----|
| 2.4.3 | Conservation context | 33 |
| 2.5 | Discussion..... | 34 |
| 2.5.1 | Potential applications..... | 35 |
| 2.6 | Conclusion..... | 36 |
| Chapter 3 | Vegetation classification in subtle landscapes – implications for predicting climate change impacts. | 37 |
| 3.1 | Abstract | 37 |
| 3.2 | Introduction | 37 |
| 3.3 | Methods | 38 |
| 3.3.1 | Study site | 38 |
| 3.3.2 | Floristic data | 39 |
| 3.3.3 | Environmental data..... | 40 |
| 3.3.4 | Data cleaning and analysis | 41 |
| 3.3.5 | Cluster analysis and evaluation | 42 |
| 3.4 | Results | 42 |
| 3.4.1 | Simulation of final groups | 43 |
| 3.4.2 | Testing old units against the new: SVT vs ALOC | 44 |
| 3.4.3 | ALOC: two vs 30-group classification..... | 45 |
| 3.4.4 | Environmental variables..... | 47 |
| 3.5 | Discussion..... | 51 |
| 3.5.1 | Vegetation pattern and mapping in the NJF | 52 |
| 3.5.2 | Explaining heterogeneity..... | 52 |
| 3.5.3 | Study limitations..... | 54 |
| 3.6 | Conclusion..... | 54 |
| Chapter 4 | Quantifying hydrologic refugia in a Mediterranean-climate ecosystem | 56 |
| 4.1 | Abstract | 56 |
| 4.2 | Introduction | 56 |
| 4.3 | Methods | 59 |
| 4.3.1 | Study site | 59 |

| | | |
|-----------|--|----|
| 4.3.2 | Floristic data | 59 |
| 4.3.3 | Remotely sensed topographic data. | 62 |
| 4.3.4 | Streamflow elasticity estimator and rainfall | 62 |
| 4.3.5 | Data preparation | 63 |
| 4.3.6 | Modelling & accuracy assessment | 64 |
| 4.3.7 | Definition and prediction of refugia | 65 |
| 4.3.8 | Conservation assessment | 65 |
| 4.4 | Results | 65 |
| 4.4.1 | Topographic wetness index | 67 |
| 4.4.2 | Model performance | 68 |
| 4.4.3 | Quantifying hydrologic refugia. | 69 |
| 4.4.4 | The conservation reserve system and hydrologic refugia..... | 70 |
| 4.5 | Discussion..... | 72 |
| 4.5.1 | Automation of site vegetation types | 72 |
| 4.5.2 | Potential hydrologic refugia in the forest | 74 |
| 4.5.3 | Hydrologic refugia in an SCP framework | 76 |
| 4.6 | Conclusion..... | 77 |
| Chapter 5 | An indicator species approach to defining hydrologic refugia in a drying Mediterranean forest..... | 78 |
| 5.1 | Abstract | 78 |
| 5.2 | Introduction | 78 |
| 5.3 | Methods | 80 |
| 5.3.1 | Study site | 80 |
| 5.3.2 | Study species | 81 |
| 5.3.3 | Environmental variables | 81 |
| 5.3.4 | Streamflow elasticity estimator and rainfall | 84 |
| 5.3.5 | Modelling & accuracy assessment | 84 |
| 5.3.6 | Defining species vs community refugia, variability and thresholds..... | 85 |
| 5.4 | Results | 85 |

| | | |
|------------|---|-----|
| 5.4.1 | Species distribution models | 85 |
| 5.4.2 | Species vs community responses..... | 89 |
| 5.4.3 | Stable vs transient hydrological refugia | 91 |
| 5.5 | Discussion..... | 95 |
| 5.5.1 | Species distribution models | 96 |
| 5.5.2 | Reliability of species predictions..... | 96 |
| 5.5.3 | Species vs community responses..... | 97 |
| 5.5.4 | Stable vs transient hydrological refugia | 98 |
| 5.5.5 | Defining thresholds | 99 |
| 5.6 | Conclusion..... | 99 |
| Chapter 6 | General discussion..... | 100 |
| 6.1 | The value and limitations of a hydrologic refugia approach | 100 |
| 6.1.1 | Statistical limitations | 102 |
| 6.1.2 | Data limitations | 104 |
| 6.1.3 | Biotic interactions..... | 105 |
| 6.2 | Fire and weed management | 105 |
| 6.3 | Improved conservation outcomes in the northern jarrah forest. | 106 |
| 6.4 | Concluding remarks..... | 107 |
| References | | 109 |
| Appendix A | | 131 |
| A1 | The current reserve system for forested region of SWA meets minimum CAR (Comprehensive, Adequate and Representative) requirements. Major creek lines are recognised in the ‘CAR Informal Reserve’ category. Source: Department of Biodiversity, Conservation and Attractions: Map 1, CAR Reserve Map, Regional Forest Agreement for the South-West Forest Region of Western Australia, May 1999. | 131 |
| Appendix B | | 132 |
| B1 | Alcoa raw data, access database. | 132 |
| B2 | Alcoa traits data, access database. | 132 |
| B3 | Northern jarrah forest traits data, access database..... | 132 |

| | | |
|------------|--|-----|
| B4 | Family data, all. | 132 |
| B5 | Vegetation complexes, raw data. | 132 |
| B6 | Conservation indices, raw data. | 132 |
| Appendix C | | 132 |
| C1 | Correlation matrix for potential environmental descriptors for community classification analyses. | 132 |
| C2 | Choice of clustering algorithm – additional information. | 133 |
| C3 | R script, classification. | 134 |
| C4 | Species frequency per group, 30-group classification (Excel file). | 134 |
| C5 | Results from two-group ALOC vegetation classification, overlain over current vegetation mapping systems available for the study area: (a) site vegetation types, (b) vegetation complexes, (c) forest ecosystem units; and (d) the National Vegetation information System (NVIS). | 134 |
| C6 | Two-step, two way table. | 134 |
| Appendix D | | 135 |
| D1 | Site vegetation types metadata. | 135 |
| D2 | Environmental metadata. | 135 |
| D3 | R script, multinomial logistic regression modelling. | 135 |
| D4 | Correlation matrix for potential topographic predictors. | 135 |
| D5 | SAGA wetness index (SWI) surfaces for (a) 2006 and (b) 2090, based on RCP 4.5 rainfall projections and streamflow-rainfall elasticity calculations. | 135 |
| D6 | Binomial logistic regression maps for current (2006) conditions for each site vegetation type. | 136 |
| D7 | Binomial logistic regression maps of change from 2006 to 2090, for each site vegetation type. | 137 |
| D8 | Summary of additional binomial and multinomial logistic regression accuracy assessment results. Binomial models show higher individual accuracy, but were not used due to overlapping distributions and non-independence. | 138 |

| | | |
|------------|---|-----|
| D9 | Forest ecosystem units reflect changes in major vegetation units between upland and stream zone sites. Refugia track the most restricted units, but are also modelled across the general forest type ('Jarrah – North West'), particularly in the north east and south west of the study area. | 139 |
| Appendix E | | 140 |
| E1 | (A) Key habitat characteristics for the indicator species used in modelling and; (B) a summary of key species responses (percent contribution to refugia area). | 140 |
| E2 | Rainfall calculations for historic, current and future climate for southwestern Australia, 1970-2090. CSIRO sustainable yield reporting (CSYP), Climate Change in Australia website (CCA). | 141 |
| E3 | Permutation importance (percent) for all variables used in modelling | 141 |
| E4 | Boxplots of the key predictors used in species distribution modelling show distinct difference in SWI and aspect (northness), but more subtle variation in eastness and convexity. SAGA wetness index (SWI), eastness (EAST), northness (NORTH) and convexity (CONVEX). | 142 |
| E5 | Habitat suitability response curves (P) for the SAGA Wetness Index (SWI), for the species predicted to lose range from 2006 to 2090. | 142 |
| E6 | The indicator species used in modelling each represent a particular site vegetation type, but occur across a larger number of types than expected. | 143 |
| E7 | The identification of refugia at two scales shows both broad and fine scale patterns and highlights the importance of micro-topography in the landscape. (A) community, 30 m scale; (B) species, 2 m scale. | 143 |
| E8 | SAGA Wetness Index (SWI) values for 2006 and 2090. Max SWI is projected to be 0.72 in 2090, indicating the possibility that no areas will provide stable hydrologic refugia under the RCP 4.5. future climate change scenario. | 144 |
| E9 | Comparison of environmental predictors for stable vs other areas of refugia. Convexity = CONVEX, eastness = EAST, northness = NORTH, topographic roughness index = TRI, SAGA wetness index = SWI, distance to valley depth (VALDEP). | 144 |
| E10 | Variation in raw (x-axis) vs percent (y-axis) change in habitat suitability for species predicted to contract in response to drying. | 145 |

| | | |
|------------|---|-----|
| E11 | Table with technical information on how the percent change in future rainfall was calculated. Weighted mean of average annual rainfall projections for the NRM super-region ‘Southern and SW Flatlands’ under RCP 4.5 for 2030, 2050, 2070 and 2090. Values were obtained from CSIRO’s Climate Futures website. Models (%): the number of GCM's (global climate models) that agree on a particular trajectory; Weighted Value: model % multiplied by the rainfall category (e.g. 7% x -15%); <i>Weighted Total</i> : sum of weighted values; and Total Expected (Exp) Change in rainfall from baseline (%): Weighted Total divided by 100. | 145 |
| E12 | Cumulative probability layer, showing the sum of species predicted occurrence in 2090, for taxa that are indicators of refugia (a positive correlation with SWI). See also Figure 5.5, which shows cumulative results based on binarised outputs. | 146 |
| E13 | R scripts, maxent species distribution modelling. | 146 |
| Appendix F | | 147 |
| F1 | Conceptual, methodological and computing challenges – additional information. | 147 |
| F2 | Attribution statement..... | 150 |
| F3 | Copyright permissions..... | 150 |

List of Figures

- Figure 1.1 The Mediterranean region of southwestern Australia (hatched), is renowned for its incredible endemism and is an internationally-recognised biodiversity hotspot. This study is based in the northern jarrah forest of the region. Basemap source: ESRI, Sep 2019.21
- Figure 1.2 a) Percent change in minimum streamflow discharge rates (m^3/s) for a selection of southwestern Australian rivers from 1975 to 2010 shows general decline across the region. While the largest change is seen in the smaller, northern streams, a 25-75% drop has also been recorded across the large south coast rivers. Points are the location of gauges, scaled to log of total discharge per year. b) and c) An example of the inter-annual variability recorded by the gauges for b) low and c) high flow streams.22
- Figure 2.1 Overview of the study area: (A) the northern jarrah forest (NJF, black) is situated in the southwest corner of Western Australia; (B) the extent of mapped vegetation complexes (black) and the Alcoa dataset (grey) in the NJF; (C) close-up of the Alcoa dataset including water bodies (blue), the 120 m survey grid (grey) and areas of no data within the study zone (white).28
- Figure 2.2 Species area curve for cleaned dataset. Almost half of the species recorded occur within a 10 km² area and 80% were captured in <100 km² area. 30
- Figure 2.3 The twenty most abundant families in the northern jarrah forest, with the number of species per family in the NJF flora (grey) vs the Alcoa dataset (black).30
- Figure 2.4 The number of species per habit type in the NJF flora (grey) vs the Alcoa dataset (black).31
- Figure 3.1 Vegetation mapping of site vegetation types in the northern jarrah forest has occurred at two scales: 1) The original survey by Havel (1975) to delineate the SVTs covered a large extent of the forest (hatched) in 320 plots (circles). In contrast, the plot data collected by Alcoa (1991-2015, blue) covers a much more restricted area but was intensively sampled (30,000 plots).40
- Figure 3.2 Histogram of number of species vs. abundance for the Alcoa dataset indicating the dominance of a few key species and occurrence of many uncommon species.43

Figure 3.3 NMDS of 77 and 162 groups (black circles) with species scores (red).
Stress is much lower for the 77-group classification.....44

Figure 3.4 UPGMA dendrogram of ALOCs final 2 and 30-group classifications,
showing the large degree of separation at the two-group scale and more
subtle patterns within the groups. Branch labels = major (A and B) and
minor (1-30) groups. Based on Bray Curtis of mean species abundance
for the 77-group partition.46

Figure 3.5 NMDS with Bray Curtis on mean species abundance per 30-group, with
species (red), groups (black) and significant environmental variables
(blue). Stress = 0.172. SRICH = species richness, CURV = curvature,
FIRE = years since fire, SHAN = Shannon diversity index, SINASP =
eastness, STWI = SAGA topographic wetness index.....48

Figure 3.6 The most statistically robust classification for the study area is a two-group
typology, that has a strong correlation with aspect (northness). Group 1
(black triangles, 1421 sites) is a riparian group sitting low in valleys
along south-facing slopes, while group 2 (clear, 27,000 sites) forms a
large cluster with several hundred species that spans the mid-upper
slopes and plateaus.50

Figure 4.1 Map of study area in the northern jarrah forest (NJF) of south-western
Australia showing Alcoa of Australia’s site vegetation type (SVT)
mapping , which follows a 120 m drilling grid used for ore exploration.
Many of the larger water bodies in the forest are dams, sitting higher in
the landscape than natural lakes. Types are ranked by wetness (C –
purple, wettest; G – red, driest).60

Figure 4.2 Site vegetation types were originally described by Havel (1975), to identify
sites of high productivity for forestry and later adapted to ecological
management. (a) Type ‘G’, exposed and shallow soils on granite
outcrops; (b) the dominant type in the forest ‘S’, includes *Eucalyptus*
marginata with *Banksia grandis*; (c) & (d) the stream-zones dominated
by *E. rudis* most closely match type ‘C’ however, Havel’s original ‘C’,
(e) was usually dominated by *E. megacarpa*; (f) type ‘A’, *Melaleuca*
preissiana and *B. littoralis* in a shallow depression; and (g) type ‘P’,
dominated by *Allocasuarina fraseriana* (Sheok) on lateritic or heavy
gravels with sand. Photos: a-d, f-g, S. Luxton; e, J. Havel 1975a).....61

Figure 4.3 The topographic signal for each site vegetation types varies, with both clear (e.g. channel base, wetness) and subtle (longitudinal curvature) differences between the groups (riparian (A, C, D, E, W), upland (Q, R, S, T), granite outcrops (G) and casuarina (P)).....66

Figure 4.4 Principal components (PCA) biplot of final model environmental variables showing the wide occurrence of several types across gradients (e.g. G, S) and separation at the edges of tolerances (e.g. A, Q). Plot is based on a sub-sample of 825 sites (75 per site vegetation type). CHNBSE = distance to channel base network, LAT = latitude, LON = longitude, WETNESS = SAGA wetness index, CONVEX = plan convexity (aspect), VALDEP = distance to valley depth, LSFAC = LS-factor, LONCRV = longitudinal/profile curvature (slope). Larger circles are means. Types are riparian (A, C, D, E, W), upland (Q, R, S, T), granite outcrops (G) and casuarina (P)67

Figure 4.5 Predicted areas of stable and transient hydrologic refugia for 2090 form a network following the rivers and streams of the forest. The current IUCN Class 2 reserve system includes refugia, e.g. (A) in the Serpentine National Park, but large areas are also not captured; B) hydrologic refugia in areas with little dry(upland)-type expansion; C) refugia around the edges of a major water body; and D) matrix of high dry-type expansion interspersed with large patches of refugia, likely to be broad, swampy valleys.....71

Figure 4.6 Vegetation complexes are a broader unit than the site vegetation types and provide a basis for interpreting hydrologic refugia modelling results in the forest. Refugia are well-represented by the swamp complexes (Swamp and Yarrigal), with very few refugial sites found in the main upland unit (Dwellingup, dark blue).....76

Figure 5.1 Selection of species used in hydrologic refugia modelling: the tree-like monocot a) *Kingia australis*; key swamp-related trees b) *Melaleuca preissiana* and c) *Banksia littoralis*. d) *Hakea lissocarpha*, e) *Trymalium ledifolium* and (f) *Acacia alata* occupy a wide range of conditions. (g) *Pericalymma ellipticum* is found in elevated areas on seasonally swampy platforms and (h) *Babingtonia camphorosmae* occurs mainly on heavy soils. The dominant upland overstorey trees in the northern jarrah forest are i) *Eucalyptus marginata* and j) *Corymbia calophylla*. k) *E. rudis* is associated with the wetter parts of south western Australia, but it can also be found on hillsides, while l) *Taxandria linearifolia* borders swamps and watercourses. Photos, S. Luxton.82

Figure 5.2 (a) Lidar data is available at 2 m resolution for a large section of the study area, enabling elevation and other topographic variables to be derived; (b) exemplar showing *Melaleuca preissiana*, a riparian shrub or tree and it's predicted habitat suitability in 2090. Differences in the species presence (cross) and the site vegetation type it represents (yellow, SVT 'A') are evident; (c) a close-up of habitat suitability in 2090. *M. preissiana* prefers sandy soils and swamps, forming open stands in shallow depressions.83

Figure 5.3 Stacked bar chart of permutation importance (%), grouped by site vegetation type (SVT), with convexity showing higher importance in the drier types and saga wetness in the wettest. Abbreviations: CONVEX, convexity; EAST, eastness; SWI, SAGA wetness index; LONGC, longitudinal curvature; NORTH, northness; TPI, topographic position index; TRI, topographic ruggedness index.87

Figure 5.4 Species response curves developed from standardised environmental variables (x-axis) depicting habitat suitability (y-axis) at relative levels for six of the predictors used in modelling. Species are: (a) *Acacia alata*, (b) *Acacia divergens*, (c) *Acacia extensa*, (d) *Astartea scoparia*, (e) *Babingtonia camphorosmae*, (f) *Banksia littoralis*, (g) *Boronia molloyae* (h) *Corymbia calophylla*, (i) *Eucalyptus marginata*, (j) *E. megacarpa*, (k) *E. patens*, (l) *E. rudis*, (m) *Gahnia decomposita*, (n) *Gahnia trifida*, (o) *Hakea lissocarpha*, (p) *Hakea prostrata*, (q) *Hakea varia*, (r) *Hypocalymma angustifolium*, (s) *Kingia australis*, (t) *Lepidosperma squamatum*, (u) *Lepidosperma tetraquetrum*, (v) *Melaleuca preissiana*, (w) *Pericalymma ellipticum*, (x) *Taxandria linearifolia*, (y) *Trymalium ledifolium*. Abbreviations: CONVEX, convexity; EAST, eastness; SWI, SAGA wetness index; LONGC, longitudinal curvature; NORTH, northness; TRI, topographic ruggedness index.89

Figure 5.5 Species modelling provides further insight into the dynamics of potential stable and transient hydrologic refugia in the study area: a) the areas around dams include refugia but may not be suitable due to human influences like clearing to enhance run-off and manipulation of water-levels; b) the Murray River flows through Lane Pool Conservation Park and represents a major tributary in the study area; c) stacked species and community models of potential hydrologic refugia showing how d) species results both correspond with and show a more nuanced perspective to the (e) community model.92

Figure 5.6 (A) Stable, transient and non-refugia areas show differences in the topographic predictors used in modelling. Convexity decreases as refugial capacity increases, while there is a marked increase in SAGA wetness between non-refugia and refugia areas. (B) Species contributions to stable, transient and non-refugia – all species contribute similarly to stable refugia (0.7-1.4%), but not non-refugial areas (0.3-9%).93

Figure 5.7 (a) The subdued topography of the northern jarrah forest lends itself to complex patterns of potential hydrologic refugia; (b, c) small creeks include many riparian species but are identified to lose taxa with continued drying; (d) some upland creeks already show evidence of encroachment by species associated with dry conditions like (e) *Banksia grandis*; and (f) particularly following fire (*Corymbia calopylla*); (g) The banks of larger streams will likely hold the moisture conditions required for riparian species to 2090 and are the obvious location of potential hydrologic refugia. However, how drying will actually impact more marginal sites (e.g. (h) wet, leached acid sands, underlain by an impermeable horizon and waterlogged in winter) that include mixed suits of upland and riparian taxa requires further work. The areas identified by modelling provide a framework for monitoring, decision-making and further work in the forest. Photos, S. Luxton.....94

Figure 5.8 a) Changes in percent and raw habitat suitability across the range-contracting species (2006 to 2090) may assist with defining the threshold between stable and transient refugia; b) in some taxa (e.g. *Acacia extensa*) raw and percent change are consistent but in others (*Melaleuca preissiana*, *Boronia molloyae*) a percent drop in suitability corresponds to both small and large decreases in raw values. See Appendix E10 for all species.....95

Figure 6.1 Potential hydrologic refugia at the landscape scale show distinct zones of refugia coupled with a complex pattern of variation in water availability and riparian vegetation throughout the forest matrix. This process-based description of important conservation areas sits in contrast to current reserves (dark grey), which are based on the percent-area systematic conservation planning criteria ‘CAR’ (comprehensive, adequate and representative). An increase in reserve area of only 3% is required to formally protect potential zones of refugia.....108

List of Tables

| | |
|--|----|
| Table 2.1 A comparison of plant traits captured by the Alcoa dataset (525 species) relative to the full NJF flora (3089 species), including: life form, mode of perennation, mode of nutrition, naturalised status and conservation status..... | 32 |
| Table 2.2 Vegetation complexes give a broad understanding of vegetation patterns in the dataset relative to the larger landscape and highlight under-sampled areas. Each unit was divided into two categories: (1a) the area and (1b) percent captured by the dataset relative to the northern jarrah forest (NJF); and (2a) the area and (2b) percent within the survey zone that was not sampled. The swamp complex was excluded to a higher degree than other types..... | 33 |
| Table 2.3 Ranked conservation statistics for 10 IBRA subregions in Western Australia. Highly cleared areas in the wheatbelt have both a high number of priority taxa and low protection (AVW01, AVW02). The northern jarrah forest (NJF) ranked 5 th , with an adequate area protected (10.4%) but moderate number of priority taxa (322). IBRA = Interim Biogeographic Regionalisation for Australia. | 34 |
| Table 3.1 Confusion matrix for Havel’s site vegetation types (SVT) vs an 11-group ALOC-derived classification. ALOC does not separate the variation described by SVTs well, only delineating two core groups across the categories..... | 44 |
| Table 3.2 The 25 most abundant taxa are almost identical when overall frequencies for the dataset are compared to group A results (the dominant group in a 2-type classification of the data). Group B is largely distinct, capturing the main riparian species in the forest. Grey = shared species, bold = species unique to group B only..... | 46 |
| Table 3.3 Summary statistics for the 2-group (A, B), and evaluation tests for the 30-group classifications. Aspect, years since fire, curvature and wetness are significantly different between groups in the 30-group classification, as is richness and diversity. PCC: Principle Component Correlation, KW: Kruskal-Wallis, MCAO: Monte-Carlo attributes in an Ordination (% Permuted $R^2 > \text{Actual } R^2$), nMDS – non-metric multidimensional scaling (R^2) with significance (p-val)..... | 49 |

| | |
|--|----|
| Table 4.1 Metadata for the eleven site vegetation types used in this study (riparian (A, C, D, E, W), upland (Q, R, S, T), granite outcrops (G) and casuarina (P)). All variables are taken from Havel (1975a), except for elevation (meters above sea level) and area (km ²) which were calculated for this study (see Appendix D1 for full details). Depth WT = depth to water table. | 62 |
| Table 4.2 Rainfall calculations used in wetness modelling. Historic (pre-1975), “current” (2006) and projected (2030-2090) rainfall for the northern jarrah forest. AR = Annual Rainfall, CCA = Climate Change in Australia, CSYP = Commonwealth Scientific and Industrial Research Organisation Sustainable Yield Project. | 63 |
| Table 4.3 Model coefficients and standard error for each site vegetation type (SVT), ranked by wetness (highest to lowest). Multinomial logistic regression calculates coefficients relative to a reference category (S) for which all variables equal zero. Types are riparian (A, C, D, E, W), upland (Q, R, S, T), granite outcrops (G) and casuarina (P). | 68 |
| Table 4.4 Multinomial logistic regression results: (A) model accuracy assessment and (B) confusion matrix (% correct = percent of types correctly predicted as known). Overall accuracy = 0.44, kappa = 0.13, p-value = <0.001. . | 69 |
| Table 4.5 Confusion matrix for changes in site vegetation type (SVT) occurrence from 2006 – 2090, predicted using the multinomial logistic classifier. Values are area (km ²) and (percent change between types). NET AREA CHANGE = total change in area from 2006 to 2090 (km ²). NET % CHANGE = total percent change in cover for each SVT from 2006 to 2090; grey = no change. Types are ranked in wetness order (C = wettest, G = driest). Type description: riparian (A, C, D, E, W), upland (Q, R, S, T), granite outcrops (G) and casuarina (P). | 72 |
| Table 5.1 Modelling results and summary statistic for each species, including accuracy and percent change from 2006 to 2090. Site vegetation types (SVT) were used to group data. Grey = the species used to define hydrologic refugia, this chapter. | 90 |

List of Abbreviations

| | |
|-------|--|
| Alcoa | Alcoa of Australia Ltd, a bauxite mining company |
| CAR | Comprehensive, Adequate and Representative |
| CSIRO | Commonwealth Scientific and Industrial Research Organisation |
| EVA | European Vegetation Archive |
| HR | Hydrologic Refugia |
| IBRA | Interim Biogeographic Regionalisation for Australia |
| LIDAR | Light Detection and Ranging |
| MCAO | Monte-Carlo Attributes in an Ordination |
| NMDS | Non-metric Multidimensional Scaling |
| NJF | Northern Jarrah Forest |
| OSL | Old, Stable Landscape |
| PATN | A statistical software package for finding patterns in data |
| R | A free software environment for statistical computing and graphics |
| SCP | Systematic Conservation Planning |
| SRTM | NASA Shuttle Radar Topography Mission |
| SWA | Southwestern Australia |
| SWAFR | Southwestern Australia Floristic Region |
| SVT | Site Vegetation Type |
| TWI | Topographic Wetness Index |
| UPGMA | Unweighted Pair-Group Method with Arithmetic mean |

Chapter 1 General Introduction

In the Mediterranean forests of southwestern Australia, a regional decrease in rainfall has caused significant streamflow decline. How these changes are affecting riparian plant communities is unknown. To provide insight into potential change, and as a basis for climate-ready conservation decision making, this thesis sought to model and map areas of potential hydrologic refugia.

1.1 A drying Mediterranean forest

The drying forests of southwestern Australia (SWA) are a high priority for adaptive conservation action. One of the world's vulnerable Mediterranean ecosystems, significant conflict exists between development and conservation goals, despite its recognition as an international biodiversity hotspot (Myers *et al.* 2000; Wardell-Johnson *et al.* 2015) (Figure 1.1). The impacts of climate change are also already evident, including regional declines in rainfall and streamflow, with unknown implications for riparian plant communities (Figure 1.2) (Petroni *et al.* 2010). While the current conservation reserve system meets federal and state government requirements (Janis 1997; Conservation Commission of Western Australia 2013), it does not account for these changes in a formal way. Developing a scientific understanding of how climate change will impact the forest will enable the identification of areas of high conservation concern and inform adaptive management (Groves *et al.* 2012; Keppel *et al.* 2015; Reside *et al.* 2018). It will also provide the basis for a reserve system that protects both current and future forest values (Wardell-Johnson *et al.* 2015; Wardell-Johnson *et al.* 2019).

Worldwide, Mediterranean ecosystems are an outstanding conservation priority (Sala *et al.* 2000; Moreira *et al.* 2019). Characterised by cool, wet winters and hot, dry summers, they house a fifth of known vascular plant species in only 2% of the world's land area (Klausmeyer and Shaw 2009; Cox and Underwood 2011). They are also one of the biomes predicted to experience the greatest biodiversity change over the next 80 years, through both climate and human land-use pressures (Sala *et al.* 2000). In Australia, SWA meets the international 'biodiversity hotspot' criteria of having both extraordinary biodiversity values and high levels of threatening processes (Myers *et al.* 2000; Bradshaw 2012). Home to over eight thousand native vascular plants (of which nearly half are endemic) the region has been extensively cleared, and as of 2000, almost 90% of primary vegetation has been removed (Bradshaw 2012; Gioia and Hopper 2017) (Figure 1.1). Scientifically informed adaptive management is essential to help mitigate further losses and ensure reserve systems are sufficient and effective (Reside *et al.* 2018).

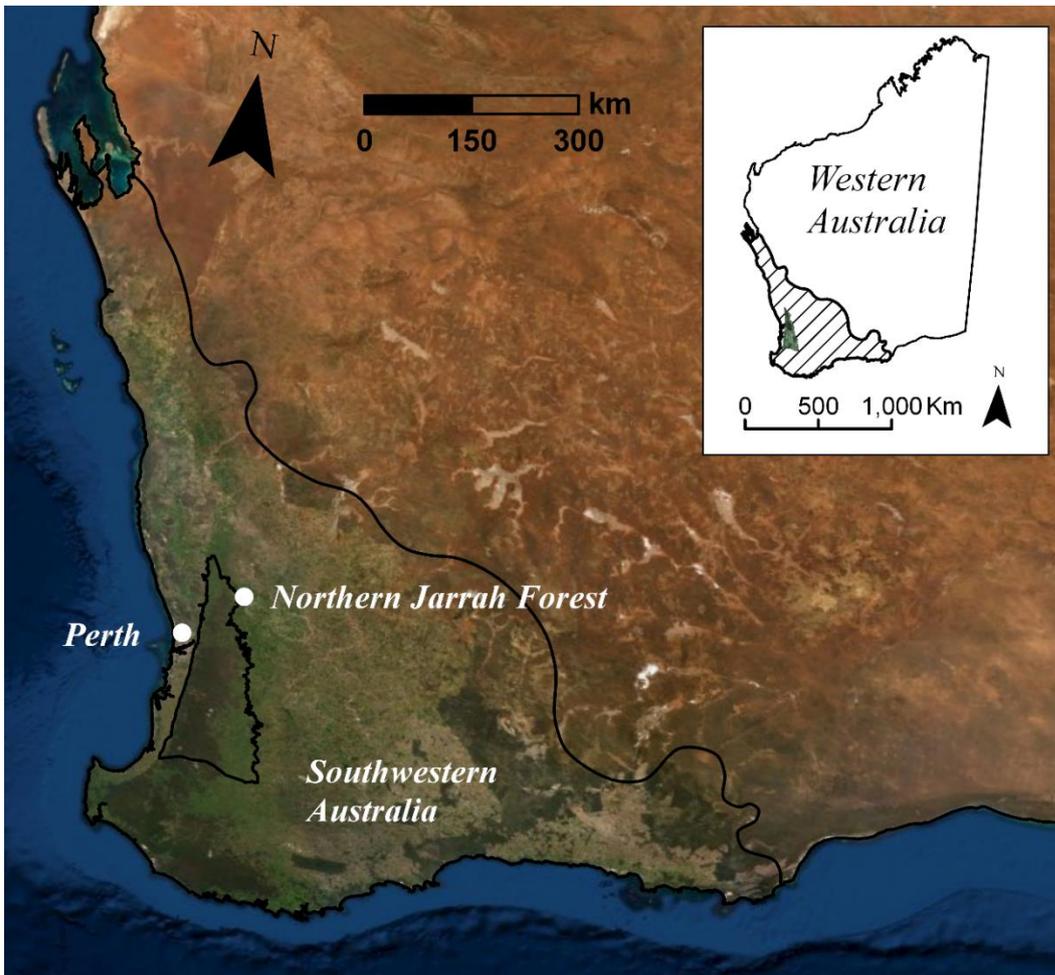


Figure 1.1 The Mediterranean region of southwestern Australia (hatched), is renowned for its incredible endemism and is an internationally-recognised biodiversity hotspot. This study is based in the northern jarrah forest of the region. Basemap source: ESRI, Sep 2019.

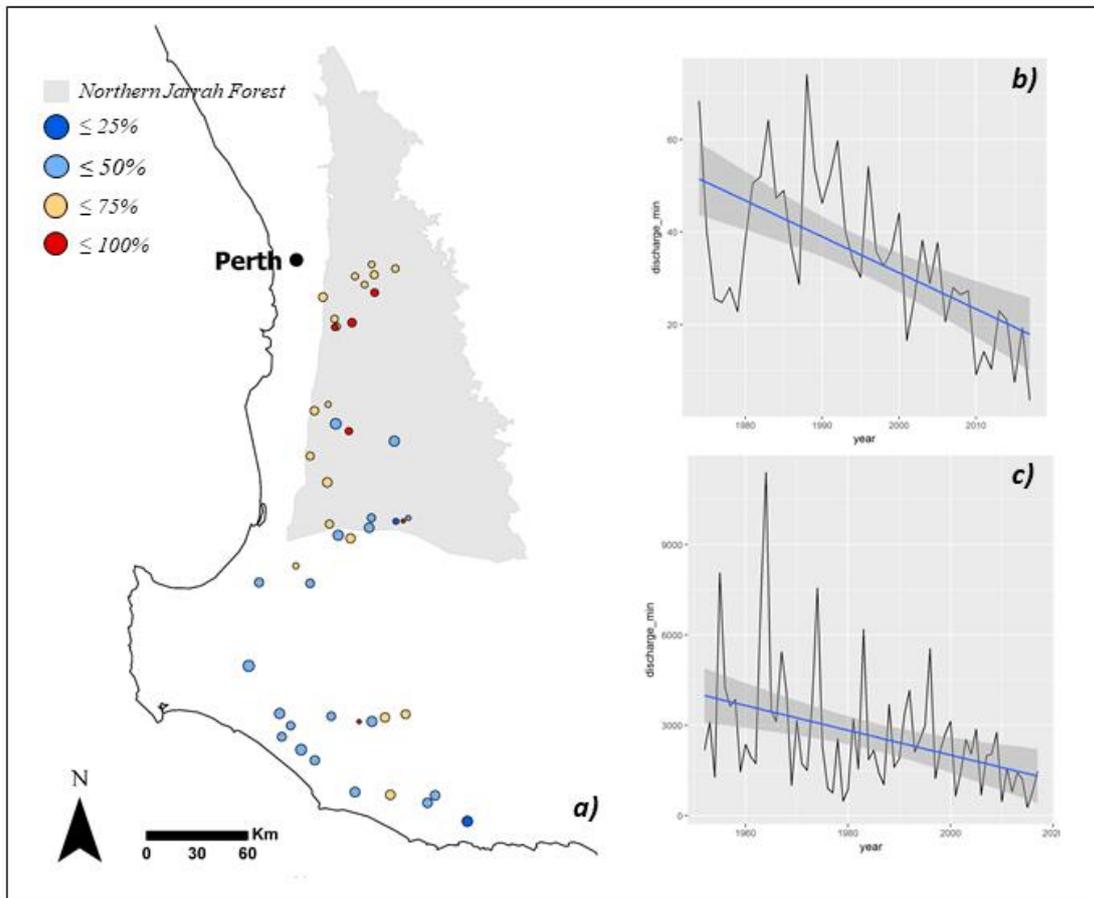


Figure 1.2 a) Percent change in minimum streamflow discharge rates (m^3/s) for a selection of southwestern Australian rivers from 1975 to 2010 shows general decline across the region. While the largest change is seen in the smaller, northern streams, a 25-75% drop has also been recorded across the large south coast rivers. Points are the location of gauges, scaled to log of total discharge per year. **b)** and **c)** An example of the inter-annual variability recorded by the gauges for **b)** low and **c)** high flow streams.

1.2 Conservation planning and climate change

The principles of systematic conservation planning underpin the conservation reserve system in SWA (Janis 1997; Conservation Commission of Western Australia 2013; Australian Government 2015). While now one of the most well-recognised conservation planning methods internationally (Margules and Pressey 2000; Moilanen *et al.* 2009; McIntosh *et al.* 2017), criticisms include that it over-simplifies real-world problems (Langford *et al.* 2011) and that climate change impacts need to be incorporated into the framework (Groves *et al.* 2012; Groves and Game 2016; Reside *et al.* 2018). Strategies to do this (outlined by Groves *et al.* (2012)) include: better accounting for geodiversity (Gray 2004; Anderson *et al.* 2015; Hjort *et al.* 2015); landscape connectivity (Hodgson *et al.* 2009; Doerr *et al.* 2011; McGuire *et al.* 2016); and ecosystem processes and function (Hobbs *et al.* 2009; Harvey *et al.* 2017); and protecting climatic refugia (Groves *et al.* 2012; Keppel *et al.* 2012; Keppel and Wardell-Johnson 2012; Keppel *et al.* 2015). Refugia provide a particularly relevant approach in SWA, given the drying trends already recorded across the region (Figure 1.2).

1.3 Refugia

The identification of refugia provides a means to assess the adequacy of current reserve systems and prioritise conservation decisions under climate change (Keppel *et al.* 2012; Keppel and Wardell-Johnson 2012; Reside *et al.* 2014; Schut *et al.* 2014; Tapper *et al.* 2014a; Tapper *et al.* 2014b; Keppel and Kavousi 2015; Keppel *et al.* 2015; Keppel and Wardell-Johnson 2015; Mokany *et al.* 2016; Morelli *et al.* 2016; Keppel *et al.* 2017; Yates *et al.* 2019; Barrows *et al.* 2020; Cartwright *et al.* 2020; Morelli *et al.* 2020; Thorne *et al.* 2020). Defined as habitats that species can retreat to, persist in and potentially expand from under changing environmental conditions (Keppel *et al.* 2012), the concept originates in palaeobiological research, but has wider implications (Stewart and Lister 2001; Gómez and Lunt 2007; Weiss and Ferrand 2007; Byrne 2008; Médail and Diadema 2009; Gavin *et al.* 2014).

More recently, it has been adapted to study the impacts of anthropogenic climate change, to predict whether critical resources and biodiversity will co-contract under climate pressures (Keppel *et al.* 2012). Theoretically, refugia are spatiotemporal ‘points’ within the multidimensional space of ecosystems that taxa may retract to and persist in under unfavourable conditions. They operate by buffering regional changes in climate at fine scales, providing increased habitat through environmental heterogeneity (e.g. micro-topography, cold air pooling) (Ashcroft *et al.* 2009; Ashcroft and Gollan 2013; Keppel *et al.* 2017). The concept provides a context for, and methods to, determine the impacts of climate change and identify taxa and areas that will act as safe havens into the future. To be effective however, the scope, scale and resolution of refugia needs to be defined at the outset of a study (Keppel *et al.* 2015).

In Australia, identifying and assessing the quality of refugia at the level of individual species and species assemblages has been recognised as a critical to their usefulness as targets for conservation and management (Reside *et al.* 2013; Reside *et al.* 2014; Schut *et al.* 2014; Keppel *et al.* 2017). Multiple forms of refugia are relevant at this scale (including refugia for temperature and fire) (Reside *et al.* 2014) however, in SWA, identifying moisture refugia is the highest priority. Termed ‘hydrologic refugia’ by McLaughlin *et al.* (2017), their identification provide a means to incorporate the effects of regional drying trends into conservation plans across SWA (Janis 1997; Petrone *et al.* 2010; Hughes *et al.* 2012).

Hydrologic refugia are sites of higher moisture availability in a landscape, providing a mesic microenvironment that persists in the face of general drying trends (McLaughlin *et al.* 2017). Hydrologic refugia may be stable (de-coupled from regional drying), or relative and/or transient (tracking regional changes in climate and disappearing as certain thresholds are crossed). Distinct processes usually support the occurrence of riparian species in these areas, by improving water input and/or limiting moisture losses from the root zone (McLaughlin *et al.* 2017). These include: topographic heterogeneity (drainage, channels, slope variation, fractures/faults, flood plains), shallow/perched aquifers due to geological and edaphic variation (aquicludes, aquitards, clay/caliche lenses) (McLaughlin *et al.* 2017), deep soils or regolith acting as water-storage (Dawson and Pate 1996), and seeps and springs (McLaughlin *et al.* 2017).

In the SWA, mesic habitats that may provide hydrologic refugia include the fringes of granite outcrops, swamps, winter-wet valley floors and the boundaries of streams and rivers. It is hypothesised that granite outcrops were refugia during the Pleistocene (Hopper 1979; Byrne 2008; Tapper *et al.* 2014b). They have

much higher microtopographic complexity than the surrounding landscape and capture and direct water (Laing and Hauck 1997; Schut *et al.* 2014; Tapper *et al.* 2014a). They also support high floristic biodiversity, occupying less than 1% of the region but supporting approximately 17% of its native vascular flora (Hopper *et al.* 1997; Yates *et al.* 2019). Conversely, work to identify and define the refugia capacity of swamp and stream zones has not been undertaken in SWA. These areas house a distinct suite of species including the trees *E. rudis* (Flooded gum), *Melaleuca preissiana* (Moonah) and *Banksia littoralis* (Swamp Banksia), plus various riparian shrubs and herbs. Understanding the risk that drying poses to these zones, and where thresholds of change vs stability may lie, will provide an important bridge between current and future priority conservation areas.

1.4 Thesis aims

The overall aim of this thesis was to determine priority areas for conservation action under climate change in SWA, through modelling and mapping potential hydrologic refugia. To limit over-stating model results and encourage practical applications (Addison *et al.* 2013) this thesis takes a conservative (Chapter 2) and multiscale (Chapter 4 and 5) approach to working with the data.

In summary, **Chapter 2** introduces the Alcoa vegetation dataset that underpins the analytical work of this thesis and frames it within the taxonomic, geographic and conservation context of the wider region. The aim was to define the limits of the Alcoa dataset and ensure that it would be appropriately used to model refugia. Given limits, **Chapter 3** is based on the premise that community types may be a more practical approach to climate change management, averaging individual species responses into recognisable units. Species plot data was classified and compared to site vegetation types (the current management unit for the forest) to determine which typology would be best for modelling community-level change. As clear community groupings were not found using the Alcoa data, site vegetation types were selected. **Chapter 4** uses site vegetation type mapping for the study area and 30 m topographic variables to model community occurrence. Streamflow elasticity, CSIRO's 2090 rainfall calculations and the topographic wetness index were then used to project community types, with the aim of identifying areas of potential hydrologic refugia under climate change. **Chapter 5** builds on the work of Chapter 4 with the aim of improving confidence in overall model results, by modelling the current and 2090 distribution of riparian indicator species for the site vegetation types. An effort to improve the nuance and realism of community results was done by defining and quantifying the difference between areas of stable and transient refugia. Based on LiDAR data at 2 m resolution, the models provide a more in-depth analysis of change, with overlap and differences to community responses identified and a final adjusted map of potential hydrologic refugia made. **Chapter 6** puts findings into context and discusses the limitations and strengths of a hydrologic refugia modelling approach. Recommendations for further work and potential management issues are touched on, and an argument made for the inclusion of hydrological refugia into the current conservation reserve system.

Chapter 2 A practical approach to ensuring a large vegetation database is ‘fit-for-purpose’.

2.1 Abstract

Modelling requires careful consideration of whether data is ‘fit-for-purpose’ and outputs suitable for their intended applications. Large amounts of vegetation data are now available through online platforms like S-plot and the European Vegetation Survey however, screening data can be difficult. Metadata including how and why data was collected may be missing, obscuring the appropriateness of potential applications. A simple approach is taken to ensuring vegetation data is fit for purpose, by taking a vegetation dataset that is to be used for climate change modelling and parametrises it against a government herbarium database for the wider region. Alcoa of Australia Ltd, a bauxite mining company have been collecting data in the northern jarrah forest for 25 years. With over 30,000 plots and 700 species, the dataset is the most detailed survey in the forest to date. However, while comprehensive, the original purpose of data-collection was to meet statutory (not research) requirements and limitations are evident. Plots are not full a full enumeration of all species and presence-absence for some taxa (shrubs and trees) but not others (annuals, biennials, non-perennial graminoids). When compared to data for the region a select proportion of families are well-represented, including a dominance of Fabaceous, Proteaceous and Myrtaceous shrubs. Additionally, data covers a restricted area of the forest, particularly concentrated on two vegetation types correlated with high laterite (and therefore bauxite) presence; Dwellingup Complex I and II. These limitations place the dataset in context and inform appropriate modelling strategies going forward, namely – the occurrence of shrub, tree and perennial herb species and how they may respond to drying. Additionally, it is recommended that projections are restricted to the same vegetation complexes as found in the study area. These precautions will ensure that data use is fit-for-purpose and improve confidence in model results.

2.2 Introduction

Survey data is time-consuming and expensive to collect and once obtained not traditionally reused outside of its original research purpose (Hampton *et al.* 2013). However, the value of data is increasingly being recognised as ecology meets the ‘information age’ and groups around the world are working to provide platforms to archive and share data. Multi-institutional collaborations to further use and develop data range from national vegetation classification systems to cross-continental climate change observatories. These include VegBank and EcoVeg (the United States national vegetation database and classification system), the European Vegetation Archive (EVS), sPlot (a repository for plant community data), the Botanical Information and Ecology Network (BIEN) and the National Vegetation Survey databank, New Zealand (Wiser *et al.* 2001; Chytrý *et al.* 2015; Chytrý *et al.* 2016; Enquist *et al.* 2016; Goldsmith *et al.* 2016; Walker *et al.* 2016) see also Table 1. (Wiser *et al.* 2001). Within Australia – ÆKOS-TERN (Advanced Ecological Knowledge and Observation System, Terrestrial Ecosystem Research Network Australia) and the National Vegetation Information System (NVIS) provide nation-level platforms for vegetation data. While State and Territory databases hold significant floristic and environmental plot data

that can be used for quantitative analyses. These include: BioNET (NSW), COREVeg (QLD), NatureMap (WA), the Vegetation Site Database (NT) and more; which together hold over 180,000 records (Gellie *et al.* 2018).

These databases can be used to support environmental research and management without the costs of additional data-collection (e.g. 17 studies have used the EVS to date (Chytrý *et al.* 2016) and 16 publications BIEN). However, as data is being re-adapted from its original purpose there are limits to useability that need to be identified. These include variation in the scale (grain and extent/coverage) of data (Stein and Kreft 2015), raw data format (e.g. occurrence vs presence-absence vs abundance data (Rondinini *et al.* 2006)) and that there may be a lack of metadata indicating how and why the data was collected (Wiser *et al.* 2001). Appropriate uses and analysis options depend on these factors (Hampton *et al.* 2013). Re-adapted data is often used in the conservation space where tight funding limits large-scale data collection however, a model is dependent on the quality of data inputs (McKenzie *et al.* 1996; Pressey 2004; Ware *et al.* 2018).

The northern jarrah forest (NJF) of southwestern Australia (SWA) is a biodiversity hotspot of on-going conservation concern (Thackway and Cresswell 1995; Myers *et al.* 2000). A Mediterranean forest, it is dominated by the trees *Eucalyptus marginata* (jarrah) and *Corymbia calophylla* (marri) and has a diverse shrub and herbaceous flora (approx. 3000 taxa) with over 300 priority taxa (Western Australian Herbarium 1998-). Climate, topography, water availability and fire act together to structure species distributions, while forest structure and composition have been heavily impacted by western anthropogenic disturbance (fire, logging, mining) (Wardell-Johnson *et al.* 2015). The climatic and hydrological regimes of the forest are also shifting. Rainfall has decreased 15-20% and temperature increased 0.4% since mid-1970 (Bates *et al.* 2008) resulting in decreased surface run-off, a falling water-table and shifts from perennial to ephemeral streams (Petroni *et al.* 2010; Hughes *et al.* 2012).

Protected areas in the NJF occur within a mosaic of public land-use (mining, logging, water run-off and storage, firewood collection, recreation). Ten percent of the forest is gazetted conservation estate, meeting the minimum requirement under the “*Nationally Agreed Criteria for the Establishment of a Comprehensive, Adequate and Representative Reserve System for Forests in Australia*” (Janis 1997). However, 10% is an arbitrary value with a complex history, often based on socio-political rather than ecological values (McKenzie *et al.* 1996; Rundle 1996; Carwardine *et al.* 2009). With a further decline in rainfall expected (6-9% by 2100) (CSIRO and Bureau of Meteorology 2015) improving basic and climate change research in the forest is essential. Data-driven climate-ready conservation plans will ensure that ecological values underlie decision making (Groves *et al.* 2012; Wardell-Johnson *et al.* 2015).

Alcoa of Australia Ltd (Alcoa) have released a large amount of floristic plot and mapping data for an area of their mining tenement in the NJF. This data provides the largest, most detailed survey in the NJF to date with 30,000 plots collected over 436 km², on a 120 m² grid. The Alcoa dataset provides the opportunity to further scientific understanding of this system however, the original purpose of data-collection was to meet statutory, not ecological research requirements.

To provide a way forward and ensure that research questions based on the dataset are appropriate, a quantitative context for the data is developed. Data from the broader NJF and WA is used to describe: (a)

floristic indices (taxonomic, habit, form, mode of perennation and nutrition, and naturalised status); (b) geographic areas (vegetation complexes); and (c) conservation indices. Limitations of the dataset are discussed in light of the regional data-pool, potential applications and their statistical requirements, and the merit of this approach for working with repurposed datasets generally.

2.3 Methods

The NJF has a Mediterranean climate with hot, dry summers, cool, wet winters and an average annual rainfall of 600-1200 mm per annum (Gentilli 1989). The area is of low topographic relief due to intense weathering with little rejuvenation over millions of years (Mucina and Wardell-Johnson 2011). Granite, laterite and dolerite form the base rock beneath soils, which are critically nutrient poor. The flora is diverse (3089 taxa) and sclerophyllous, with specialist adaptations in response to nutrient poverty and fire (Hopper and Gioia 2004; Orians and Milewski 2007; Hopper 2009). The vegetation of the forest has been mapped at several scales including site vegetation types (Havel 1975a; 1975b), vegetation complexes (Heddle *et al.* 1980; Mattiske and Havel 1998; Havel and Mattiske 2000), forest ecosystem units (Bradshaw *et al.* 1997; Regional Forest Agreement Steering Committee 1998) and the National Vegetation Information System (NVIS) (NVIS Technical Working Group 2017).

2.3.1 Data collection

Alcoa have mined bauxite ore in the NJF since 1963 (alcoa.com). To meet statutory requirements (identify rare and threatened flora, map vegetation and provide plant species lists for restoration) floristic surveys have been conducted since 1991 (Figure 2.1). The dataset now spans 432 km², 25 years (1991 – present) and consists of 31,084 plots and 550,597 species records. Floristic plots overlap the grid system used by geologists for ore and chemical sampling (120 m x 120 m) (Figure 2.1). At each plot, understory taxa (five m radius) and tree species (20 m radius) were recorded and cover-abundance (1-5) and stress (1-5) scored (Mattiske 2012). Collected plant specimens were dried, fumigated and identified in accordance with requirements of the West Australian Herbarium. The conservation status of taxa was checked using Government Gazette lists (Western Australian Herbarium 1998-; Mattiske 2012).

Additional data for each plot includes environmental variables (soil, topography, outcrop type and amount, dieback (*Phytophthora cinnamomi*) presence/absence and impact, years since fire, log debris, number of stumps) and geological data (soil and overburden depth (m), depth to pallid zone clays (m) and soil chemistry readings).

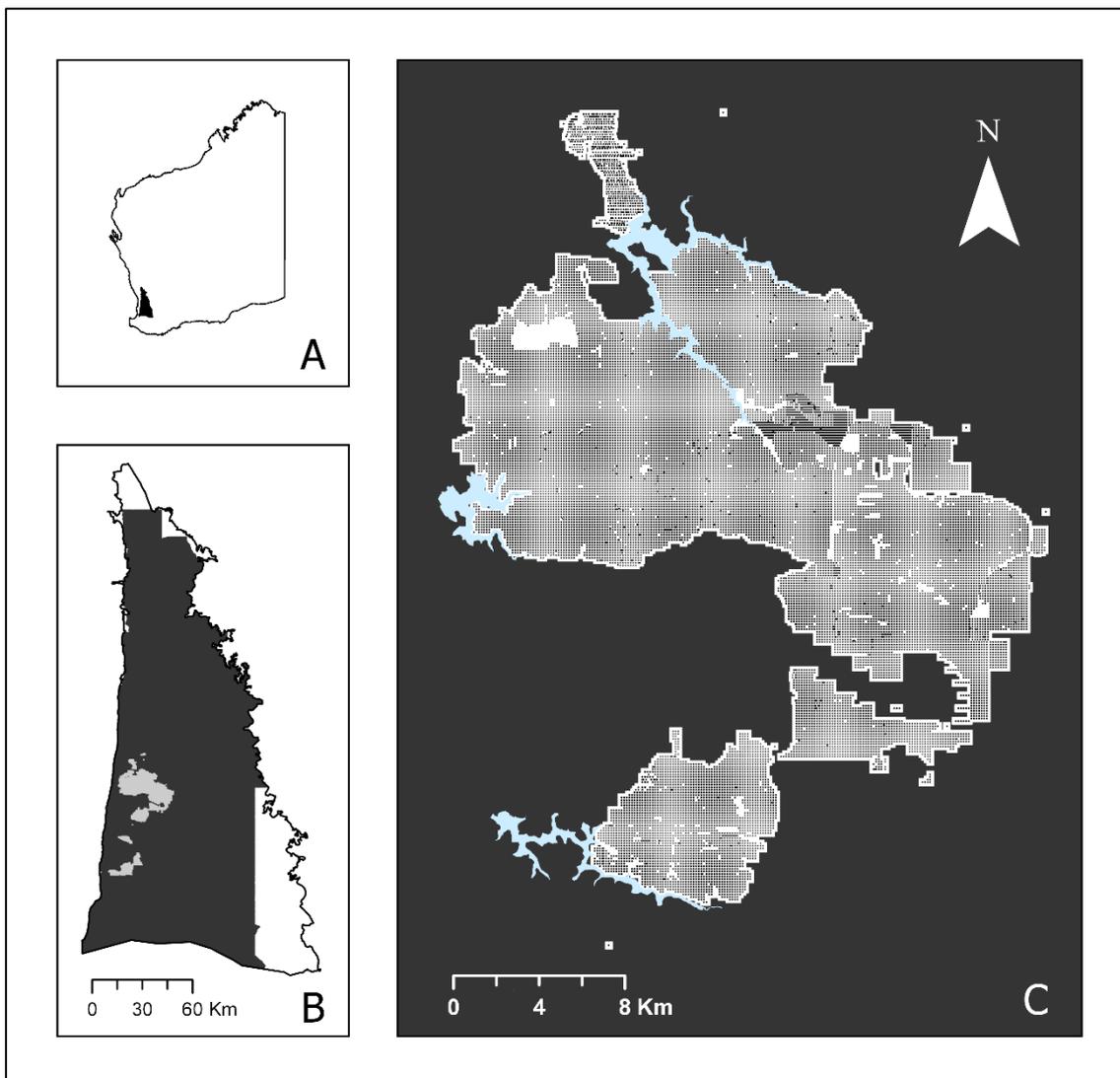


Figure 2.1 Overview of the study area: (A) the northern jarrah forest (NJV, black) is situated in the southwest corner of Western Australia; (B) the extent of mapped vegetation complexes (black) and the Alcoa dataset (grey) in the NJV; (C) close-up of the Alcoa dataset including water bodies (blue), the 120 m survey grid (grey) and areas of no data within the study zone (white).

2.3.2 Data cleaning

A floristic management database was developed to prevent data corruption and ensure future data additions fit system requirements (Trotter *et al.* 2018). In addition to the technical data-cleaning process, 337 species were excluded from the original Alcoa species list due to data-entry errors, sub-species updates and incorrect species identification. Singletons and doubletons were also removed from the dataset for this analysis (186 species) but remain in the original database. In total 525 species were available for analysis (Appendix B1).

2.3.3 Floristic context

To determine what element of the NJV flora is captured by the dataset, a traits database was created using the WA online herbarium 'FloraBase' (Western Australian Herbarium 1998-). All species in the IBRA

Sub-region “Northern Jarrah Forest” (IBRA_SUB: JF1) were uploaded (3089 taxa, retrieved 27 July 2016). For each species: family, habit, life form, nutritional mode, mode of perennation and naturalised status were recorded (Appendix B2, Appendix B3). The comparison of traits for the dataset relative to the full flora (525 vs 3089 vascular plant taxa) was done in Excel (Microsoft Office, Excel 2013). Chi-square analysis was used to test for significant differences (R, version 3.3.2) (R Core Team 2017).

2.3.4 *Geographic context*

Vegetation complexes were used to provide a geographic context for the dataset, comparing both the survey area with the entire NJF, and unsampled areas within the survey bounds to the survey area. Forty-one vegetation complexes have been mapped for the NJF based on aerial photography and survey work, in relation to soil-landform units (Churchward and McArthur 1980; Heddle *et al.* 1980). The area (square kilometres) of each vegetation complex was recorded (ArcGIS 10.3.1, Excel 2013) and chi-square analysis used to test for significant differences (R).

2.3.5 *Conservation context*

Each species conservation status was recorded from FloraBase (Priority 1-4 or not threatened) (Appendix B2, Appendix B3) (Western Australian Herbarium 1998-). The number of priority taxa and percent area protected was also recorded for each IBRA subregion (47 in total) from the 2015 Collaborative Australian Protected Area Database (Australian Government 2015). The number of priority taxa per percent area reserved was calculated and plotted in Excel (2013).

2.4 **Results**

2.4.1 *Floristic context*

The Alcoa dataset contains a skewed proportion of the full NJF flora. Seventeen percent of vascular plant species were captured (525/3089) in half of the families (74/138), with an average of 17 species per plot (SE = 0.03). Turn-over is high, within almost half of the species captured found within a 10 km² area and 80% within a quarter of the study area (Figure 2.2).

In the 20 most well-represented families, the number of taxa captured was significantly different to that expected for the full flora ($X^2 = 82.7$, $df = 19$, $p = <0.05$). Fabaceae and Proteaceae were the most abundant (374 and 291 taxa), representing a quarter of the surveyed flora, while Orchidaceae (12%, 23 taxa), Poaceae (15%, 23 taxa) and Goodeniaceae (15%, 11 taxa) were the least-well surveyed (Figure 2.3). Dilleniaceae, Thymelaeaceae, Restionaceae and Apiaceae were especially well-represented with approximately half of available taxa captured for each family (Figure 2.3). Twenty-three families were represented by a single species including: Araliaceae, Loranthaceae and Scrophulariaceae (Appendix B4).

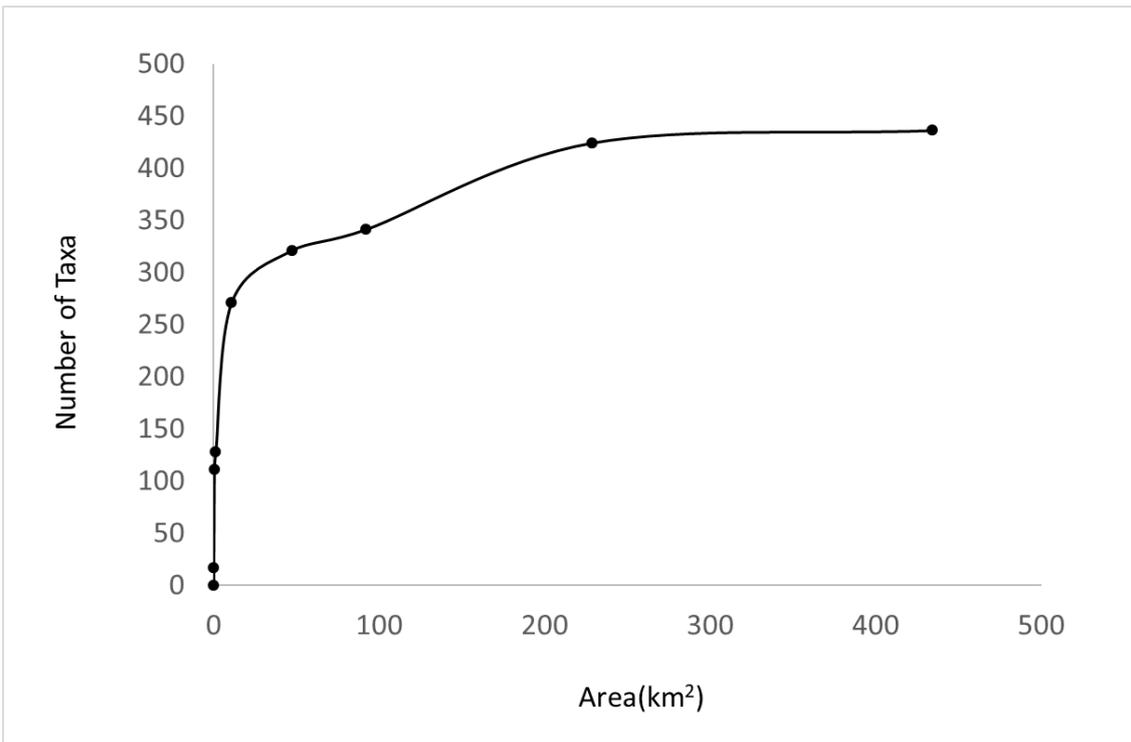


Figure 2.2 Species area curve for cleaned dataset. Almost half of the species recorded occur within a 10 km² area and 80% were captured in <100 km² area.

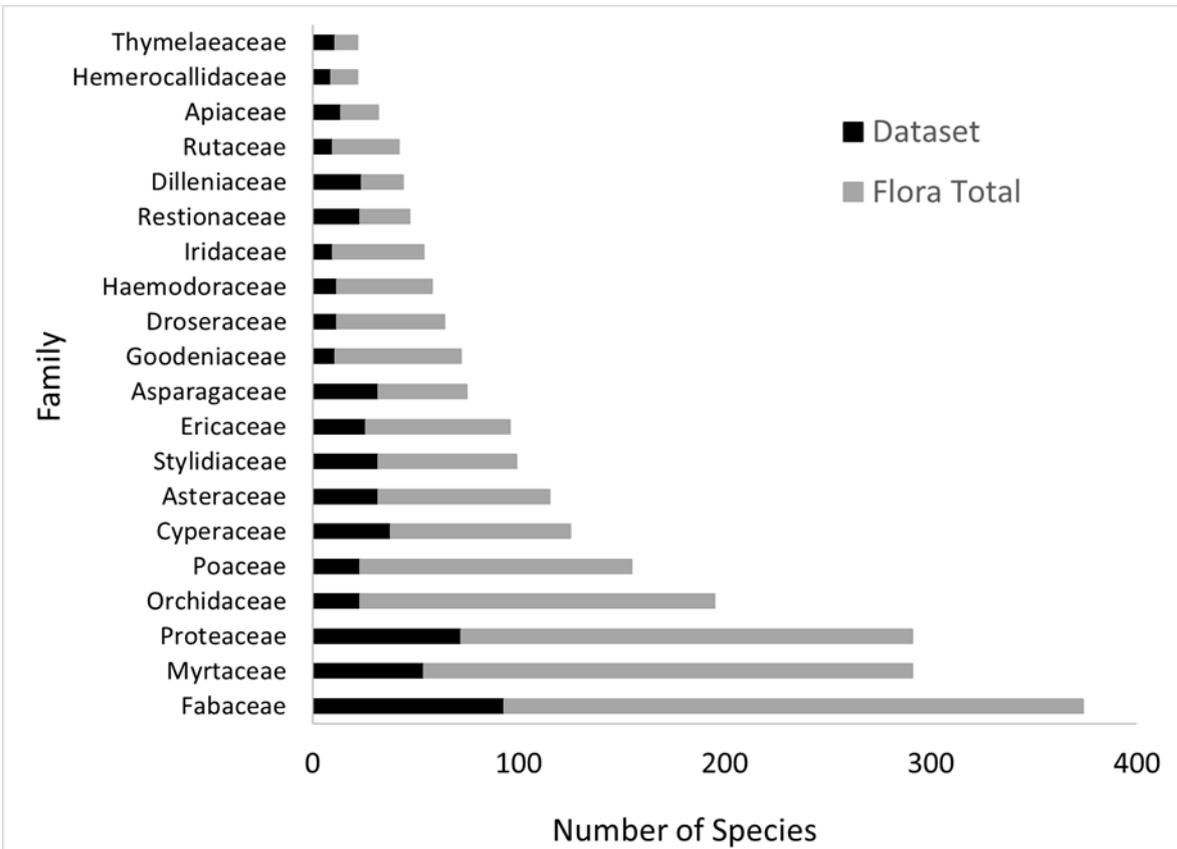


Figure 2.3 The twenty most abundant families in the northern jarrah forest, with the number of species per family in the NJF flora (grey) vs the Alcoa dataset (black).

The Alcoa dataset includes 21% of the NJF's shrub flora, 14% of herbs, 18% of graminoids and 40% of tree species (Figure 2.4). Annual taxa were proportionally under-represented in the dataset, with 10% of annuals captured, compared to 20% of perennials (Table 2.1). A quarter of rhizomatous species were captured, but only 5% of the cormous (2 from 38). Four of the 6 parasitic species for the forest were recorded, with the dominant number 'free-living' (Table 2.1). Five percent of the dataset is alien species, compared to 15% in the full flora, and a quarter 'mixed' (native in parts of range, naturalised elsewhere).

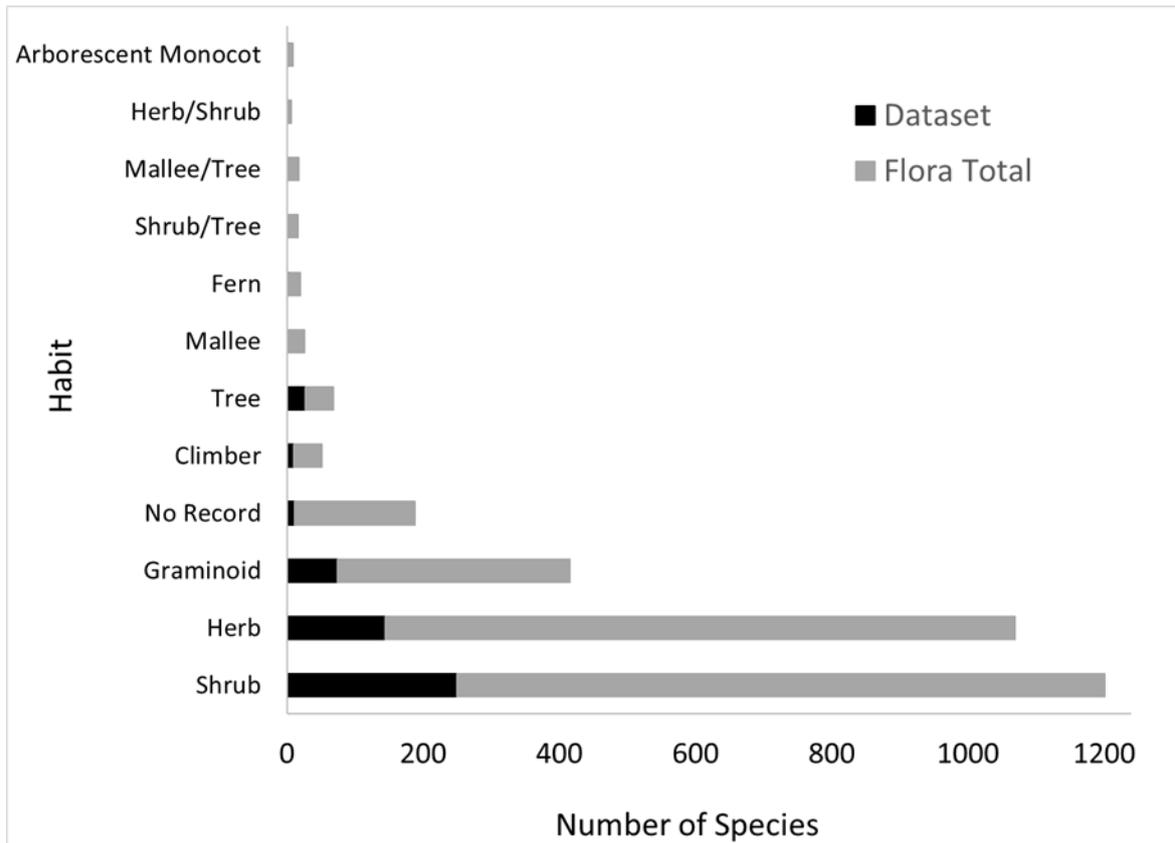


Figure 2.4 The number of species per habit type in the NJF flora (grey) vs the Alcoa dataset (black).

Table 2.1 A comparison of plant traits captured by the Alcoa dataset (525 species) relative to the full NJF flora (3089 species), including: life form, mode of perennation, mode of nutrition, naturalised status and conservation status.

| Trait | Dataset | NJF | Percent (%) |
|-----------------------------|----------------|-------------|--------------------|
| Life Form | | | |
| Annual | 36 | 390 | 9 |
| Annual/Biennial | 0 | 15 | 0 |
| Biennial | 0 | 3 | 0 |
| Annual/Perennial | 0 | 46 | 0 |
| Biennial/Perennial | 1 | 4 | 25 |
| Perennial | 463 | 2390 | 19 |
| NA | 25 | 240 | 10 |
| Mode of Perennation | | | |
| Aerial | 76 | 371 | 20 |
| Bulbaceous | 5 | 31 | 16 |
| Cormous | 2 | 38 | 5 |
| Cormous/Bulbaceous | 0 | 4 | 0 |
| Rhizomatus | 73 | 303 | 24 |
| Rhizomatus/Bulbaceous | 0 | 2 | 0 |
| Rhizomatus/Tuberous | 1 | 5 | 20 |
| Tuberous | 27 | 248 | 11 |
| NA | 341 | 2087 | 16 |
| Mode of Nutrition | | | |
| Carnivorous | 9 | 64 | 14 |
| Free-living | 486 | 2812 | 17 |
| Hemi-parasitic | 2 | 11 | 18 |
| Parasitic | 4 | 6 | 67 |
| Semi-parasitic | 0 | 1 | 0 |
| NA | 24 | 195 | 12 |
| Naturalised Status | | | |
| Alien to Western Australia | 25 | 464 | 5 |
| Mixed | 6 | 25 | 24 |
| Native to Western Australia | 494 | 2600 | 19 |
| Total # Species | 525 | 3089 | |

2.4.2 Geographic context

Forty-one vegetation complexes have been mapped for the NJF, with 11 falling in the study area (Table 2.2). The 11 types captured are dominant types for the forest - covering 49% of the NJF. Dwellingup complex 1 is the most abundant in the NJF, within the study area and in the unsampled area within the survey bounds (full NJF = 2290 km², sampled = 185 km², unsampled = 4.4 km²) (Table 2.2). Ninety percent of study sites fall within five vegetation complexes (Dwellingup complex 1 & 3, Yarragil

complex 1 & 2 and Murray complex). Within the survey bounds 17.8km² was excluded from sampling, covering 10 vegetation types. Most unsampled types were relative to the area sampled except for Swamp complex (2.4% of NJF captured but 8% unsampled within study area) (Table 2.2, Appendix B5).

Table 2.2 Vegetation complexes give a broad understanding of vegetation patterns in the dataset relative to the larger landscape and highlight under-sampled areas. Each unit was divided into two categories: (1a) the area and (1b) percent captured by the dataset relative to the northern jarrah forest (NJF); and (2a) the area and (2b) percent within the survey zone that was not sampled. The swamp complex was excluded to a higher degree than other types.

| Vegetation Complex | Total area - NJF (km²) | (1a) Study area (km²) | (1b) Percent captured (%) | (2a) Unsampld (km²) | (2b) Percent excluded (%) |
|---------------------------|--|---|----------------------------------|---------------------------------------|----------------------------------|
| Helena complex | 171 | 1 | 0.7 | NA | NA |
| Goonaping complex | 173 | 0 | 0.2 | 0.01 | 4 |
| Cook complex | 227 | 20 | 9 | 0.6 | 2.8 |
| Swamp complex | 456 | 11 | 2.4 | 0.9 | 8 |
| Yarragil complex 2 | 603 | 66 | 11 | 3.4 | 5 |
| Murray complex | 705 | 47 | 7 | 3.0 | 6 |
| Dwellingup complex 3 | 811 | 24 | 3 | 1.6 | 6.6 |
| Yarragil complex 1 | 848 | 71 | 8.3 | 3.9 | 5.5 |
| Dwellingup complex 2 | 1254 | 1 | 0.09 | 0.03 | 2 |
| Pindalup complex | 1967 | 6 | 0.3 | 0.1 | 1.5 |
| Dwellingup complex 1 | 2291 | 185 | 8.1 | 4.4 | 2.4 |
| Other | 13054 | NA | NA | NA | NA |
| Total | 22560 | 432 | NA | 18 | NA |

2.4.3 Conservation context

Ten of the 322 priority (P) taxa were captured including a P1 (*Hemigenia rigida*) and P2 species, plus six P4's. This constitutes just 2% of the database, relative to priority taxa forming 10% in the full flora (Appendix B6). The NJF is ranked 5th out of 47 IBRA subregions based on the number of priority taxa as a function of protected area (Table 2.3). The Katanning region ranks first, with over 400 priority taxa and just 1.7% of the region protected, and the Fitzgerald 10th with a very high number of priority taxa (479) that are quite well-protected (28%).

Table 2.3 Ranked conservation statistics for 10 IBRA subregions in Western Australia. Highly cleared areas in the wheatbelt have both a high number of priority taxa and low protection (AVW01, AVW02). The northern jarrah forest (NJF) ranked 5th, with an adequate area protected (10.4%) but moderate number of priority taxa (322). IBRA = Interim Biogeographic Regionalisation for Australia.

| IBRA subregion | IBRA code | Number of priority taxa / protected area (%) | Number of priority taxa | Percent of IBRA subregion protected (%) |
|------------------------|-----------|--|-------------------------|---|
| Katanning | AVW02 | 257 | 429 | 1.7 |
| Merredin | AVW01 | 164 | 553 | 3.4 |
| Fortescue | PIL02 | 90 | 50 | 0.6 |
| Western Mallee | MAL02 | 39 | 390 | 10.0 |
| Northern Jarrah Forest | JAF01 | 31 | 322 | 10.4 |
| Perth | SWA02 | 30 | 359 | 11.9 |
| Southern Jarrah Forest | JAF02 | 25 | 421 | 17.0 |
| Eastern Murchison | MUR01 | 25 | 167 | 6.8 |
| Fitzroy Trough | DAL01 | 22 | 38 | 1.8 |
| Lesueur Sandplain | GES02 | 20 | 356 | 18.0 |
| Dandaragan Plateau | SWA01 | 17 | 132 | 7.5 |
| Fitzgerald | ESP01 | 17 | 479 | 27.8 |

2.5 Discussion

As ecology meets the information age platforms to archive and share data are in development globally (Hampton *et al.* 2013). This data provides enormous opportunities for collaborative and desktop research, but data elements (e.g. grain and extent, presence-absence vs presence-only) are not consistent across projects or geographic locations (Kujala *et al.* 2018). Defining well-framed research questions answerable by available data reverse-engineers the typical scientific process but, is an important element of modern science (Hampton *et al.* 2013; Rosenheim and Gratton 2017). We take a large survey dataset that was collected for pre-mining checks and quantify its limits within a larger ecological context. Using regionally available data, the taxonomic, trait and geographic limits of the dataset were defined, and a conservation context developed. Our framework provides one approach to dealing with the limitations of re-purposed data.

The scale and extent of data was medium resolution (120 m²) but restricted. Initial expectations due to the size of the dataset (30,000 plots) were tempered by the limited taxonomic and geographic portion of the flora captured (Table 2.1 and 2.2, Figure 2.2, Figure 2.3). With the shrub and tree flora more well-represented than the herbaceous, questions related to patterns of diversity or endemism became inappropriate, unless specifically within a woody context. The dominance of just five vegetation complexes from 41 warns against extrapolating outside of the study area for community-level analyses. The low number of priority taxa in the study area (10 of 322) means the dataset also cannot be used to

study conservation species, even though the NJF is ranked as 5th within WA as an area with a high number of priority species (322) relative to area protected (10%). Finally, the data is not temporal (no repeat measurements were made and many areas were cleared post-survey), preventing the use of the data as benchmarks for ongoing climate-change work.

2.5.1 *Potential applications*

Potential applications of the database include classification work, species distribution modelling (SDMs), vegetation community modelling and using woody species to investigate heterogeneity differences between major vegetation complexes. Questions in these areas will be constrained by the extent of plots, but plots are well-structured and occur at a similar scale to widely available topographic data. Statistical considerations include that any classification work will not be based on the full enumeration of plots (Peet and Roberts 2013). However, this is also the case for previous classifications (Havel 1975a) and provides an opportunity to investigate past typologies with vastly more data. Pros and cons for species and community modelling follow current best-practice, including that presence-absence data is more robust (e.g. common woody species) than presence-only (annuals) (Elith *et al.* 2006). The dataset also occurs at a similar spatial resolution to available topographic layers (120 m vs 30 m) but not climate data (1 km, CSIRO), which will impact model decision-making (Miller and Franklin 2010; Kujala *et al.* 2018). Unfortunately, the database faces the same issues as much correlative modelling; in that there is a lack of eco-physiological measurements, physical and chemical soil data and biotic variables (Wisz *et al.* 2013; Mod *et al.* 2016).

Data is now regularly used outside of its original research despite challenges in amalgamating different datatypes, units and layers (Tichý *et al.* 2014; Franklin *et al.* 2016; Kujala *et al.* 2018). The American system for vegetation data and classification ‘VegBank’ and ‘EcoVeg’ successfully brought together data from multiple plot sizes (0.01-0.1 to 1 ha in tropical vegetation), environmental and site variables (Faber-Langendoen *et al.* 2014). It now forms the basis of the American, Canadian and NatureServe’s International Vegetation Classification systems (Faber-Langendoen *et al.* 2018). Applications are being integrated into ecological monitoring and assessment programs across the Americas, with feedback into the system to further refine classification units (Faber-Langendoen *et al.* 2018). Studies based on VegBank include a revision of longleaf pine vegetation associations (Palmquist *et al.* 2015) and an analysis of Great Lakes alvar vegetation (Reschke *et al.* 1999). Studies derived from BIEN include drivers of plant functional group dominance (Engemann *et al.* 2016), linking environmental filtering to biogeography (Blonder *et al.* 2015) and inferences of species associations (Morueta-Holme *et al.* 2016). In Australia, the Biodiversity and Climate Change Virtual Laboratory (BCCVL) acts as both an archive for species and environmental data, a training portal and an analysis tool for a range of SDM (Hallgren *et al.* 2016).

Conservation planning encapsulates many of these issues. The wide-spread use of optimisation software has powerfully brought together multiple datatypes while also being hampered by it (Lehtomäki and Moilanen 2013). SDMs are an important source of information when occurrence data is lacking, but with variable margins of error relative to the quality of input data (Elith *et al.* 2006). Results should always be interpreted in the context of data input quantity and quality, but the ideal is often hard to achieve in

practical decision making (Lehtomäki and Moilanen 2013). Surrogates are essential, each with their own value and limits (Pressey 2004). However, knowing exactly what data is missing is a sharp tool for promoting further conservation effort, especially in areas where data-infrastructure is still lacking (McKenzie *et al.* 1996; Gellie *et al.* 2018).

2.6 Conclusion

The merit of this approach for working with repurposed datasets is that it provides a relatively simple and structured template to define limits. Online herbaria and museum collections, or the global biodiversity information facility (GBIF) are options for baseline data. Simple elements that form good modelling practise can be outlined quickly (e.g. datatype, coverage, consistency with environmental layers and appropriateness for different statistical methods) (Elith and Graham 2009). Framing data also means model applications sit within these limits, potentially improving end-user confidence in outputs (Addison *et al.* 2013; Guisan *et al.* 2013). We have used a single dataset within a region, but creating a context when amalgamating datasets will also help to define appropriate research questions when working at larger scales. Finally, a framework highlights where ecoinformatics research needs validation by traditional, experimental research and complementary areas between the two (Rosenheim and Gratton 2017). Understanding the limits of available data (and its applications) can then be used to rationalise further survey work and argue the case for experimental studies. This is especially pressing when climate change models can't realistically capture on-ground effects at a resolution relevant to threatened or vulnerable species.

Chapter 3 **Vegetation classification in subtle landscapes – implications for predicting climate change impacts.**

3.1 Abstract

In the northern jarrah forest of southwest Australia, a regional decrease in rainfall has caused significant streamflow decline. Modelling provides one approach for understanding how these changes will affect riparian plant communities. In the northern jarrah forest, community-level vegetation mapping has been done using site vegetation types (SVT), while a large amount of additional plot-based survey data has recently become available. To determine the best-available units for climate change modelling work, an updated classification using this data was compared to SVTs. Non-hierarchical clustering with monte carlo simulation was used to group the data. Results showed indistinct types, with dominant species overlapping between groups and uncommon species classified to both core and outlying types. Aspect, curvature and topographic wetness were significantly different between groups, but the effect of other remotely sensed topographic variables was small. Additionally, the groups did not correlate well to SVTs, clustering into three groups in the SVT typology. Results suggest that classification based purely on floristics in this system is difficult, due to continuity in vegetation patterns and their underlying environmental gradients. Consequently, SVTs provide a stronger framework for modelling potential climate change impacts at the community level and are recommended for further work.

3.2 Introduction

The forefront of modern vegetation mapping and modelling is now at the interface of big data and machine learning tools (Franklin *et al.* 2016), while a key challenge is synthesising new data with previous knowledge (Tichý *et al.* 2014). Shared data archives include the European Vegetation Archive (EVA), VegBank, EcoVeg, the Botanical Information and Ecology Network (BIEN), S-plot and more (Wiser *et al.* 2001; Enquist *et al.* 2009; Faber-Langendoen *et al.* 2014; De Cáceres *et al.* 2015; Chytrý *et al.* 2016; Faber-Langendoen *et al.* 2018; Bruelheide *et al.* 2019). VegBank and EcoVeg are enabling the consistent classification of vegetation associations across Northern America within a hierarchical framework, describing floristic and structural attributes at multiple scales (Faber-Langendoen *et al.* 2018). In Australia, over-arching mapping systems include the Interim Biogeographic Regionalisation for Australia (IBRA) and the National Vegetation Information System, which provide information at the continental scale (Thackway and Cresswell 1995; McKenzie 2008; NVIS Technical Working Group 2017). However, a lack of cohesion between federal and state, and inter-state vegetation systems; and the sheer scope of mapping continental Australia at fine scales has prevented consistent mapping at lower resolutions (McKenzie 2008; Gellie *et al.* 2018).

In addition to the improvements in the availability and analysis of vegetation data, remote-sensing provides environmental data at scale, enabling the automation and mapping of community types across large areas (e.g. Su *et al.* 2016). NASA's global one-arc second (30 m) resolution Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) is the most downloaded dataset in NASA history (NASA 2011; Buckley 2019), providing topographic data at medium resolution. More locally,

LiDAR data can now take measurements at cm to meter resolution to generate a range of terrain variables (e.g. Keppel *et al.* 2017). Correlation between remote and ground-survey data is constantly improving, providing support for multiple management needs across areas too large to field survey (Harwood *et al.* 2016; Lawley *et al.* 2016; Ware *et al.* 2018).

One application of automated vegetation mapping is predictive modelling to study the effects of climate change. While less common than species modelling, a community-approach has several benefits over individual models, particularly if many species are recorded infrequently (Ferrier and Guisan 2006). These include averaging individual species responses to describe broad patterns of change and providing a distinct unit for interpretation and management. Three methods are possible: predict then assemble, assemble then predict, or simultaneously (Ferrier and Guisan 2006; House *et al.* 2012; Molloy *et al.* 2016). Updating the accuracy of local typologies is important if climate change modelling exercises are to be meaningful and relevant to decision-makers (Addison *et al.* 2013; Guisan *et al.* 2013).

The northern jarrah forest region of SWA provides a test case for updating a typology and further climate change modelling work. Part of the SWA Mediterranean climate ecosystem (Sala *et al.* 2000; Klausmeyer and Shaw 2009), the forest have been mapped at several scales (Appendix C5) (Churchward and McArthur 1980; Havel 2000). Recently, a large amount of plot-based vegetation data (30,000 plots) has become available from bauxite mining company Alcoa. This dataset, with NASA's 30 m SRTM DEM, provide the opportunity to run an updated classification and investigate whether improvements in data-availability affect the definition of community types.

To create an updated typology, test its utility, make a comparison to previous work and determine the best units for further modelling this chapter aimed to:

- 1) identify an appropriate and practical clustering strategy that would handle the size of the dataset while creating a robust partition. Due to the number of plots (30,000) most common, hierarchical clustering algorithms were expected to be less effective than a non-hierarchical approach.
- 2) use topographic variables generated from NASA's 30 m SRTM DEM to describe differences in groups (1). Changes in key factors like slope, elevation and aspect are expected to reflect differences in community composition.
- 3) compare the results found in (1) with current community-classification units (site vegetation types) while accounting for differences in data-input and methodology between the typologies. As previous work occurred over a larger spatial area, using different input data and analyses methods, both similarities in general groupings and new, distinct community types are expected.

3.3 Methods

3.3.1 Study site

The NJF has a Mediterranean climate, with an average annual rainfall of 600-1300 mm that falls across an East-West gradient in the forest (Havel 1975b; Gentilli 1989). A dry-sclerophyll forest, it is significant globally, occurring in the SWA biodiversity hotspot (Myers *et al.* 2000; Hopper *et al.* 2015). The

vegetation is characterised by subdued topography, an east-west rainfall gradient and regular fire incidence (Burrows *et al.* 2019). The forest also has a history of logging and mining (1834-current) that has significantly impacted structure (Churchward and McArthur 1980; Wardell-Johnson *et al.* 2015). Floristically, the forest overstorey is dominated by two tree species (*Eucalyptus marginata*, (jarrah) and *Corymbia calophylla* (marri)) but has a complex and diverse understorey (3096 taxa) (Dell *et al.* 1989; Western Australian Herbarium 1998-). Understorey vegetation has been described as a multidimensional continuum with semi-discrete community types (Havel 1975a).

The climatic and hydrological regimes of the forest are shifting; rainfall has decreased 12-15% and temperature increased 0.4 C since the mid-1970's (Bates *et al.* 2008). This has led to decreased surface run-off (20-100%), falling water-tables and shifts from perennial to ephemeral streams (Petrone *et al.* 2010; Hughes *et al.* 2012). High summer evapotranspiration offsets winter rainfall, which is predicted to further decrease under climate change. Alcoa, a bauxite mining company, has a mining tenement covering a large proportion of the forest (Figure 3.1).

3.3.2 Floristic data

To meet statutory requirements Alcoa has been collecting floristic data in the NJF for almost 30 years (1991-ongoing). Vegetation mapping now spans 432 km² with 31,000 plots on a 120 m² grid that follows ore exploration (Figure 1). At each grid point, tree species are recorded within a 20 m radius and understorey taxa within a 5 m radius, and cover-abundance (1-5) and stress (1-5) scored (Mattiske 2012). Collected plant specimens were dried, fumigated and identified in accordance with the requirements of the West Australian Herbarium (Mattiske 2012). The conservation status of taxa were checked using current Government Gazette lists (Western Australian Herbarium 1998-). Each plot in the dataset was also assigned a 'Site Vegetation Type' (SVT) (Havel 1975a; 1975b) (Figure 3.1).

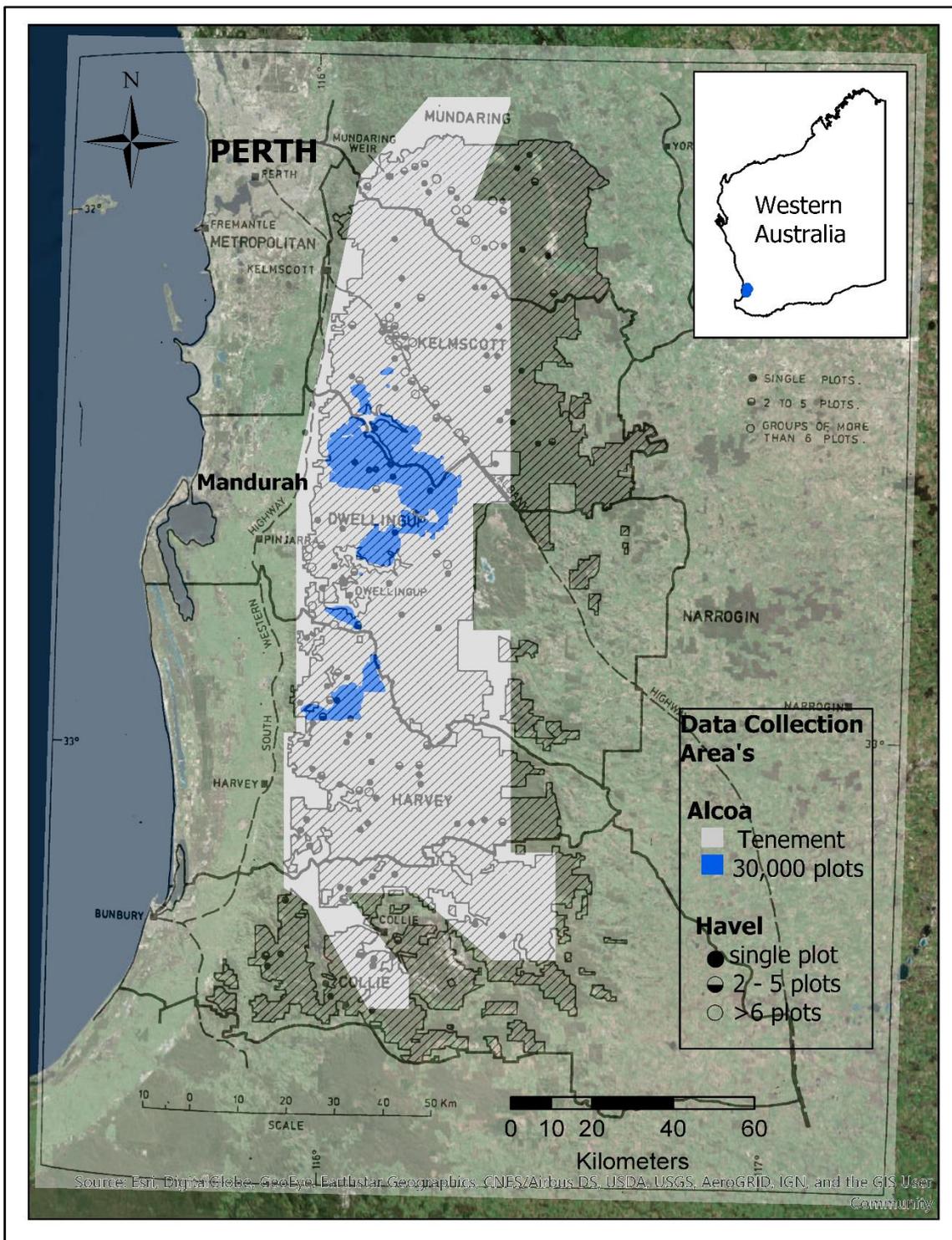


Figure 3.1 Vegetation mapping of site vegetation types in the northern jarrah forest has occurred at two scales: 1) The original survey by Havel (1975) to delineate the SVTs covered a large extent of the forest (hatched) in 320 plots (circles). In contrast, the plot data collected by Alcoa (1991-2015, blue) covers a much more restricted area but was intensively sampled (30,000 plots).

3.3.3 Environmental data

Environmental variables were collected at each plot: soil type, topography, outcrop type and amount, dieback (*Phytophthora cinnamomi*) presence/absence and impact, years since fire, log debris and number

of tree stumps. Additional environmental variables were generated in the System for Automated Geoscientific Analyses (SAGA, v 2.1.2) (Conrad *et al.* 2015) using NASA's one arc-second (30 m) Shuttle Radar Topography Mission (STRM) hydraulically enforced digital elevation model (DEM) (NASA 2011).

Topographic variables include 'distance to channel base network', which is calculated from an inverted DEM and negatively correlated with elevation (high valley depth values indicate a larger distance from ridge lines, or lower valley position). Distance to valley depth is an inverted ridge-top distance. Curvature and convexity both measure the roundness or hollowness of a surface, with convexity indicating the number of convex shapes within a set radius while curvature measures the convex or concavity of a slope. The SAGA wetness index (SWI) was used to represent flow across the study area. SWI is similar to the 'Topographic Wetness Index' (TWI) (Beven and Kirkby 1979) but is based on a modified catchment area calculation which does not conceptualise flow as a thin film (Boehner *et al.* 2002).

Aspect in degrees must be transformed for further analysis (Roberts 1986). Northness is derived from $\cos(\text{aspect})$, with values close to 1 if the aspect is northward, or -1 if the aspect is southward. Alternatively, eastness is a function of $\sin(\text{aspect})$ with values close to 1 represent east-facing slopes. Environmental data were tested for co-correlation using the package 'corrplot' (Wei *et al.* 2017) and each variable from a pair with a correlation of more or less than 0.5 removed (Appendix C1).

3.3.4 Data cleaning and analysis

A floristic information management system was developed to prevent data corruption, data-entry error and ensure future data additions fit system requirements (Trotter *et al.* 2018). In addition to the technical data-cleaning process, 337 species were excluded from the original Alcoa species list due to data-entry errors, sub-species updates and incorrect species identification. Singletons (a single species record) and doubletons (only two species records) were also removed from the dataset for this analysis (186 species) but remain in the original database. Annuals, biennials and alien taxa were removed as surveys were not conducted consistently in the flowering season. In total 435 species remained in the modified dataset prepared for this analysis. Floristic data was then summarised in R (version 3.4.2) and species richness, Shannon diversity and a histogram of species abundance produced.

Bray Curtis with non-hierarchical clustering was used to analyse the data (PATN, version 3.12). Bray Curtis is a similarity (rather than 'distance') metric and accounts for both species similarities (common presence) and differences (common absence) (Roberts and Peres-Neto 2017). However, it tends to underestimate true distance and gives the same importance to absolute differences in abundance, irrespective of the order of magnitude of the abundances (e.g., the difference between one and two, and 456 and 457 are weighted the same) (Borcard *et al.* 2011).

Cluster analysis was done using the non-hierarchical allocation algorithm 'ALOC' in PATN (version 3.12) (Belbin 1987; Belbin and McDonald 1993). Non-hierarchical algorithms have benefits over hierarchical methods in large datasets and are appropriate for the data type (presence-absence) and size (30,000 plots). By calculating object-group associations (not all object-object associations) ALOC is effective in clustering large datasets. The algorithm allocates objects to a set of seeds – starting with the

first object in the dataset and iterating from there (although prior groups can also be used). The number of groups produced is controlled by an allocation radius (an association value between zero and one that acts as a threshold for the generation of additional seeds). The number of iterations is \sqrt{N} – with more iterations required to reach optimality as the size of the dataset increases. The algorithm is reasonably insensitive to seeds and has been tested against hierarchical methods and shown to have good recovery of clusters (Belbin and McDonald 1993).

PATN also has visualisation options (including a dendrogram of group associations), statistics for cluster evaluation and a fast processing time (one simulation took an average of 20 minutes to complete) relative to other packages (Roberts 2015).

3.3.5 Cluster analysis and evaluation

A site by species matrix was imported into PATN and the non-hierarchical clustering algorithm ‘ALOC’, with Bray Curtis dissimilarity (Bray and Curtis 1957) used to produce 99 groups. To minimise variation in cluster output, plot order was randomised in R and the algorithm run 100 times (C3). To select the optimal number of final groups from the simulations; quadrat membership per group, for each run, was amalgamated into a master matrix and a true/false loop used to assign presence. Membership was then compared at an 80, 50, 40 and 30% threshold (the number of times out of 100 that a quadrat occurred in the same group) (C3) (Harrell 2019). Species frequency, mean species abundance and the mean of each environmental value were calculated per group. nMDS was then used to select the final partition (Oksanen *et al.* 2015).

The optimal number of clusters for the final partition was determined using silhouette width (a measure of the degree of membership of an object to its cluster (package ‘cluster’, R)) (Maechler *et al.* 2012) (Borcard *et al.* 2011). Cluster evaluation was done using PATN’s in built evaluation tools. These include the Kruskal-Wallis statistic (significance of the discriminating environmental variables), Principal Component Correlation (PCC) and Monte-Carlo attributes in an Ordination (MCAO). MCAO is a permutation test that tests the ‘robustness’ of PCC. Boxplots and nMDS were used to visualise group association with environmental variables (‘vegan’, R).

The difference between previous and current vegetation mapping units was examined using summary tables and a confusion matrix in R (C3). Basic tables were produced in Excel (2016) and map figures in ArcGIS Pro (version 2.2).

3.4 Results

Species frequency is heavily skewed, with the 20 most common species comprising 55% of occurrence observations. Of the 436 taxa present, only 20% were recorded more than 1000 times, but comprise 92% of all records. Conversely, 67% of species were recorded less than 200 times and represent 2% of total records (Figure 3.2). *Eucalyptus marginata* (jarrah) and *Corymbia calophylla* (marri), the dominant overstorey tree species were recorded in almost all quadrats (27,228 and 23,204 records). Other key dominants were *Macrozamia riedleii*, a cycad (19610 records), *Banksia grandis* a mid-story tree (16,428), *Lasiopetalum floribundum* a Malvaceae shrub (14682) and *Hibbertia commutata* a shrub (14405). Less

common species include the trees *E. rudis* (133), *B. seminuda* (36), and *Allocasuarina huegeliana* (3), plus *Nuytsia floribunda* (20), a small parasitic tree. Fifty families are represented across the least common species (<200 records) including Fabaceae, Proteaceae and Cyperaceae (Appendix B3). All of the priority species (taxa listed as requiring further survey work, with the potential to be listed as a rare species) collected in the database consist of <200 records, including *Tetratheca phoenix* (a priority two). Mean species richness per plot was 17 (SD = 4.5), and mean Shannon Diversity was 0.033 (max = 3.16, min = 0.0004).

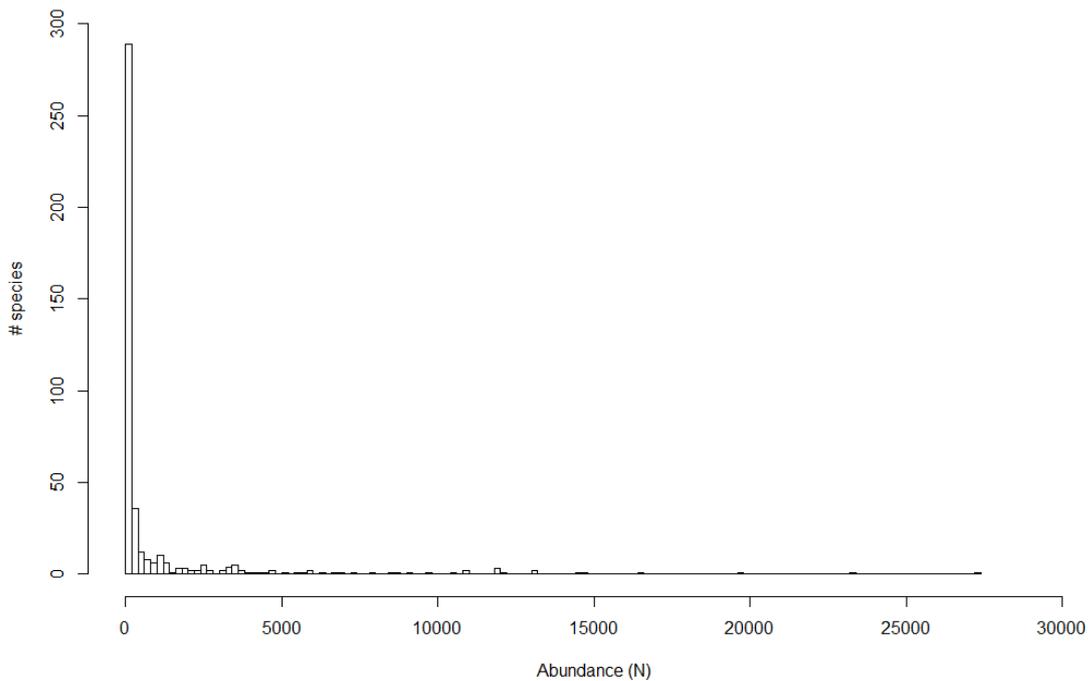


Figure 3.2 Histogram of number of species vs. abundance for the Alcoa dataset indicating the dominance of a few key species and occurrence of many uncommon species.

3.4.1 Simulation of final groups

At an 80% threshold, 6518 final groups were produced, compared to 50% with 574 final groups and many singletons (single quadrat per group), 40% (162) and 30%, with 77 final groups and few singletons (Appendix C3). An nMDS of the 77 and 162-group partition showed little difference in group structure (Figure 3.3). The number of outliers remains consistent, with variability within the central groups smoothed in the 77-group approach. Stress was significantly lower for 77 groups (stress = 0.15 c.f. 162 group stress = 0.22) (Figure 3.3).

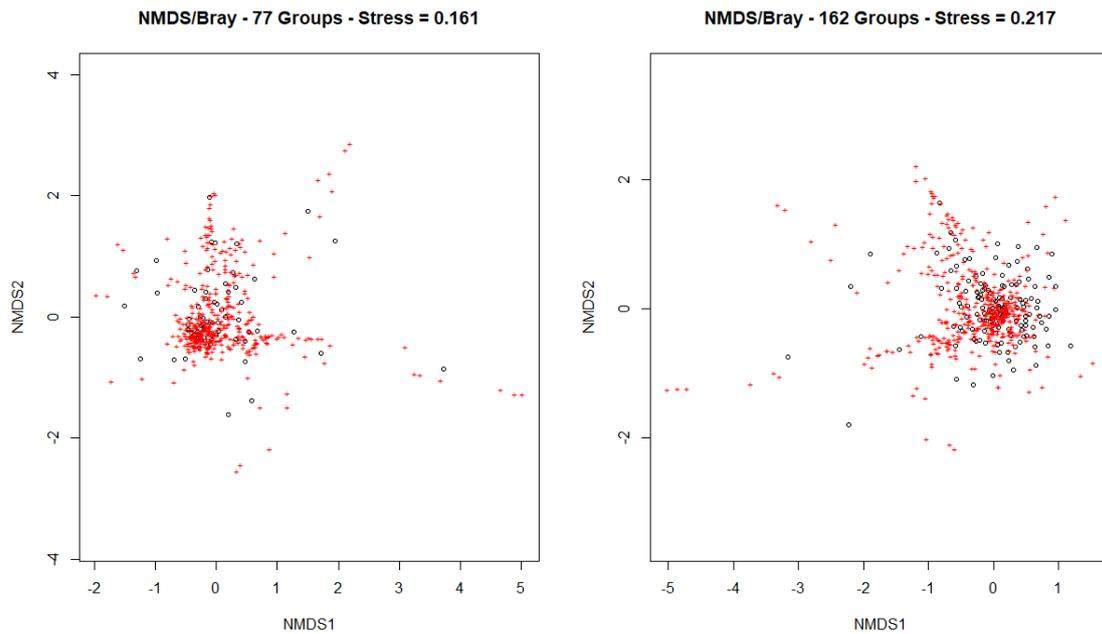


Figure 3.3 NMDS of 77 and 162 groups (black circles) with species scores (red). Stress is much lower for the 77-group classification.

3.4.2 Testing old units against the new: SVT vs ALOC

A comparison of SVTs to an 11-group ALOC typology shows the majority of SVTs clustering into one ALOC major and two minor groups (Table 3.1). While SVTs are heavily skewed towards type ‘S’ and ‘T’ (64% of plots), the 11-group ALOC classification is even more unbalanced, with 95% of plots in group ‘11’ and seven groups with <10 quadrats each (Table 3.1). Group 1 contains the largest proportion of SVTs associated with riparian conditions (A, C, W) (Table 3.1).

Table 3.1 Confusion matrix for Havel’s site vegetation types (SVT) vs an 11-group ALOC-derived classification. ALOC does not separate the variation described by SVTs well, only delineating two core groups across the categories.

| SVT | ALOC Group | | | | | | | | | | |
|----------|------------|---|---|---|---|----|---|---|---|----|-------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| <i>A</i> | 471 | 0 | 0 | 0 | 2 | 4 | 0 | 6 | 0 | 0 | 104 |
| <i>C</i> | 487 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 251 |
| <i>D</i> | 86 | 1 | 0 | 0 | 0 | 76 | 1 | 2 | 0 | 0 | 897 |
| <i>E</i> | 6 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 61 |
| <i>G</i> | 1 | 0 | 0 | 0 | 1 | 52 | 0 | 0 | 1 | 26 | 94 |
| <i>P</i> | 19 | 0 | 0 | 0 | 0 | 12 | 0 | 0 | 0 | 0 | 5960 |
| <i>Q</i> | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 80 |
| <i>R</i> | 4 | 0 | 0 | 0 | 0 | 6 | 0 | 0 | 0 | 0 | 187 |
| <i>S</i> | 26 | 4 | 0 | 0 | 2 | 6 | 0 | 1 | 0 | 1 | 12215 |
| <i>T</i> | 9 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 5985 |
| <i>W</i> | 98 | 0 | 1 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 1199 |

3.4.3 ALOC: two vs 30-group classification

Silhouette width, an indicator of cluster fidelity, found the optimal number of groups to be two (silhouette width = 0.3). However, based on visual inspection a 30-group classification was also selected for further analysis (UPGMA, row-fusion = 0.491, Figure 3.4). Sites are unequally distributed: group B holds 95% of sites and the largest sub-group (26) contains 404 species and 24253 sites (Figure 3.4). The dominant species in group B mirror the most-common species in the database (jarrah, marri, zamia), interspersed with other well-known taxa like *Xanthorrhoea preissii* (grasstrees), *Bossiaea aquifolium*, *Allocasuarina fraseriana*, *Persoonia longifolia* and the bracken fern *Pteridium esculentum* (Table 3.2). One sub-group (29) also includes *B. littoralis*, *H. angustifolium* and *Lomandra sonderi*, moisture-associated species (Appendix C4). Species unique to group A (1421 sites) include *E. patens*, *Taxandria linearifolia*, *Banksia littoralis* and *Hypocalymma angustifolium* which are all associated with stream banks, valleys and seasonally wet areas (florabase 26/5/19). Key dominants from group A are also abundant in B, notably jarrah, marri, grasstrees, zamias and bracken fern (Table 3.2).

At the 30-group level, species differentiation is difficult, with mixing in the proportion of dominant species present across the groups, interspersed by the occurrence of uncommon taxa (e.g. group 18 v 19, Appendix X). Additionally, half of the groups are very small with a low abundance of taxa (≤ 60 records) relative to the size of major groups (1000 – 400,000 records) (Appendix X). Despite the lack of clear distinctions in the species data, there were significant environmental differences between groups including years since fire, curvature, aspect (eastness) and topographic wetness (Table 3.3).

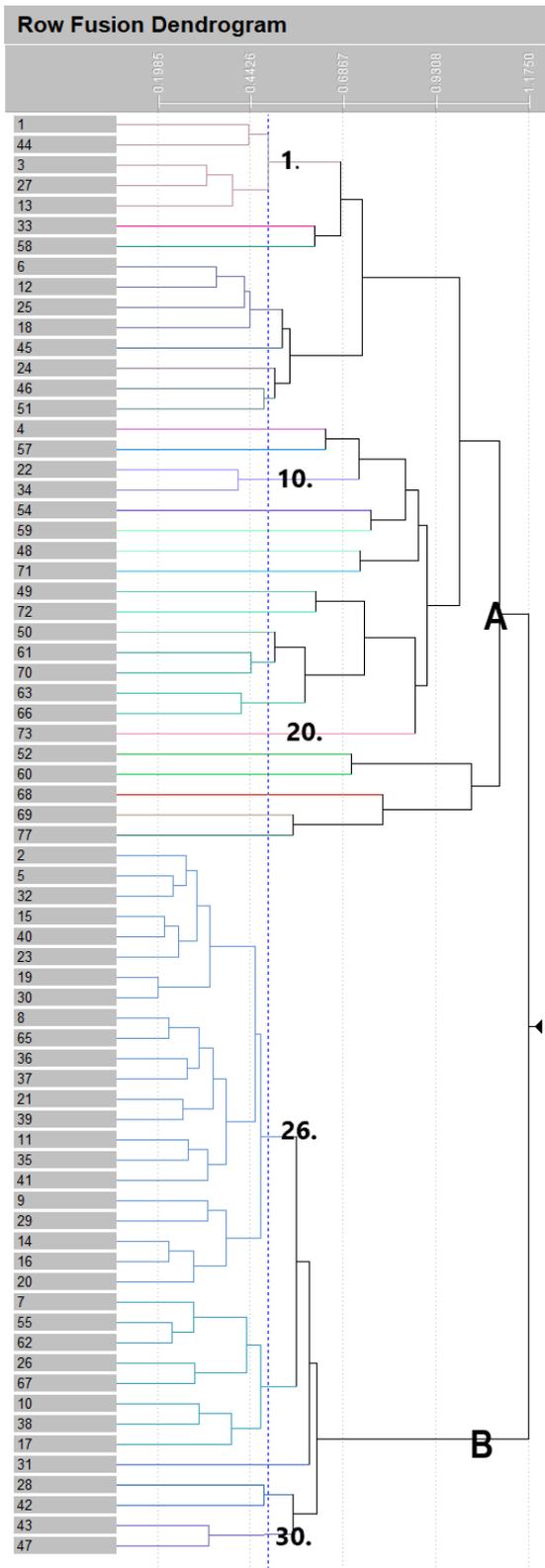


Figure 3.4 UPGMA dendrogram of ALOCs final 2 and 30-group classifications, showing the large degree of separation at the two-group scale and more subtle patterns within the groups. Branch labels = major (A and B) and minor (1-30) groups. Based on Bray Curtis of mean species abundance for the 77-group partition.

Table 3.2 The 25 most abundant taxa are almost identical when overall frequencies for the dataset are compared to group A results (the dominant group in a 2-type classification of the data). Group B is largely distinct, capturing the main riparian species in the forest. Grey = shared species, bold = species unique to group B only.

| All Sites | Freq | Group A | Freq | Group B | Freq | Name |
|-----------|-------|---------|-------|---------------|------------|----------------------------------|
| EUCMAR | 27273 | EUCMAR | 26473 | EUCPAT | 884 | <i>Eucalyptus patens</i> |
| CORCAL | 23247 | CORCAL | 22542 | TAXLIN | 874 | <i>Taxandria linearifolia</i> |
| MACRIE | 19649 | MACRIE | 19265 | EUCMAR | 800 | |
| BANGRA | 16448 | BANGRA | 16340 | XANPRE | 769 | |
| LASFLO | 14701 | LASFLO | 14427 | CORCAL | 705 | |
| HIBCOM | 14427 | HIBCOM | 14311 | BANLIT | 474 | <i>Banksia littoralis</i> |
| BORFAS | 13179 | PERLON | 13139 | HYPANG | 453 | <i>Hypocalymma angustifolium</i> |
| PERLON | 13178 | BORFAS | 13028 | MACRIE | 384 | |
| LOMSON | 12179 | LOMSON | 12116 | MIRDIL | 379 | <i>Mirbelia dilatata</i> |
| TRYLED | 12000 | TRYLED | 11803 | PTEESC | 358 | |
| TETCAP | 11975 | TETCAP | 11659 | TETCAP | 316 | |
| BOSAQU | 11925 | BOSAQU | 11650 | THOPAN | 305 | <i>Thomasia paniculata</i> |
| XANGRA | 10948 | HIBAMP | 10817 | MELPRE | 289 | <i>Melaleuca preissiana</i> |
| HIBAMP | 10893 | XANGRA | 10726 | LEPTET | 281 | <i>Lepidosperma tetraquetrum</i> |
| XANPRE | 10533 | XANPRE | 9764 | BOSAQU | 275 | |
| PENPEL | 9816 | PENPEL | 9675 | LASFLO | 274 | |
| ALLFRA | 9057 | ALLFRA | 8962 | ACAALA | 257 | <i>Acacia alata</i> |
| BOSORN | 8735 | BOSORN | 8661 | HYPCOR | 244 | <i>Hypocalymma cordifolium</i> |
| PHYCAL | 8539 | PHYCAL | 8442 | EUCMEG | 243 | <i>Eucalyptus megacarpa</i> |
| ADEBAR | 7851 | ADEBAR | 7840 | ASTSCO | 237 | <i>Astartea scoparia</i> |
| PTEESC | 7287 | PTEESC | 6929 | BABCAM | 234 | <i>Babingtonia camphorosmae</i> |
| BANDAL | 6849 | CONSET | 6665 | XANGRA | 222 | |
| CONSET | 6725 | BANDAL | 6663 | GAHDEC | 212 | <i>Gahnia decomposita</i> |
| HOVCHO | 6306 | HOVCHO | 6284 | CYAAVE | 211 | <i>Cyathochaeta avenacea</i> |
| LEUVER | 5970 | LEUVER | 5959 | ACADIV | 207 | <i>Acacia divergens</i> |

3.4.4 Environmental variables

The dominant groups in the classification (26-30, 'B') occupy a restricted environmental space and are not well-separated by the ordination axes (Figure 3.5). The variability in groups A (1-25) occurs across gradients of fire, surface curvature, wetness and aspect. Sin(aspect) ("eastness") is the most-strongly related nMDS variable ($R^2 = 0.47$, $p = 0.001$), with cos(aspect) ("northness") showing a significant correlation using PCC ($R^2 = 0.3$, MCAO = 0%) (Figure 3.6; Figure 3.7). Both sin and cos-aspect are positive for group A (0.32, 0.35 = north-east facing slopes), and slightly negative in group B (-0.04, -0.03, a very slight south-west tendency). Solar radiation is highly correlated with cos-aspect ($cor = 0.77$) and slope and saga wetness index are also well-correlated ($cor = -0.77$) (Figure 3.5). Together, group B indicates higher wetness and valley depth (a negative corollary of elevation, $cor = -0.52$) on reduced slopes (Table 3.3; Figure 3.7). Wetness and convexity are also negatively correlated, with convex areas having lower wetness ($cor = -0.67$). Finally, species richness and diversity are significantly lower in A in both nMDS and PCC (Table 3.3).

The outlying groups are small (7-27 species records) and contain rarer taxa (Appendix C4). Patterns between outlying taxa, remotely sensed variables and environmental factors are evident. *Verticordia huegelii* and *Hakea erinacea* (HAKERI, VERHUE) which occur at high curvature values, are associated with granite outcrops (florabase). Group 23 is distinct, with key species generally found in dry, rocky outcrop sites: *Hakea petiolaris* (HAKPET) outcrops, *Neurachne alopecuroidea* (NEUALO), a rhizomatous perennial grass associated with dry conditions (sandplains, hillslopes, outcrops) and *Cheilanthes austrotenuifolia* (CHEAUS), a fern are found on exposed rocky areas and granitic outcrops (Florabase). *Stylidium pulchellum* is rare, with only two records, while *Kunzea micrantha* and *Aotus cordifolia* are shrubs found in peaty soil, marshes and swamps. *Lasiopetalum cardiophyllum* (LASCAR) is a multi-stemmed priority four shrub with only 23 records.

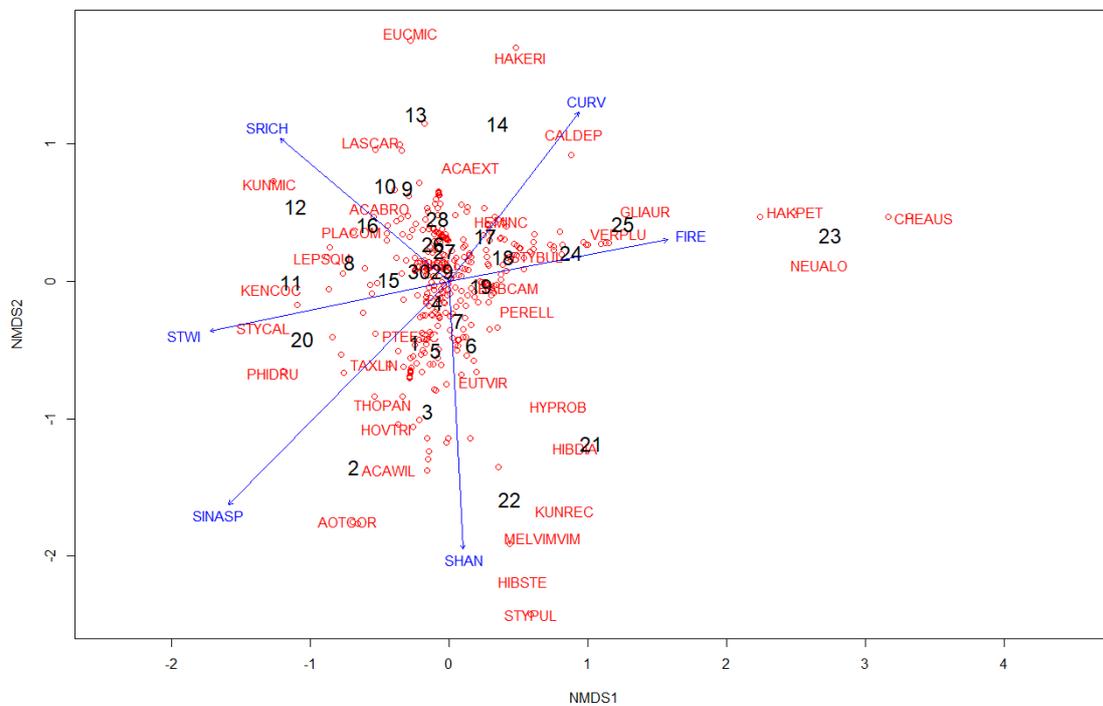


Figure 3.5 NMDS with Bray Curtis on mean species abundance per 30-group, with species (red), groups (black) and significant environmental variables (blue). Stress = 0.172. SRICH = species richness, CURV = curvature, FIRE = years since fire, SHAN = Shannon diversity index, SINASP = eastness, STWI = SAGA topographic wetness index.

Table 3.3 Summary statistics for the 2-group (A, B), and evaluation tests for the 30-group classifications. Aspect, years since fire, curvature and wetness are significantly different between groups in the 30-group classification, as is richness and diversity. PCC: Principle Component Correlation, KW: Kruskal-Wallis, MCAO: Monte-Carlo attributes in an Ordination (% Permuted $R^2 > \text{Actual } R^2$), nMDS – non-metric multidimensional scaling (R^2) with significance (p-val).

| Variable | Code | Group A | Group B | 30 Group Classification Results | | | | |
|---------------------------|--------|---------------------------|-------------------------|---------------------------------|-------------|----------|----------------|--------------|
| | | min-max (mean) | min-max (mean) | PCC (R^2) | KW | MCAO (%) | nMDS (R^2) | nMDS p-val |
| Species Richness | SRICH | 1-26 (12) | 1-41 (17) | 0.53 | 50.5 | 0 | 0.24 | 0.03 |
| Number of Stumps | STMP | NA | NA | 0.44 | 60.4 | 0 | NA | NA |
| Mean Annual Rainfall | RAIN | 904-1197 (1076) | 890-1200 (1077) | 0.32 | 38.4 | 0 | 0.08 | 0.3 |
| Cos-aspect | COSASP | -1-1 (0.35) | -1-1 (-0.03) | 0.30 | 59.0 | 0 | 0.13 | 0.15 |
| Years Since Fire | FIRE | 0-17 (6.4) | 0-22 (5.7) | 0.28 | 48.2 | 0 | 0.24 | 0.03 |
| Shannon Diveristy Index | SHAN | NA | NA | 0.25 | 60.6 | 0 | 0.35 | 0.005 |
| Distance to Valley Depth | VALD | 0-171 (76) | 0-189 (54) | 0.21 | 49.2 | 0 | 0.18 | 0.09 |
| Curvature | CURV | -2.4-1.5 (-0.003) | -3.9-3.0 (0.006) | 0.08 | 46.1 | 11 | 0.22 | 0.04 |
| Convexity | COVX | 36-61 (46.4) | 21-68 (50.5) | 0.05 | 56.1 | 22 | NA | NA |
| Sin-aspect | SINASP | -1-1 (0.32) | -1-1 (-0.04) | 0.04 | 58.1 | 35 | 0.47 | 0.001 |
| SAGA Wetness Index | STWI | 3-8.5 (5.8) | 2.8-9.8 (5.7) | 0.04 | 51.7 | 45 | 0.29 | 0.01 |
| Topographic Wetness Index | WET | NA | NA | 0.02 | 55.6 | 68 | NA | NA |
| Longitudinal Curvature | LCRV | -0.0051-0.0052 (-2.8e-04) | -0.007-0.007 (-2.1e-05) | 0.02 | 45.8 | 82 | 0.12 | 0.18 |
| Number of Sites | | 1421 | 27033 | | | | | |

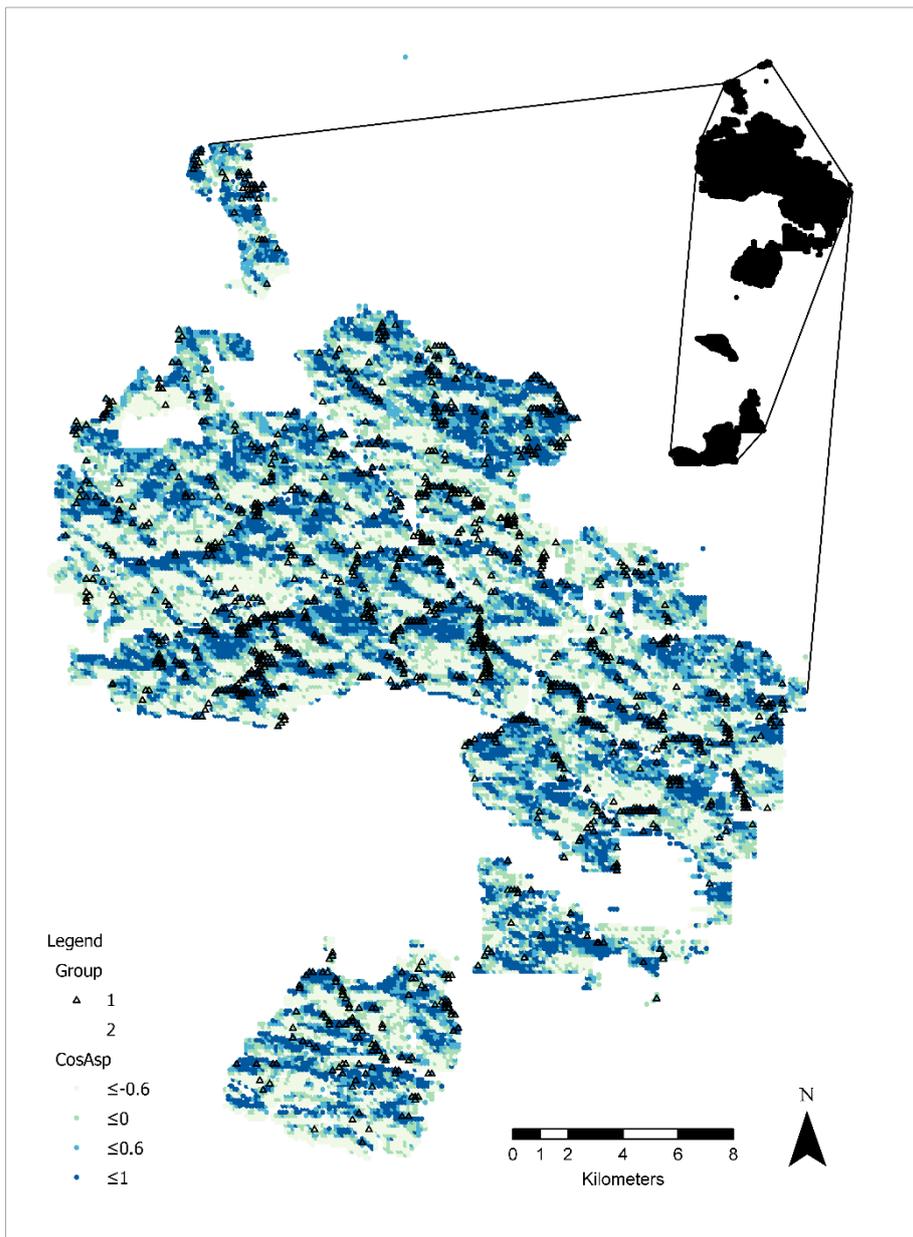


Figure 3.6 The most statistically robust classification for the study area is a two-group typology, that has a strong correlation with aspect (northness). Group 1 (black triangles, 1421 sites) is a riparian group sitting low in valleys along south-facing slopes, while group 2 (clear, 27,000 sites) forms a large cluster with several hundred species that spans the mid-upper slopes and plateaus.

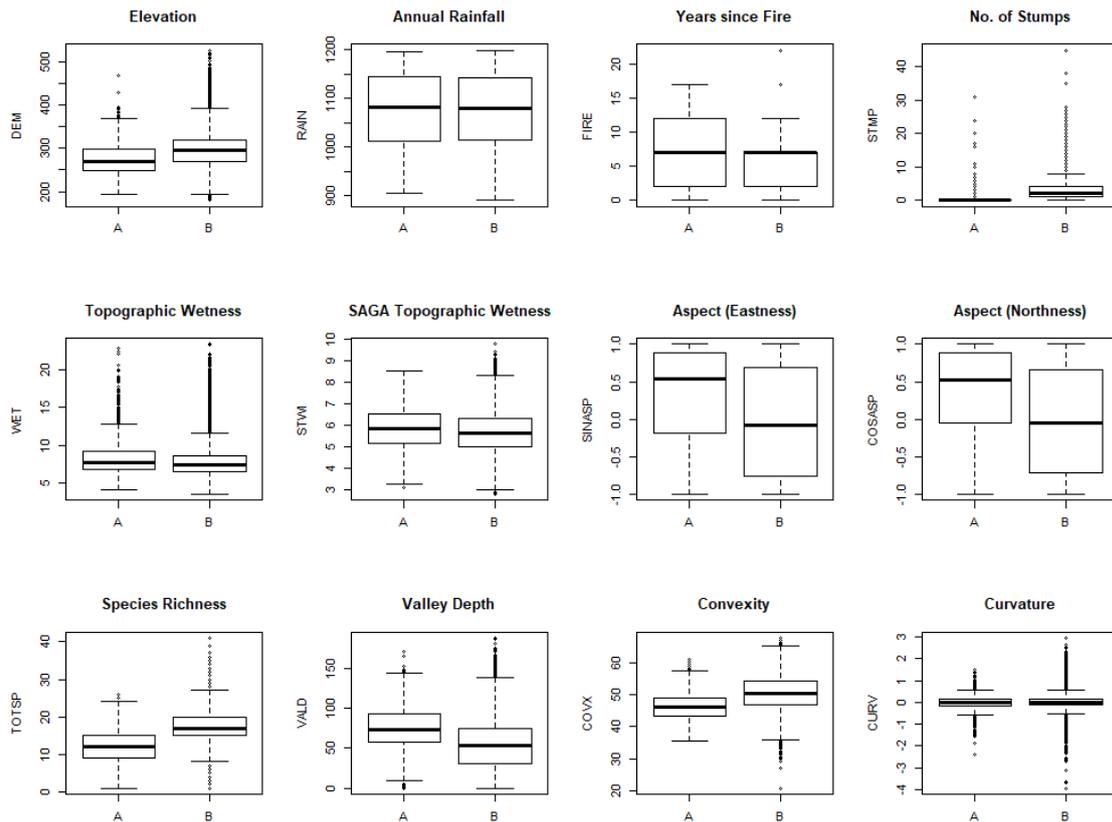


Figure 3.7 Boxplots of environmental variables for the two-group typology. Group A (1421 sites) includes predominantly riparian species and is found lower in the landscape on less convex, north-east facing slopes vs Group B (27,000 sites) on mid-slope to uplands and plateaus.

3.5 Discussion

Results do not follow previous vegetation classification work in the forest, which maps least 11 site-vegetation types within the study area (Figure 3.7 B, Appendix C5) (Havel 1975a; 1975b). Unexpectedly, non-hierarchical clustering with monte-carlo simulation split the dataset into two major groups. Further classification to a 30-group split showed indistinct types, with dominant species overlapping between most groups and uncommon species classified to both core and outlying types (Appendix C4). At the two-group level, group A is clearly associated with riparian and moisture-related taxa found on the south-facing slopes of valleys. Conversely, group B contains the majority of all species and sites, including the dominant trees and shrubs of the forest (Figure 3.4, Figure 3.6). Aspect (northness) best predicted two-group differences, and species richness and diversity were lower in the moisture-associated group. The difficulty in defining vegetation types based on floristics alone is likely due to the heterogenous nature of vegetation pattern in the NJF and lack of distinct environmental gradients at the plot scale (Havel 1975a; 1975b). Given these findings, it is recommended that the SVTs be used as the basis for further modelling work.

3.5.1 *Vegetation pattern and mapping in the NJF*

The different results found for the ALOC and Havel classifications are in-part due to differences in methodology. Havel's typology was originally designed to be used by foresters to identify sites of high productivity, only later being adapted to ecological mapping. The types are based on indicator species and detailed environmental measurements taken in the field, including geological and soil descriptions and measurements (pH, N, P, K) (Havel 1975a). Conversely, the ALOC classification is based entirely on floristics and environmental variables only used to help describe patterns post hoc. Havel's survey was also stratified, with sample sites selected from government monitoring locations that targeted different habitat types (Figure 3.1). The Alcoa dataset however, is based on a systematic survey that follows the drilling grid used for ore-exploration. Finally, Havel's data is based on the full enumeration of 20 m² plots, versus the rapid-fire sampling of Alcoa that focused largely on perennial shrub (5 m radius plots) and tree species (20 m radius).

Vegetation mapping at larger scales across the NJF provides a context for results (Appendix C5). Systems include vegetation complexes (Mattiske and Havel 1998), forest ecosystem units (Commonwealth of Australia 1998) and the National Vegetation Information System (NVIS) (NVIS Technical Working Group 2017) (Appendix C5). These systems rely on environmental factors to help define units, with vegetation complexes highly correlated with landform and NVIS remotely sensed data and structural attributes (Havel 1975a; 1975b; Heddle *et al.* 1980). The forest ecosystem units match most closely the results found in this work, separating stream zone communities from the rest of the flora (e.g. Jarrah – Sandy Basins and Shrub, herb and sedge lands, Appendix C5). As the basis for defining the conservation reserve system, this amalgamates considerable heterogeneity found outside of stream zones into one unit. Given the difficulties in classifying this system, alternative ways to describe and understand patterns are important for conservation management (e.g. Di Virgilio *et al.* 2012).

3.5.2 *Explaining heterogeneity*

Why is species turn-over high and consequently vegetation pattern difficult to describe in the NJF? Climate and topography have been identified as determinants of broad patterns in the NJF (Table 3.3) however, these do not fully explain the turnover of vegetation across the plateaus. One hypothesis is that of old climatically buffered and infertile landscapes (OCBILs), or old, stable landscapes (OSLs) (Hopper 2009; Mucina and Wardell-Johnson 2011). Through a long history of weathering with little rejuvenation by large-scale disturbances (volcanic activity, glaciation), OSLs are defined by complex edaphic patterns, muted topography and soil nutrient poverty. Selection for persistence (stress-tolerance traits) under these conditions, contrasts with the competitive and successional dynamics that operate more strongly in recently glaciated, fertile, topographically distinct landscapes. These factors, together with the flora evolving in-situ (limited dispersal) (Hopper 2009; Tapper *et al.* 2014a) may contribute to the high turnover and indistinct community groupings.

Fire is another important driver. Fire-related traits are well-studied in the SWA flora (Burrows and Wardell-Johnson 2003; Enright *et al.* 2011; Enright *et al.* 2014; Wardell-Johnson *et al.* 2018), but influences on community composition are less understood (Havel 2000; Williams *et al.* 2009). Surprisingly, Burrows *et al.* (2019) found that fine-scale temporal turnover of jarrah forest understorey

vegetation assemblages is independent of the fire regime. Despite this, fire likely influences community composition in complex ways (e.g. from stochastic effects to competitive exclusion and environmental filtering (Liu *et al.* 2018)) and requires further work.

Additional factors that may also influence patterns, but are not well-investigated in the NJF include soil pH, stochastic or random effects and biotic interactions. Laliberté *et al.* (2014) found that variation in plant species richness was almost entirely explained by soil pH, acting as an environmental filter between regional and local species pools. (Although this is likely not a driver in the NJF (Havel 1975a)). In the kwongan heathlands of SWA, multiple environmental factors didn't explain differences between restored sites, with unmeasured stochastic or random effects important (Riviera 2019). Finally, biotic interactions including herbivory and below-ground facilitative relationships are starting to be incorporated into models where traditional variables cannot fully explain observations and is an avenue for further investigation (Araújo and Luoto 2007; Wisz *et al.* 2013; le Roux *et al.* 2014; Pausas and Bond 2019).

More broadly, understanding the drivers of vegetation heterogeneity is a major topic in ecology globally. Pattern and process are linked and variability dependent on the scale of description (Levin 1992; Chave 2013; Stein *et al.* 2014). Hierarchical classifications of vegetation provide a basis for grouping vegetation pattern across scales (Mucina 2019). Macroclimate drives the formation of biomes at large scales; while topography, soils, water availability and disturbance are key at finer scales (Wardell-Johnson *et al.* 2004; Wardell-Johnson *et al.* 2007; Mucina 2019). In many northern hemisphere systems, biotic interactions (particularly competition and more recently, below-ground soil interactions) also have an important impact on local-scale interactions and patterns (De Deyn and Van der Putten 2005; Jablonski 2008; Bardgett and Wardle 2010; Le Roux *et al.* 2013; Wisz *et al.* 2013).

In Mediterranean systems worldwide, plant communities are recognized as being diverse and rich, with high spatial heterogeneity (Lavorel 1999). Climate, topography, fire and human intervention are key drivers of vegetation patterns in these systems (Mazzoleni *et al.* 2004). Thuiller *et al.* (2006) found topographic heterogeneity to be the most powerful explanatory variable for plant species richness in megadiverse South Africa. Recovery from fire is also important across the Mediterranean flora, with a high proportion of resprouting species present (Paula and Pausas 2006). In the oak forests of Portugal, human influences (wildfires, population density, and land accessibility) were the most important driver of forest change (Acácio *et al.* 2017).

Finally, local scale drivers vary between systems and biotic interactions are increasingly seen as important in structuring vegetation pattern. Rainfall and temperature gradients are major drivers in alpine systems (Kikvidze *et al.* 2005), while in the Tropics one mechanism for high local tree diversity is negative density-dependent interactions, shaped by herbivores (Forrister *et al.* 2019). In grasslands, long-term competition experiments have shown how resource limitation sustains diversity (Harpole and Tilman 2007). In boreal North America, geodiversity has been mapped with high correlation to species richness, endemism and rarity (Anderson and Ferree 2010). Process-explicit models that account for physiology, dispersal, demography and biotic interactions are vouched as a way forward (Briscoe *et al.* 2016; Briscoe *et al.* 2019). Overall, key ecological processes may also operate at a different scale to the one being measured (Levin 1992; Chave 2013; Wan *et al.* 2020) and multi-scale studies (temporal and

spatial) are becoming necessary for confidence in climate change modelling predictions (Briscoe *et al.* 2019).

3.5.3 Study limitations

The difficulties in defining groups at the community scale are likely due to natural vegetation patterns and a lack of environmental gradients, exacerbated by the survey and analysis method used in this study. Previous mapping work in the NJF describes the vegetation as a multidimensional continuum with semi-continuous patterns (Havel 1975a; 1975b). Despite surveying across larger environmental gradients than this study, collecting site-based attributes (soil pH, N, K), and using semi-supervised classification with environmental data; community groups were still not clear. Vegetation patterns occur along a gradient from continuous to discrete (Whittaker 1967; Austin 2013). Therefore, regression and other forms of gradient analysis (rather than classification) may be more useful for describing vegetation in this system (Ferrier *et al.* 2007; Austin 2013).

The use of alternative clustering methods may also improve the detection of community groups. Non-hierarchical clustering has advantages over hierarchical methods, as it allows objects to change group membership until an overall optima is reached (Belbin 1990; Belbin and McDonald 1993; Gülagiz and Sahin 2017). However, cluster number must be pre-set and the starting configuration affects final results (Gülagiz and Sahin 2017). The suite of tools available for cluster evaluation is also less well-developed than for hierarchical approaches. Other algorithms that may have been more effective include neural networks, fuzzy clustering, random forest, boosted regression trees, self-organising maps and generalised dissimilarity modelling (Foody 1999; Breiman 2001; Ferrier *et al.* 2007; Crawley 2012; Wisser and De Cáceres 2013) (see also C2). Boosted regression trees have been shown to have higher predictive performance than other methods (Elith *et al.* 2008), while fuzzy/noise clustering provides a probabilistic approach in data that allows and describes the overlap between groups (Kaufman and Rousseeuw 2008; Wisser and De Cáceres 2013). Additional remotely sensed data (e.g. Sentinel-2, Landsat, the soil and landscape grid of Australia) would also probably improve predictions (Xie *et al.* 2008; Frampton *et al.* 2013; Randin *et al.* 2020).

Finally, the evaluation statistics used in the study are based on and biased to Bray Curtis (Table 3.3). There are now available more advanced methods for cluster assessment including indicator species analysis (Indval, Óptimclass), which measures species-to-cluster fidelity, and assessments of cluster stability (De Cáceres and Legendre 2009; De Cáceres *et al.* 2010; Tichý *et al.* 2010; Lubomír *et al.* 2011; Lötter *et al.* 2013; Tichý *et al.* 2014).

3.6 Conclusion

The purpose of this chapter was to create an updated typology for the study area, as a basis for further modelling work. However, the classification of vegetation units for the study area did not work as expected. While an appropriate classification algorithm and analysis process was used for the dataset, distinct units were only discernible at the landscape scale. Comparison to current mapping units (SVTs) was also not successful. Difficulties reflect the high species turn-over found in the NJF and lack of distinct boundaries in vegetation change. In the words of Havel “*The feasibility of using plant indicators*

to define sites in the northern jarrah forest has been demonstrated. The task has not been without problems, in that the vegetation and associate environmental factors form a complex continuum which does not lend itself to easy manipulation.” (pg. 77, Havel 1975a).

Given the difficulties in creating an updated classification for the study area, and extensive field-mapping data already available, SVTs will be used for further modelling work. SVTs provide the opportunity to link expert, ground-based field mapping with remotely sensed data, and a basis for understanding how riparian plant communities may respond to climate change.

Chapter 4 Quantifying hydrologic refugia in a Mediterranean-climate ecosystem

4.1 Abstract

Hydrologic refugia are important areas for biodiversity conservation under climate change, providing a safe-haven for moisture-dependant taxa in drying landscapes. Particular sensitive are Mediterranean-climate ecosystems, which house 20% of vascular plant species in 2% of the worlds area. Mapping hydrologic refugia requires quantification of the relationship between reduced water availability and biotic persistence. This is challenging due to the complexity of plant-climate-hydrological interactions and data-availability. The northern jarrah forest of south-western Australia has undergone strong climatic and hydrological changes since the 1970's. Rainfall has decreased 15-20%, causing reductions in surface run-off (40-100%) and shifts from perennial to ephemeral streams. Using medium-resolution vegetation data (31,000 plots, 432 km², 120 m² grid) and widely available topographic variables (30 m²), we automated community types to model their distribution under projected drying. Multinomial logistic regression modelling and accuracy assessment were carried out in R and ArcGIS. To circumvent mismatches in the resolution of available climate (1 km) and topographic (30 m) data, we used a non-parametric streamflow elasticity estimator modulated by rainfall to detect areas of refugia to 2090. Community types were classified with weak accuracy (0.44, kappa = 0.13, p = <0.001) however, relative results show meaningful change. Overall trends met expectations, with upland types expanding and wetter types contracting. However, type P showed a slight correlation with moisture, which was not expected due to its correlation with *Allocasuarina fraseriana* and *Banksia grandis*. Changes in extent varied from 99% contraction in a riparian type found on sandy loams ('C', 160-380 m elevation, key indicator species: *Eucalyptus patens* and *E. megacarpa*), to 300% expansion in the driest type, found along ridge-tops and plateaus ('G', 210-480 m elevation, key species: *Allocasuarina fraseriana* and *E. marginata*). Patterns of expansion and contraction were combined to indicate areas of potential hydrologic refugia, and were mapped for the study area. Perennial stream-zones provide the basis of hydrologic refugia, with vulnerable floristic communities contracting tightly into remaining waterways. Our work quantifies the importance of riparian areas as habitat for at-risk plant communities in the region and provides a data-driven basis for extensions to the current reserve network under climate change. Streamflow-elasticity estimations are available for over 500 catchments worldwide, enabling medium resolution hydrologic refugia modelling and mapping in landscapes where climate data only exists at coarser resolutions.

4.2 Introduction

Refugia are important for biodiversity conservation under anthropogenic climate change, but quantification and application to planning has been slow (Keppel *et al.* 2015). Defined as maintaining favourable habitat conditions under change, refugia have enabled species persistence over evolutionary time scales (Stewart *et al.* 2010; Keppel *et al.* 2012). This has prompted research into the processes that form and maintain refugia, and methods to identify their occurrence (Keppel *et al.* 2015). To date, refugia have been quantified using paleological data (Rull 2009; Tzedakis *et al.* 2013), genetic data (Tapper *et al.*

2014a), remote-sensing (Klein *et al.* 2009; Mackey *et al.* 2012; Reside *et al.* 2013) and species distribution modelling (Forester *et al.* 2013; Keppel *et al.* 2015; Leipold *et al.* 2017). While much research has focused on temperature and fire refugia, identifying hydrologic refugia is important for adaptive conservation planning in drying systems (Groves *et al.* 2012; McLaughlin *et al.* 2017).

Mediterranean-climate ecosystems (MCE's) are high priorities for mapping hydrologic refugia as they are generally drying and are one of the biomes predicted to experience greatest biodiversity change over the next 80 years (Sala *et al.* 2000). They house 20% of known vascular plant species in 2% of the world's land area but protection is low, with 5% of MCE land-area protected compared to an average of 12% in other biomes (Hoekstra *et al.* 2005). Of the world's five MCE's, Australian MCE's (southern and south-western Australia (SWA)) are a global priority for conservation investment (Klausmeyer and Shaw 2009). This is due to high diversity and endemism (71 plant species per 1000 km²) and high habitat conversion (housing, industry, agriculture) (Cox and Underwood 2011). Strong drying has already occurred, with a 12-15% reduction in rainfall and 40-100% decline streamflow recorded from pre-1975 levels in SWA (Figure 1.2) (Petroni *et al.* 2010). This shift, with a further 6% drop in rainfall, 49-77% contraction in the Australian MCE and only 10% of current protected areas predicted to remain stable by 2100 (Klausmeyer and Shaw 2009; CSIRO and Bureau of Meteorology 2015) underline the urgency of identifying and protecting areas that will retain MCE and mesic conditions.

Systematic conservation planning (SCP) is the basis of conservation-prioritisation and the selection of protected area networks for many local to international government and non-government organisations. Twelve principles guide SCP, including comprehensiveness (spatial objective), adequacy (likelihood of species persistence) and representativeness (the spatial solution) (CAR) (Kukkala and Moilanen 2013). CAR forms the basis of Australia's strategy for the national reserve system (2009-2030) and the SWA reserve system (2014-2023) (Commonwealth of Australia 2010; Conservation Commission of Western Australia 2013). However, the SCP/CAR framework was developed prior to widespread recognition of anthropogenic climate change. Further, most applications of SCP do not include explicit criteria for climate change resilience or adaptation (Groves *et al.* 2012). Hydrologic refugia mapping is one way to incorporate the impacts of climate change into SCP's current reserve-networks and adaptive management (Morelli *et al.* 2016).

Hydrologic refugia are sites of elevated water availability in a landscape, providing a mesic microenvironment that persists in the face of drying trends (McLaughlin *et al.* 2017). The capacity of hydrologic refugia to protect species will depend on the physical characteristics of the site and their continuity under further drying. An organism's physiological tolerances and biotic interactions must also continue to be supported (e.g. pollination, dispersal). Additionally, refugia may be stable (de-coupled from regional drying trends), or relative and/or transient (track regional changes in climate, disappearing as certain thresholds are crossed) (Keppel *et al.* 2015; McLaughlin *et al.* 2017). Physical characteristics include: topographic heterogeneity (drainage, channels, slope variation, fractures/faults, flood plains), shallow/perched aquifers due to geological and edaphic variation (aquicludes, aquitards, clay/caliche lenses) (McLaughlin *et al.* 2017), deep soils or regolith acting as water-storage (Dawson and Pate 1996), and seeps and springs (McLaughlin *et al.* 2017). In SWA, outcropping of springs, the fringes of granite

outcrops, and the edges of creeks and streams (Wardell-Johnson and Roberts 1993; Wardell-Johnson and Horwitz 2000) have been identified as potential hydrologic refugia.

Potential hydrologic refugia can be identified using remotely sensed topographic data and extant vegetation communities (Davis *et al.* 2013; McLaughlin *et al.* 2017; Cartwright *et al.* 2020). In particular, riparian vegetation often forms distinct communities separate from surrounding terrestrial vegetation and provides a clear signal between xeric and mesic conditions (Naiman and Decamps 1997; Richardson *et al.* 2007). Quantification of the hydrologic niche for riparian vegetation types using community modelling enables prediction of change based on future climate scenarios. Distinctions along moisture gradients can also be used to partition between stable, relative and transient elements of a refugium (McLaughlin *et al.* 2017). While species distribution modelling is well-established for predicting climate change responses, community modelling provides integrated management units and synthesizes complex data (Ferrier and Guisan 2006).

Areas of potential hydrologic refugia were modelled for current (2006) and future (2090) climate scenarios using widely available topographic data and a large ground-based vegetation dataset in the northern jarrah forests (NJF) of SWA. Vegetation-environmental relationships were quantified, then patterns of change modelled for predicted drying and warming trends. The overall aim was to determine areas of hydrologic refugia, in order to enable adaptive conservation management under drying.

The four specific aims and associated expectations are:

1. To model the current distribution of vegetation types using medium resolution (30 m) terrain variables. Automation is expected to reveal the underlying drivers of differences between SVTs – including the topographic habitat preferences for riparian vs upland vegetation. Riparian types are expected to occur at lower elevations on shady slopes.
2. To predict and describe how climate drying and warming will impact the distribution of modelled SVTs. Riparian communities are predicted to contract to areas of higher moisture, e.g., larger streams, rivers and wet flats, while upland types will expand into drying areas of the landscape.
3. To define an index for, and map the location of, potential hydrologic refugia in the forest. Refugia are expected to be an emergent property of the collective changes in riparian SVTs distributions, and occur in the wettest conditions represented by these types, e.g., larger streams, rivers and wet flats.
4. To assess the proportion of potential hydrologic refugia that will be protected by the current conservation reserve system. As current reserves are largely based on percent-area measurements, it is expected that a significant amount of potential refugia will not fall in the reserve system.

4.3 Methods

4.3.1 Study site

The northern jarrah forest (NJF) has a Mediterranean climate, with an average annual rainfall of 600-1300 mm that falls predominantly in winter (Havel 1975b; Gentili 1989). A dry-sclerophyll forest, the NJF is significantly globally, occurring in the SWA biodiversity hotspot (Myers *et al.* 2000; Wardell-Johnson *et al.* 2015). Australian dry-sclerophyll forests are adapted to dry conditions, surviving summer drought due to deep regolith providing sufficient water storage to sustain trees (Tozer *et al.* 2017). The vegetation is characterised by subdued topography, an east-west rainfall gradient and regular fire incidence. The forest also has a history of logging and mining (1834-current), that has significantly impacted structure (Churchward and McArthur 1980; Wardell-Johnson *et al.* 2015). Floristically, the forest overstorey is dominated by two tree species (*Eucalyptus marginata*, (jarrah) and *Corymbia calophylla*, (marri)), but has a complex and diverse understorey (3096 taxa) (Dell *et al.* 1989; Western Australian Herbarium 1998-). Understorey vegetation has been described as a multidimensional continuum with semi-discrete community types (Havel 1975a).

The climatic and hydrological regimes of the forest are shifting; rainfall has decreased 12-15% and temperate increased 0.4% since the mid-1970's (Bates *et al.* 2008). This has led to decreased surface runoff (20-100%), falling water-tables and shifts from perennial to ephemeral streams (Petroni *et al.* 2010; Hughes *et al.* 2012). High summer evapotranspiration offsets winter rainfall, which is predicted to further decrease under climate change.

4.3.2 Floristic data

To meet statutory requirements, Alcoa of Australia Ltd (Alcoa, a bauxite mining company) has been collecting floristic data in the NJF for over 25 years (1991-ongoing). Vegetation mapping now spans 432 km² with 31,000 plots on a 120 m² grid that follows ore exploration (Figure 4.1). Communities are mapped using 'site vegetation types' (SVTs), a typology developed in the 1970's for forestry to identify NJF sites of high productivity (Havel 1975a, 1975b, Figure 4.2, Table 4.1, Appendix D1). Each type is defined by a set of indicator species and site characteristics, including topographic position, physical and chemical soil variables, soil texture, pH and depth to water table (Appendix D1). Sixteen SVTs have been mapped in the study area, with five excluded from this analysis due to small sample size (> 40 records) (Figure 4.1).

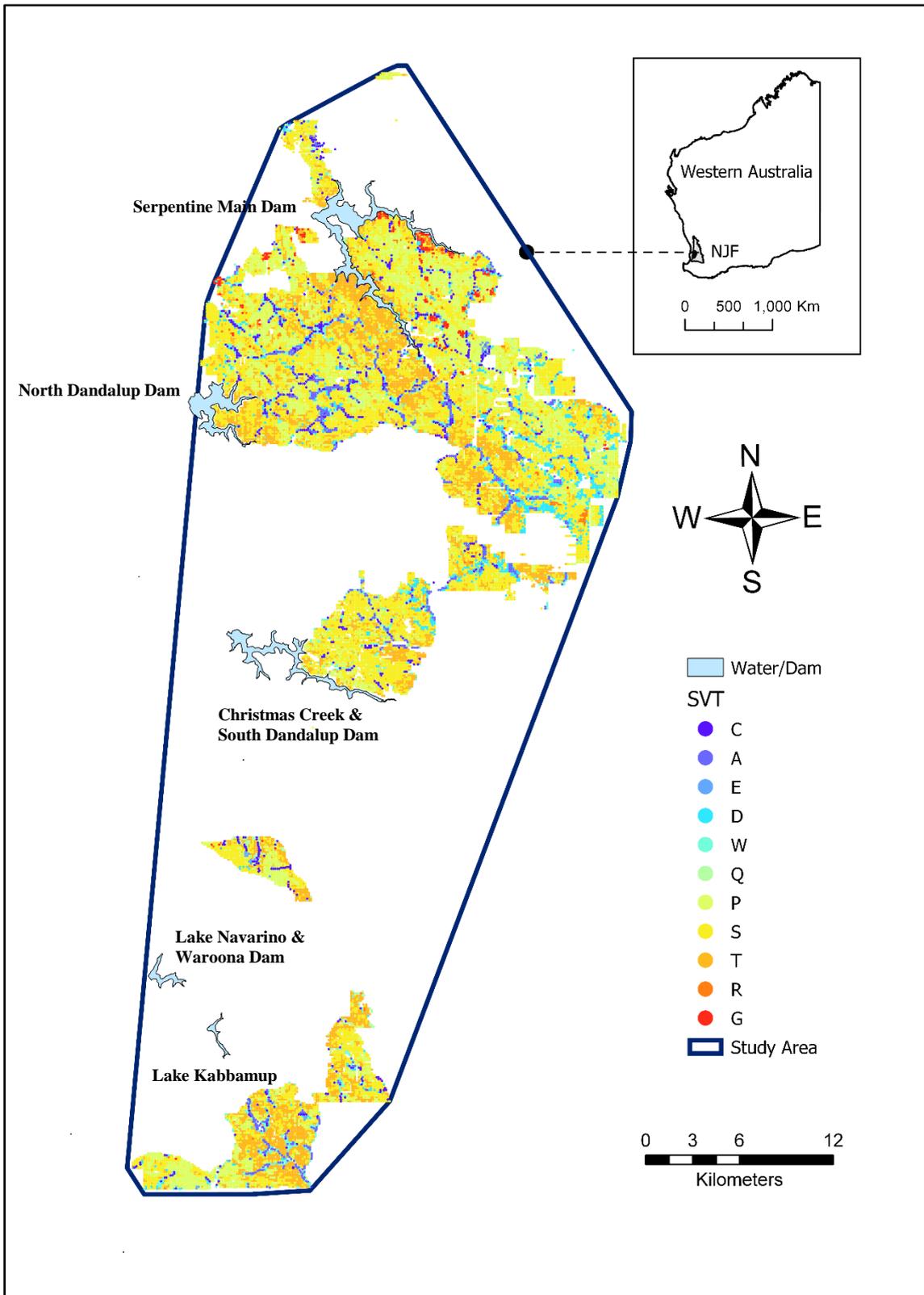


Figure 4.1 Map of study area in the northern jarrah forest (NJF) of south-western Australia showing Alcoa of Australia’s site vegetation type (SVT) mapping , which follows a 120 m drilling grid used for ore exploration. Many of the larger water bodies in the forest are dams, sitting higher in the landscape than natural lakes. Types are ranked by wetness (C – purple, wettest; G – red, driest).



Figure 4.2 Site vegetation types were originally described by Havel (1975), to identify sites of high productivity for forestry and later adapted to ecological management. (a) Type ‘G’, exposed and shallow soils on granite outcrops; (b) the dominant type in the forest ‘S’, includes *Eucalyptus marginata* with *Banksia grandis*; (c) & (d) the stream-zones dominated by *E. rudis* most closely match type ‘C’ however, Havel’s original ‘C’, (e) was usually dominated by *E. megacarpa*; (f) type ‘A’, *Melaleuca preissiana* and *B. littoralis* in a shallow depression; and (g) type ‘P’, dominated by *Allocasuarina fraseriana* (Sheok) on lateritic or heavy gravels with sand. Photos: a-d, f-g, S. Luxton; e, J. Havel 1975a).

Table 4.1 Metadata for the eleven site vegetation types used in this study (riparian (A, C, D, E, W), upland (Q, R, S, T), granite outcrops (G) and casuarina (P)). All variables are taken from Havel (1975a), except for elevation (meters above sea level) and area (km²) which were calculated for this study (see Appendix D1 for full details). Depth WT = depth to water table.

| SVT | Key Site Characteristics | Tree Height (m) | Slope | Elevation (masl) | Gravel % | Depth WT (cm) | pH | P (ppm) | Area (km ²) |
|-----|--|-----------------|-------|------------------|----------|--------------------|-----|---------|-------------------------|
| A | Very wet, leached acid sands underlain by impermeable horizon (often clay or organic-iron hardpan at depth). Waterlogged in winter. | 16-23 | 0-3 | 195-360 | nil | 0-15 | 5.4 | 15 | 10 |
| C | Moist to wet sandy loams - along creeks and swamp margins. | 20-27 | 1-2 | 160-385 | 0-78 | 10-70 | 5.6 | 46 | 12 |
| D | Sandy loams over sandy clay or secondary lateritic ironstone, seasonally waterlogged. Lower slopes and valley floors on eastern range of forest. | 16-33 | 1-3 | 200-420 | 0-26 | 2-90 | 6 | 24 | 17 |
| E | Sands/loamy sands with lateritic gravel. Lower slopes and depressions. Moist to wet in winter, dry in summer. Broad transitional segment between swamps and gravelly slopes. | 14-30 | 1-5 | 240-385 | 0-65 | 27-90+ | 5.8 | 17 | 1 |
| G | Granite outcrops. | NA | NA | 200-520 | NA | NA | NA | NA | 3 |
| P | Lateritic gravel with sand or loamy sand matrix, or sand with heavy gravel. Mid-lower slope, medium to high rainfall zone. | 24-35 | 0-12 | 165-320 | 15-73 | 90+ / not detected | 6.1 | 35 | 87 |
| Q | High fertility and moisture. Dark brown sandy or silty loam over red-brown clay loam. Slopes of major valleys in western high rainfall zone. | 24-38 | 2-6 | 200-440 | 0-44 | 120+ | 6.3 | 114 | 1.5 |
| R | Sandy loam/clay with lateritic gravel. Lower and middle valley slopes, frequently in proximity to granite outcrops. | 16-30 | 1-9 | 170-520 | 32-66 | 90 | 5.9 | 80 | 3 |
| S | Heavy lateritic gravels with sandy loam matrix. Slopes, ridges and plateaus in medium - high rainfall zone. | 23-35 | 2-9 | 210-485 | 21-76 | 90+ | 6.1 | 25 | 182 |
| T | Orange to brown gravel with sandy loam. Restricted to slopes of strongly dissected high-rainfall western zone. | 29-39 | 2-15 | 190-525 | 3-84 | 90+ | 6.1 | 89 | 90 |
| W | Moist sandy loams. Lower slopes and valley floors, tendency to excessive wetness in winter. | 18-32 | 1-4 | 210-465 | 0-26 | 27-90 | 6 | 36 | 20 |

4.3.3 Remotely sensed topographic data.

NASA's one arc-second (30 m) Shuttle Radar Topography Mission (STRM) hydraulically enforced digital elevation model (DEM) was used to derive a set of layers of topographic variables for the study area (NASA 2011). Twenty-eight variables were generated from the DEM in the System for Automated Geoscientific Analyses and ArcGIS (Appendix D2). Topographic position (a measure of terrain ruggedness) was calculated using methods described in (Cooley 2015). Total solar radiation for 2004 (median year for vegetation data) was calculated using the techniques outlined by (Fu and Rich 2002). The System for Automated Geoscientific Analyses wetness index (SWI) was used to represent flow across the study area. SWI is similar to the 'Topographic Wetness Index' (TWI) (Beven and Kirkby 1979) but is based on a modified catchment area calculation which does not conceptualise flow as a thin film (Boehner *et al.* 2002). Consequently, it predicts a higher potential soil moisture per cell and is more realistic than TWI (Böhner and Selige 2006). Latitude and longitude were re-sampled in R to the same resolution as topographic variables (30 m) (Hijmans and Van Etten 2012). All variables were created using automated tools in the System for Automated Geoscientific Analyses (SAGA 2.1.2) and ArcGIS (10.3).

4.3.4 Streamflow elasticity estimator and rainfall

Sankarasubramanian's (2001) non-parametric precipitation elasticity of streamflow estimator (Ep) was used to relate current and future streamflow and rainfall [1]:

$$[1] \quad Q_t = \bar{Q} + (\varepsilon_p \cdot \bar{Q} \cdot (P_t - \bar{P}))/\bar{P}$$

Where Q_t = change in wetness index (site); \bar{Q} = current wetness index (site); E_p = streamflow-rainfall elasticity estimator; P_t = predicted mean annual rainfall (region); \bar{P} = current mean annual rainfall

(region) (Sankarasubramanian *et al.* 2001). Ep provides a measure of the sensitivity of streamflow to changes in rainfall and has been calculated for 219 catchments in Australia (Chiew 2006) and over 500 catchments worldwide (Chiew *et al.* 2006). SWA has a mean Ep of 3.0 (a 10% decrease in rainfall corresponds to a 30% decrease in streamflow). SWI represents flow accumulation across a landscape and provides streamflow input data at 30 m resolution.

Rainfall actual values for 1975-2006 were based on gridded climate data from Commonwealth Scientific and Industrial Research Organisation’s (CSIRO) Representative Climate Futures Framework (Clarke *et al.* 2011; Whetton *et al.* 2012; Williams *et al.* 2018) (Table 4.2). Projected values were developed using the Australian Governments 'Climate Change in Australia' 'Climate Futures' tool, using representative concentration pathway (RCP) scenario 4.5. This tool compares the results of 60+ models (for SWA) and stratifies predictions (e.g. % decline in annual rainfall) by the number of models that agree with an outcome (e.g. 17/68 models predict a 'much drier' (15% decrease) annual rainfall under RCP 4.5 in 2090 (see <https://www.climatechangeinaustralia.gov.au/en/climate-projections/climate-futures-tool/projections/> and Appendix E11 for full details).

Additional projections were obtained from the CSIRO Sustainable Yield Project (CSIRO 2012) and previous work on streamflow decline in SWA (Petrone *et al.* 2010; Hughes *et al.* 2012) (Appendix E2). Both the Climate Futures tool and Sustainable Yield Project report a 12.5% decrease in rainfall from historic conditions (1986-2005 and 1975-2007 respectively), which was used as the baseline value to model the “current” (2006) distribution of vegetation (Table 4.2, Appendix E2).

Table 4.2 Rainfall calculations used in wetness modelling. Historic (pre-1975), “current” (2006) and projected (2030-2090) rainfall for the northern jarrah forest. AR = Annual Rainfall, CCA = Climate Change in Australia, CSYP = Commonwealth Scientific and Industrial Research Organisation Sustainable Yield Project.

| Year | Period | Min AR (mm) | Max AR (mm) | Mean AR (mm) | Recorded/Projected Change (%), Source |
|----------|----------|-------------|-------------|--------------|---|
| pre-1975 | Historic | 600 | 1300 | 950 | NA, Havel (1975) |
| 2006 | Current | 525 | 1138 | 831 | -12.5% cf historic, CSYP, CCA |
| 2008 | Current | 510 | 1105 | 808 | -15% cf historic, Petrone & Hughes (2010) |
| 2030 | Future | 488 | 1058 | 773 | -7% cf current, CSYP |
| 2030 | Future | 494 | 1069 | 781 | -6% cf current, CCA |
| 2050 | Future | 488 | 1058 | 773 | -7% cf current, CCA |
| 2070 | Future | 483 | 1047 | 765 | -8% cf current, CCA |
| 2090 | Future | 478 | 1035 | 756 | -9% cf current, CCA |

4.3.5 Data preparation

The original vegetation dataset held 44 SVTs, as combinations were made while mapping in areas of unclear typology (e.g. a site could be classed as PW/S). Amalgamated types were reduced to the dominant-type (the first single-letter for each record) using the “LEFT” function in Excel (2013). Sites that had the same reference coordinates (29) or no SVT record (22) were removed from the dataset.

Topographic data were imported into R and histograms for each variable were checked visually for normality and homogeneity of variance in responses between groups. Variables were then formally tested for normality using the Shapiro Wilks test (Appendix D3). Transformations followed (McDonald 2008). Multi-collinearity was detected and removed using variable inflation factor (VIF) analysis (package ‘fmsb’) (Zuur *et al.* 2010; Nakazawa 2014) and correlation plots in R (package ‘corrplot’) (Wei *et al.* 2017) (Appendix D4). Variables with a VIF over four, or > 0.5 correlation, were removed from the analysis, based on which would be less environmentally meaningful, relative to the goals of the study (e.g. the topographic roughness index and topographic wetness index were negatively correlated (-0.8), but wetness was of more ecological interest). Sixteen prospective variables were removed, with 14 remaining for the model-fitting process.

4.3.6 Modelling & accuracy assessment

Multinomial logistic regression (MLR) (Agresti 2003; Agresti and Kateri 2011) was used to fit mutually exclusive community distribution models against the 14 topographic variables. MLR is a commonly used algorithm for non-parametric datasets, as it does not assume normality, linearity or homoscedasticity (discriminant function analysis is a more powerful analysis when these assumptions can be met) (Starkweather and Moske 2011). MLR models categorical outcomes, where the log odds of the outcomes is modelled as a linear combination of the predictor variables. The sum of the probability for all outcomes occurring is one. MLR is appropriate in studies where parameterisation of predictor-response relationships is required (Starkweather and Moske 2011). Conversely, machine learning algorithms are often used when predictive power is required but can be a ‘black box’ and difficult to interpret. Couronne *et al.* 2018 found that the random forest algorithm performed significantly better than linear regression, but the gain was marginal (an increase in accuracy of 0.029%) (Breiman 2001; Couronné *et al.* 2018).

The data were split into a training or testing set (80/20) using the package ‘dplyr’ (Wickham *et al.* 2017) and MLR fitted to train data using the package ‘nnet’ (Ripley *et al.* 2016). Nnet is designed to fit artificial neural networks however, no hidden layer was specified in this study. Model building steps followed (Agresti 2003). The Wald statistic (package ‘stargazer’) was used to determine which explanatory variables were significant (Hlavac 2015). Variables that had insignificant coefficients for more than three SVTs were removed from the analysis. A manual backward selection process was used to select final model variables. At each step a variable was removed, and the model tested against the model with all variables using the likelihood ratio test (package ‘lmtest’) (Hothorn *et al.* 2019) (Appendix D3). Non-significant variables were removed ($p < 0.05$) and the process repeated until only significant variables remained. A manual forward-selection process was then used to re-test the significance of each variable removed and several re-added to the equation. Once the main explanatory variables had been selected, possible interactions were tested using the likelihood ratio test (Appendix D3). Three variations of the model were then compared using the Wald statistic, residuals and Akaike’s Information Criterion (AIC) and a final model selected.

The model was run using test data in R (package ‘raster’) (Hijmans *et al.* 2013). A confusion matrix comparing actual vs predicted data points (package ‘caret’ and ‘e1071’) (Kuhn *et al.* 2017; Meyer *et al.* 2017) produced model classification accuracy results including: sensitivity, specificity, detection

prevalence, kappa and balanced accuracy criteria. The kappa coefficient measures agreement between categorical items and is considered more robust than pure percent measures as it incorporates agreement that may occur by chance. Scores are rated - poor: < 0.20, fair: 0.21-0.40, moderate: 0.41-0.60, good: 0.61-0.80, very good 0.81-1.00 (Cohen 1960; Altman 1991).

4.3.7 *Definition and prediction of refugia*

Model coefficients for SWI (wetness) were ranked and used as the basis for quantifying SVT-water relationships. The current (2006) distribution of each type was calculated. Model coefficients and summary statistics (mean of each environmental variable, per SVT) were extracted and the relationship between environmental variables and SVTs visualised using boxplots and principle component analysis (PCA). SVTs that had a strong positive correlation with SWI (mean SWI = >2) were classed as “wet” types. Types that had a negative correlation with wetness were classed as “dry” types. Types with that sat in the medium range for SWI (0-2) were generally not referred to as either and did not form a core component of further analysis.

Using the streamflow-rainfall elasticity estimator and predicted rainfall values, wetness was re-calculated for 2090 using the non-parametric streamflow-rainfall elasticity estimator and predicted rainfall values (Table 2) in R (Appendix D5). The MLR model was used to predict SVT occurrence in 2006 and 2090 for the study area and potential refugia mapped (R, ArcGIS Pro v 2.0). Percent change in distributions was calculated for each type from 2006 to 2090, using the number of presence-pixels multiplied by pixel area (26.5 m²) (Appendix D3). A confusion matrix was produced. Stable and transient hydrologic refugia (McLaughlin *et al.* 2017) were categorised based on 2006-2090 change classes. Stable refugia were riparian types (A, C, E, D, W) that did not change, while area's that shifted to a new riparian type were classed as transient refugia.

4.3.8 *Conservation assessment*

To determine the proportion of riparian community types (and potential hydrologic refugia) in conservation areas, crown reserve system data was used (Harris 2018). Reserves in the study area are classified IUCN ‘2’ (Category II: National Park, Dudley 2008). Under the WA Land Administration Act (1997) the entire study area is a ‘Class A’ reserve (protect areas of high conservation or community value, requiring approval of Parliament to amend or cancel the reservation) however, there are varying classes of protection of use within this. Reserves were clipped to study area and proportions calculated (number of grid cells x grid cell area).

4.4 **Results**

Eight topographic variables were significant in the final model and show patterns across the types (Figure 4.3). Elevation (channel base, valley depth), convexity and wetness show distinct differences between the riparian types (A, C, D, E, W) and upland groups (G, P, Q, R, S and T) (Figure 4.3). The casuarina type (P), and upland groups (S and T) sit highest in the landscape (low valley depth), while Q (upland) is situated south-westerly within the study area and granite outcrops (G) more widely distributed in the north. Type D (a riparian type) is the most easterly distributed of the types. Granite outcrops (G) and Q

also both show a higher slope-ratio (LS-Factor) relative to all other types, indicating a higher soil erosion quotient. Finally, longitudinal curvature (curvature of the slope) showed little difference between the groups (Figure 4.3). Relative to wetness, riparian types quantified as having high moisture requirements (A, C, D and W) occurred on east facing slopes (high plan convexity as related to aspect) and lower in waterways (distance to valley depth) (Figure 4.4). Conversely, high longitudinal curvature (or 'profile' convexity, which measures slope) is associated with an increased probability of upland types (Figure 4.4).

While these patterns were found, there is also considerable overlap in environmental space, with only the casuarina type (P) and riparian type A becoming distinguishable at maximum environmental values (Figure 4.4).

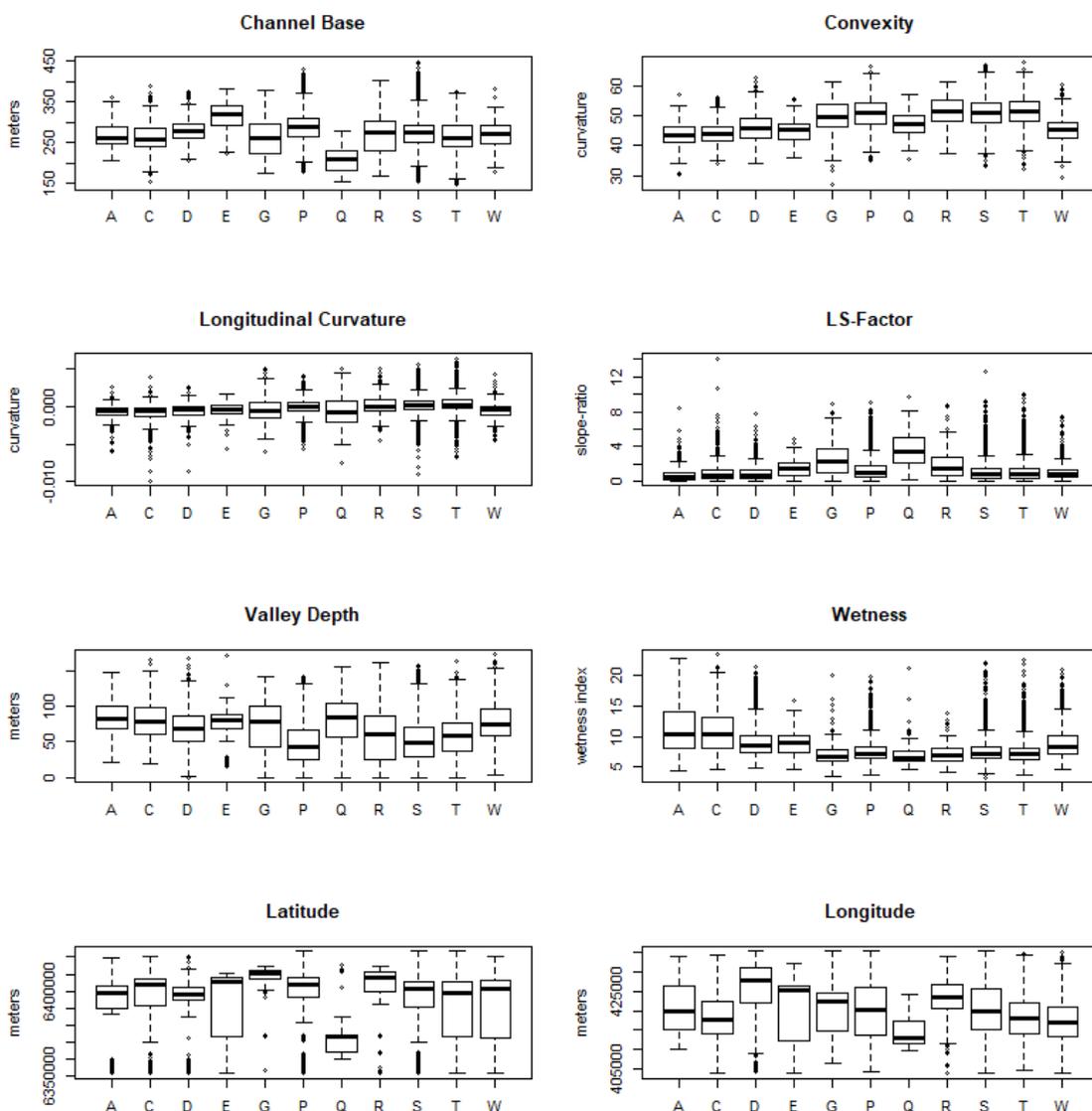


Figure 4.3 The topographic signal for each site vegetation types varies, with both clear (e.g. channel base, wetness) and subtle (longitudinal curvature) differences between the groups (riparian (A, C, D, E, W), upland (Q, R, S, T), granite outcrops (G) and casuarina (P)).

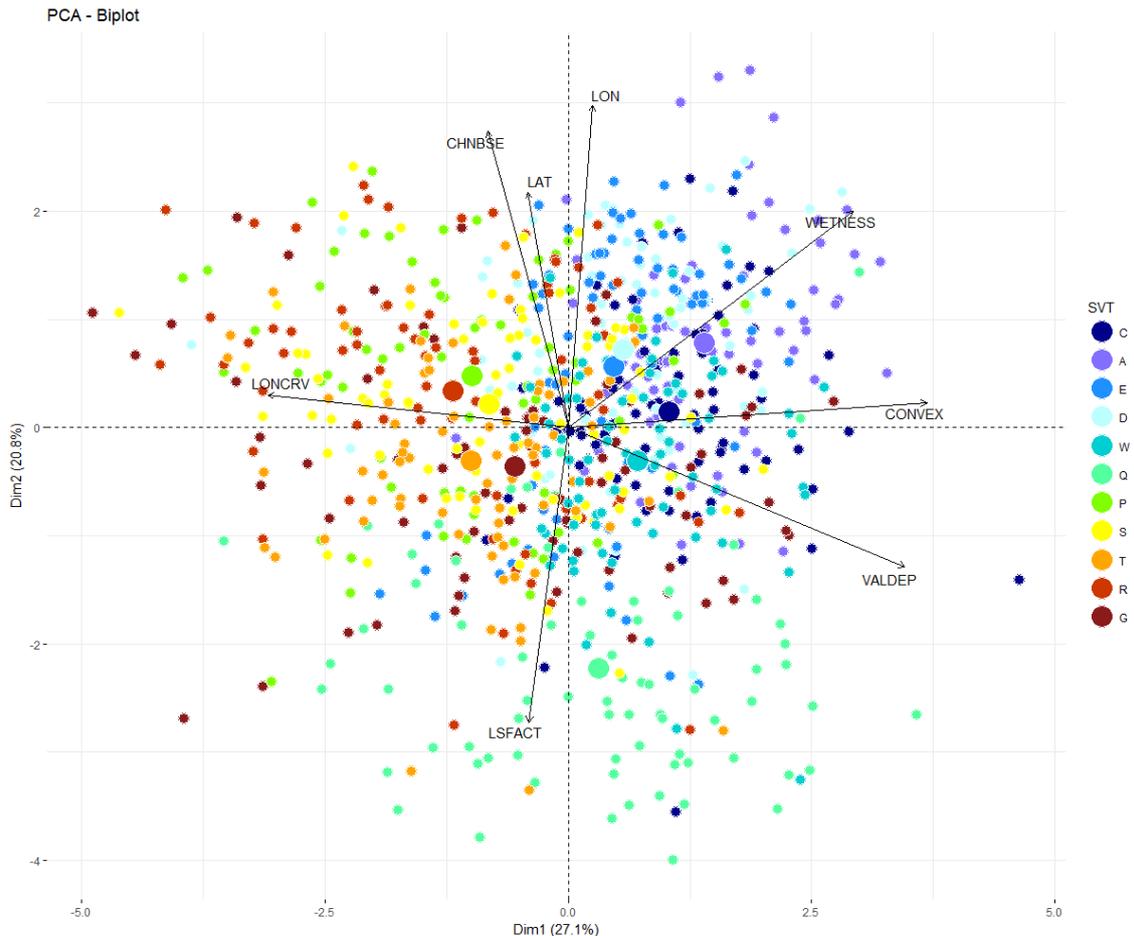


Figure 4.4 Principal components (PCA) biplot of final model environmental variables showing the wide occurrence of several types across gradients (e.g. G, S) and separation at the edges of tolerances (e.g. A, Q). Plot is based on a sub-sample of 825 sites (75 per site vegetation type). CHNBSE = distance to channel base network, LAT = latitude, LON = longitude, WETNESS = SAGA wetness index, CONVEX = plan convexity (aspect), VALDEP = distance to valley depth, LSFAC = LS-factor, LONCRV = longitudinal/profile curvature (slope). Larger circles are means. Types are riparian (A, C, D, E, W), upland (Q, R, S, T), granite outcrops (G) and casuarina (P)

4.4.1 Topographic wetness index

To identify the types most vulnerable to drying, SVTs were ranked using the wetness coefficients from modelling (Table 4.3). Riparian types C and A had the highest positive correlation with wetness (C, coef = 8.14, SE = 0.28; A, coef = 6.5, SE = 0.27; Table 4.3). Conversely, the upland types T, R and granite outcrops (G) were the most negatively correlated (Table 4.3). Type upland type Q and casuarina (P) showed a slight correlation with moisture (1.7 and 0.2 respectively). The distribution of ‘wet’/riparian SVTs corresponds to the major and minor creek lines of the forest (Figure 4.1), with topographic factors interacting to influence wetness (Figure 4.4). Longitudinal curvature and LS-factor negatively impact the probability of riparian types, while distance to valley depth and surface convexity are positively associated (Table 4.3). Geographic position is also important and riparian types are more likely to be

found in the south east of the study area (Figure 4.1, Figure 4.2). Upland type T proves an exception, as it is dominant in the southern study area and expected to expand significantly with drying (Figure 4.1).

Table 4.3 Model coefficients and standard error for each site vegetation type (SVT), ranked by wetness (highest to lowest). Multinomial logistic regression calculates coefficients relative to a reference category (S) for which all variables equal zero. Types are riparian (A, C, D, E, W), upland (Q, R, S, T), granite outcrops (G) and casuarina (P).

| SVT | Environmental variables (coefficient; standard error) | | | | | | | | | | | | | | | | | |
|-----|---|-------|--------------|-------|-----------|------|------------------------|------|-----------|------|--------------|------|---------|------|----------|------|-----------|------|
| | Intercept | | Channel Base | | Convexity | | Longitudinal Curvature | | LS-Factor | | Valley Depth | | Wetness | | Latitude | | Longitude | |
| C | -57.10 | 0.001 | -0.0001 | 0.001 | 2.39 | 0.11 | -0.40 | 0.04 | -0.07 | 0.10 | 0.31 | 0.03 | 8.14 | 0.28 | 1.06 | 0.05 | -0.82 | 0.07 |
| A | 32.61 | 0.001 | -0.0014 | 0.002 | 2.32 | 0.11 | -0.32 | 0.05 | -1.19 | 0.11 | 0.42 | 0.03 | 6.50 | 0.27 | -0.88 | 0.05 | 0.05 | 0.07 |
| E | -8.50 | 0.003 | 0.0307 | 0.004 | 1.67 | 0.32 | 0.11 | 0.15 | 0.85 | 0.26 | 0.29 | 0.08 | 5.91 | 0.03 | -0.19 | 0.14 | -0.21 | 0.21 |
| D | -20.22 | 0.001 | 0.0045 | 0.001 | 1.24 | 0.08 | -0.31 | 0.04 | -0.43 | 0.09 | 0.21 | 0.03 | 3.30 | 0.31 | -0.63 | 0.04 | 1.12 | 0.06 |
| W | 22.49 | 0.001 | 0.0020 | 0.001 | 1.90 | 0.08 | -0.18 | 0.04 | -0.27 | 0.08 | 0.33 | 0.02 | 2.94 | 0.28 | -0.13 | 0.04 | -0.73 | 0.06 |
| Q | 39.53 | 0.004 | -0.0554 | 0.005 | 0.97 | 0.29 | -0.13 | 0.08 | 1.34 | 0.22 | 0.03 | 0.08 | 1.70 | 0.02 | -0.01 | 0.20 | -0.90 | 0.31 |
| P | -76.73 | 0.000 | 0.0176 | 0.001 | -0.17 | 0.04 | -0.18 | 0.02 | 0.40 | 0.04 | 0.03 | 0.01 | 0.20 | 0.22 | 1.27 | 0.02 | -0.24 | 0.03 |
| S | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| T | 205.53 | 0.000 | -0.0163 | 0.001 | -0.04 | 0.04 | 0.14 | 0.02 | -0.23 | 0.04 | 0.03 | 0.01 | -0.86 | 0.22 | -3.11 | 0.02 | -0.06 | 0.03 |
| R | -34.18 | 0.001 | -0.0009 | 0.003 | -0.53 | 0.16 | 0.01 | 0.08 | 1.07 | 0.15 | 0.09 | 0.04 | -1.74 | 0.01 | 0.13 | 0.08 | 0.57 | 0.13 |
| G | -69.28 | 0.002 | 0.0053 | 0.002 | -0.45 | 0.17 | -0.27 | 0.06 | 1.85 | 0.14 | 0.27 | 0.05 | -2.62 | 0.01 | 1.10 | 0.09 | -0.17 | 0.14 |

4.4.2 Model performance

Overall, modelling classified SVTs with medium-low accuracy (accuracy = 0.44, kappa = 0.13, $p < 0.001$) (Table 4.4). The model under-predicted the occurrence of all types (low sensitivity and percent correct) except for the reference group 'S'. Types were regularly misclassified as S, P or T (Table 4.4). Overall accuracy improved when types S, T and P were amalgamated (acc = 0.84) however kappa was still very poor (0.18, $p\text{-val} = 0.13$) and misclassification as S high. Kappa and individual accuracies improved when classes S, T and P were amalgamated and the data subsetted to create even categories (Kappa 0.35, $p\text{-val} < 0.001$) however, overall accuracy was not improved (0.48). The model greatly underestimates the occurrence of riparian types (A, C, D, E, W) and misclassifies them as S.

Table 4.4 Multinomial logistic regression results: (A) model accuracy assessment and (B) confusion matrix (% correct = percent of types correctly predicted as known). Overall accuracy = 0.44, kappa = 0.13, p-value = <0.001.

(A)

| Statistic | A | C | D | E | G | P | Q | R | S | T | W |
|-----------------------------|------|------|------|-----|------|------|-----|------|------|------|------|
| <i>Sensitivity</i> | 0.21 | 0.2 | 0.1 | 0 | 0 | 0.14 | 0 | 0 | 0.88 | 0.15 | 0.08 |
| <i>Specificity</i> | 0.99 | 0.99 | 0.99 | 1 | 1 | 0.97 | 1 | 1 | 0.25 | 0.94 | 0.99 |
| <i>Pos Pred Value</i> | 0.39 | 0.31 | 0.38 | NaN | 0 | 0.56 | 0 | NaN | 0.45 | 0.39 | 0.24 |
| <i>Neg Pred Value</i> | 0.98 | 0.98 | 0.96 | 1 | 0.99 | 0.8 | 1 | 0.99 | 0.74 | 0.81 | 0.96 |
| <i>Prevalence</i> | 0.03 | 0.03 | 0.04 | 0 | 0.01 | 0.22 | 0 | 0.01 | 0.41 | 0.2 | 0.05 |
| <i>Detection Rate</i> | 0.01 | 0.01 | 0 | 0 | 0 | 0.03 | 0 | 0 | 0.36 | 0.03 | 0 |
| <i>Detection Prevalence</i> | 0.02 | 0.02 | 0.01 | 0 | 0 | 0.05 | 0 | 0 | 0.8 | 0.08 | 0.02 |
| <i>Balanced Accuracy</i> | 0.6 | 0.6 | 0.55 | 0.5 | 0.5 | 0.56 | 0.5 | 0.5 | 0.56 | 0.55 | 0.53 |

(B)

| KNOWN (% Correct) | PREDICTED | | | | | | | | | | |
|-------------------|-----------|----|----|---|---|-----|---|---|------|-----|----|
| | A | C | D | E | G | P | Q | R | S | T | W |
| A (24) | 41 | 18 | 8 | 0 | 0 | 1 | 0 | 0 | 88 | 2 | 12 |
| C (24) | 17 | 47 | 5 | 0 | 0 | 5 | 2 | 0 | 105 | 6 | 9 |
| D (13) | 29 | 5 | 34 | 0 | 0 | 7 | 0 | 0 | 172 | 5 | 6 |
| E (0) | 1 | 0 | 0 | 0 | 0 | 10 | 0 | 0 | 172 | 5 | 6 |
| G (0) | 1 | 1 | 0 | 0 | 0 | 9 | 0 | 0 | 35 | 1 | 2 |
| P (15) | 7 | 6 | 10 | 0 | 0 | 186 | 0 | 0 | 972 | 64 | 9 |
| Q (15) | 0 | 1 | 0 | 0 | 0 | 0 | 3 | 0 | 6 | 10 | 0 |
| R (0) | 0 | 0 | 0 | 0 | 0 | 9 | 0 | 0 | 38 | 5 | 0 |
| S (88) | 18 | 15 | 5 | 0 | 1 | 106 | 1 | 0 | 2260 | 149 | 15 |
| T (13) | 9 | 14 | 2 | 0 | 0 | 14 | 1 | 0 | 1066 | 164 | 10 |
| W (7) | 22 | 23 | 4 | 0 | 0 | 4 | 0 | 0 | 198 | 13 | 20 |

4.4.3 Quantifying hydrologic refugia.

The projected 6-9% decrease in rainfall (2006-2090) corresponds to 18-27% decrease in streamflow ($\epsilon_p = 3.0$) and wetness across the landscape (Appendix D5). The model predicts both range contraction in riparian types and expansion for upland types (Figure 4.5, Table 4.5). Riparian types generally occupy a smaller initial area and undergo more significant changes (Table 4.5). Types C and A are predicted to be the most impacted by drying (Table 4.5). Types P and S show slight decline ($P = -14\%$ and $S = -4\%$), while type T is predicted to expand significantly (Table 4.5).

Hydrologic refugia were modelled to 2090 and mapped across the study area (Figure 4.5). Riparian types contract strongly into areas of high moisture around larger water bodies, creeks (Figure 4.5 A-C) and low-lying swamps (Figure 4.5 D) while upland types expand, especially across the south of the study area (Figure 4.5 C, D). SVTs S and T will dominate, superseding up to 90% of current cover in other types (Table 4.5). Patterns of expansion and contraction overlap, decreasing areas of potential refugia (Figure 4.5). Transitions were quantified within and between types. Riparian types shift down the moisture continuum with only 1.4% of C predicted to remain (Table 4.5). Shifts in upland types vary, with 66% of Q remaining while types P, S and T retain more than 85% cover. SVT 'G' is the driest type, but is not

predicted to dominate, likely due to its limited starting extent (3 km², Table 4.1). S and T will supersede the other types by the greatest amount (Table 4.5).

Hydrologic refugia were also separated into ‘stable’ and ‘transient’ refugia. Stable refugia are areas where no change in moisture (or other) habitat conditions are expected to occur under predicted drying, whereas transient refugia occur at low levels of change, but may disappear as drying continues (Figure 4.5). The topographic signal for areas of stable and transient vs non-refugial areas is subtle but clear (Figure 4.3). Riparian types (A, C, D, E, W) generally show a lower distance to channel basin, convexity and higher wetness (Figure 4.3). While stable and transient refugia occur around the edges of the major water bodies, they are not significantly larger or more stable there than in other locations (e.g. Figure 4.5 C vs A and D).

4.4.4 The conservation reserve system and hydrologic refugia

Within the study area, three percent is expected to retain areas of stable hydrologic refugia in 2090 and 6% transient refugia (Figure 4.5). Currently, 10.2 percent of the forest is within an International Union for the Conservation of Nature (IUCN) Class 2 reserve, but only a fraction of stable refugia are captured by these in 2090 (Figure 4.5). Based on model calculations, the area required to be reserved to protect all potential stable hydrologic refugia in 2090 is 40 km², increasing the total protected area for this region by 3% to 13%.

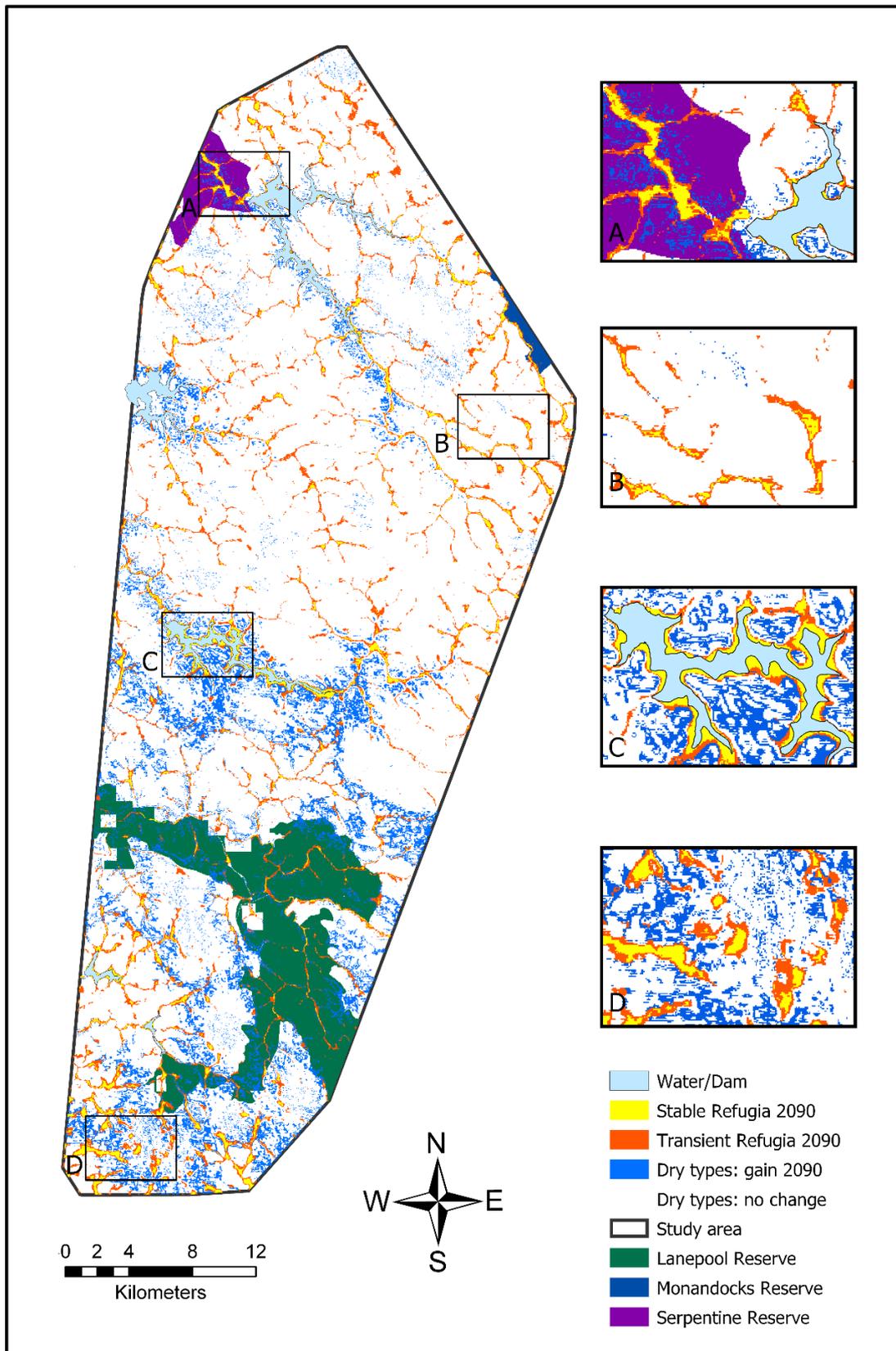


Figure 4.5 Predicted areas of stable and transient hydrologic refugia for 2090 form a network following the rivers and streams of the forest. The current IUCN Class 2 reserve system includes refugia, e.g. (A) in the Serpentine National Park, but large areas are also not captured; B) hydrologic refugia in areas with little dry(upland)-type expansion; C) refugia around the edges of a major water body; and D) matrix of high dry-type expansion interspersed with large patches of refugia, likely to be broad, swampy valleys.

Table 4.5 Confusion matrix for changes in site vegetation type (SVT) occurrence from 2006 – 2090, predicted using the multinomial logistic classifier. Values are area (km²) and (percent change between types). NET AREA CHANGE = total change in area from 2006 to 2090 (km²). NET % CHANGE = total percent change in cover for each SVT from 2006 to 2090; grey = no change. Types are ranked in wetness order (C = wettest, G = driest). Type description: riparian (A, C, D, E, W), upland (Q, R, S, T), granite outcrops (G) and casuarina (P).

| SVT 2006 | SVT 2090 | | | | | | | | | NET AREA CHANGE (km ²) | NET % CHANGE |
|----------|------------|-----------|-------------|-----------|---------------|---------------|-----------|-------------|---------------|------------------------------------|--------------|
| | C | A | D | W | Q | P | S | T | G | | |
| C | 0.2 (1.4%) | 2.3 (14%) | 0.02 (0.1%) | 3.8 (23%) | 0.23 (1.4%) | 0.1 (0.6%) | 7.5 (45%) | 2.4 (14%) | 0.04 (0.24%) | -16 | -99% |
| A | 0 | 4.6 (27%) | 0.74 (4.3%) | 1.9 (11%) | 0.03 (0.2%) | 0.005 (0.03%) | 6.4 (37%) | 3.6 (21%) | 0 | -10 | -60% |
| D | 0 | 0 | 0.36 (9.5%) | 0 | 0 | 0.014 (0.36%) | 3.4 (90%) | 0 | 0 | -2.7 | -70% |
| W | 0 | 0 | 0 | 2 (19%) | 0.003 (0.03%) | 0.025 (0.24%) | 3.6 (33%) | 5.1 (48%) | 0.002 (0.02%) | -3 | -28% |
| Q | 0 | 0 | 0 | 0 | 3.9 (66%) | 0 | 0.06 (1%) | 1.9 (32%) | 0.08 (1.4%) | -1.8 | -30% |
| P | 0 | 0 | 0 | 0 | 0 | 13.2 (85%) | 2.3 (15%) | 0.000 | 0.03 (0.22%) | -2.2 | -14% |
| S | 0 | 0 | 0 | 0 | 0 | 0 | 327 (90%) | 36 (10%) | 0.02 (0.06%) | -14 | -4% |
| T | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 123 (99.9%) | 0.02 (0.02%) | 49 | 40% |
| G | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.12 (100%) | 0.4 | 335% |

4.5 Discussion

Climate drying and warming is expected to cause range contraction into hydrologic refugia (Keppel *et al.* 2012; Keppel *et al.* 2015). Identifying hydrologic refugia requires quantifying the relationship between reduced water availability and plant persistence, which is challenging due to limitations in data-availability and the complex interactions that influence hydrology (McLaughlin *et al.* 2017). The effects of drying were modelled for the distribution of riparian and upland plant communities in a Mediterranean system undergoing significant reductions in rainfall and streamflow (Petroni *et al.* 2010; Hughes *et al.* 2012). Medium resolution vegetation (120 m) and remotely sensed topographic data (30 m) was used, with a wetness indicator modulated by rainfall to capture plant-climate-hydrology relationships and to project future distributions. This approach enabled the location of projected hydrologic refugia to be mapped and an assessment of the underlying the reserve system to be made.

4.5.1 Automation of site vegetation types

The topographic drivers of vegetation pattern show both distinct and subtle differences between the SVTs. Along the moisture continuum, types quantified as having higher moisture requirements (A, C, D and W) are distinct from ‘dry’ types (S, T, R) (Figure 4.3). Additionally, differences in elevation, terrain convexity and geographic location were seen. These distinctions enabled SVTs to be adequately modelled however, the overlap in tolerances and preferences for the groups (Figure 4.4) impacted overall detection accuracy.

The MLR model had difficulties classifying SVTs clearly. ‘S’ dominates predictions, with all other types significantly underestimated (Table 4.4). This has implications for predictions, with the current and future

area occupied by riparian SVTs likely lower than ground-based measurements (Figure 4.4, Table 4.5). Both statistical and ecological factors contribute to this. Statistically, multinomial logistic regression solves multiple binomial equations simultaneously to predict a single outcome (in this case, an SVT for each grid cell). This reduces error within the calculations, but limits the capacity to distinguish between groups and find an optimal solution, compared to running separate binomial equations (Appendix D6, Appendix D7). Binomial models for individual SVTs show greatly improved accuracy results ($\text{acc} = 0.6 - 0.88$, Appendix D8) however, predicted outputs overlap, making it difficult to select a single solution for a spatial location. Additional variables that would likely improve prediction accuracy (but were unavailable at a 30-120 m resolution) include: disturbance history (fire, management and resource use), rainfall and soil data.

Alternative modelling methods should also be investigated. Regression enabled the drivers of vegetation patterns in the forest to be parameterised and is appropriate when an understanding of the relationship between predictors and response variables is required. However, other algorithms like: Neural Networks, Random Forest Trees and Generalised Dissimilarity Modelling (GDM) may have better predictive power (Breiman 2001; Elith *et al.* 2006; Ferrier and Guisan 2006; Ferrier *et al.* 2007; Bourel and Segura 2018). Machine learning algorithms maximize predictive power using unsupervised techniques that search the natural structures in data and build relationships for prediction. However, in 'big data' they rely on 'brute force' computing and are often a black box where it is difficult to interpret drivers. Couronne *et al.* (2018) found that the random forest algorithm performed significantly better than linear regression, but the gain was only marginal (an increase in accuracy of 0.029%). GDM has a wide range of applications, including reserve network selection and incorporates both environmental and species data into predictions of beta-diversity (House *et al.* 2012; Molloy *et al.* 2016; Di Marco *et al.* 2019). The used of GDM with soil data (e.g. the soil and landscape grid of Australia (Grundy *et al.* 2015)) would likely improve predictive performance.

The philosophical basis of classification is to separate and define groups for human knowledge, use and communication (Russell 1948; Sparrow 1991). However, vegetation pattern does not always lend itself well to this task. The lack of clear vegetation classes separated along the topographic variables used in modelling corresponds with previous work in the NJF. While several types are distinct near environmental maximums (e.g. riparian-A and casuarina-P, Figure 4.4), the vegetation and its associated environmental variables generally forms a complex continuum (Havel 1975a). The lack of strong environmental gradients in the NJF includes subdued topography (elevation range = 160-525 m asl, mean = 295, SE 0.22) and (within the study area) relatively low spatial variation in rainfall.

The broad occurrence of some types is also due to the indicator species that form the groups and in general, the difficulty of deriving indicator species in this system (Havel 1975a; 1975b). For example, *E. patens*, or Swan River Blackbutt, is recognised as a stream zone species, but can also grow along the side of valley slopes. *Lepidosperma squamatum*, an indicator for C, is found in calcareous, peaty or lateritic sand, sandy clay, gravel, dunes and swamps, and is also an indicator for types D, E and W (Havel 1975a). Similarly, type P (casuarina) showed a slight correlation with moisture, which was not expected, as several of its indicators are well-known upland tree species (*Allocasuarina fraseriana* and *Banksia grandis*). The upland, expanding types (T, R, G) contain species well-correlated with the dry, rocky

(lateritic) upland soils of the ridges and plateaus, including the forest dominants *E. marginata* and *C. calophylla* and the understory tree *B. grandis*. However, widely-distributed taxa like *Clematis pubescens* (T), which is found on sites that have sandy clay or loam, to limestone, valleys and river banks, and (less-consistently) *E. patens* were also recorded in these types.

The automation of the SVTs using only terrain variable limits applications. Type G (granite outcrops) for example, warrants careful interpretation. Granite outcrops have complex microtopography, with areas of deep soil that collect runoff and retain water immediately next to very shallow soils on the exposed domes of outcropping rock (Figure 4.2 a) (Tapper *et al.* 2014a). They are restricted edaphically to localised patches with a complex vegetation ranging from low, open forests of *Allocasuarina fraseriana* to highly diverse herb lands (Hedde *et al.* 1980; Yates *et al.* 2019). The high expansion forecast for this type (+300%) is partly due to the very hot, exposed conditions that this vegetation can tolerate. However, modelling doesn't account for other factors that may influence plant occurrence between outcrops and the surrounding landscape. These include the dispersal ability of species and the effect of an outcrop's isolation on extinction and colonisation processes (Yates *et al.* 2019). Additionally, competitive exclusion by adult individuals that have the capacity to tolerate drying trends and persist in current locations may delay the recruitment of new species into these areas (McLaughlin *et al.* 2017).

4.5.2 Potential hydrologic refugia in the forest

By defining SVTs as 'riparian ('wet') or upland ('dry') based on their correlation with the topographic wetness index, groups could be amalgamated and an assessment of potential hydrological refugia made (Table 4.3, Table 4.4, Figure 4.5). Overall, riparian communities are predicted to disappear from the reaches of lower-order streams and contract strongly to major tributaries and water bodies and low-lying swamps in broad valleys (Figure 4.5 A-D). These areas of persistent moisture are likely to provide the most stable habitat conditions for riparian taxa and the location of potential hydrological refugia in the forest.

Vegetation mapping at larger scales across the forest provides a basis for the interpretation of refugia results (Figure 4.6, Appendix D9). Vegetation complexes are a more broad-scale unit than SVTs. Based on landform mapping, they delineate major and minor valley and upland units and reflect differences in soil across the plateau (Figure 1.3, Figure 4.6) (Churchward and McArthur 1980). Units support the distribution of hydrologic refugia in valleys, with zones of stable refugia particularly centred in the Swamp complexes ('Swamp and Yarrigal', Figure 4.6). Yarrigal (max) includes large areas of swamps restricted to the upper reaches of rivers, while Yarrigal (min) has less swamps, but still supports large patches of stable refugia (Figure 4.5 D, Figure 4.6). The Murray unit is represented by SVTs C and T. The moderately incised valleys of this area may underly the mixing of both significant contraction by C, the presence of refugia along the banks of the Murray River and encroachment of type T into the unit. Finally, the exclusion of refugia zones in the uplands is represented by the Dwellingup complex, which only contains several fragments of transient refugia over its extent (Figure 4.6). Forest ecosystem units also reflect changes in major vegetation units between upland and stream zone sites. Refugia track the most restricted units, but are also modelled across the general forest type ('Jarrah – North West'), particularly in the north east and south west of the study area (Appendix D9).

The formation or persistence of refugia can also be described using modelled outputs through the estimation of how each SVT will shift in response to drying (Table 4.5). The near complete loss of riparian type C indicates that the wettest conditions in the forest may not be retained, with implications for species like *Hypocalymma angustifolium* (White Myrtle). White Myrtle occurs on flats and swamps, and along watercourses (although it can also be found on outcrops and hillsides) (Western Australian Herbarium 1998-). Monitoring of taxa like White Myrtle provides an indicator – signalling both the location of the most stable areas and potential loss if soil moisture conditions start changing. As types like riparian-A and upland-S are predicted to extend into riparian-C, the encroachment of species represented by these types (e.g. *Mesomelaena tetragona* (Semaphore Sedge – A), or, more likely, *E. marginata* or *C. calophylla* (S), also provides a signal of change.

The complex over-lapping tolerances of species and effects of physiological and other factors provides a challenge to assessing actual change. In SWA, the plasticity of drought-tolerant traits has been measured in several riparian trees (White 2017). Seedlings were found to abscise leaves and re-sprout when conditions improved, suggesting a physiological resilience to drying, and *E. rudis* (Flooded Gum) is moderately drought tolerant once established. However, interactions between geology, climate and wildfire may reduce resilience and facilitate community transitions (Bond and Keeley 2005; Enright *et al.* 2015). Drying will lengthen fire seasons, increase the number of extreme fire weather days annually and select for more flammable species, mediating vegetation change (Enright *et al.* 2015; Stralberg *et al.* 2018). In the NJF, riparian zones include a disproportionately high number of fire sensitive species that are slow to reach reproductive maturity (Williams *et al.* 2009) while dry-tolerant tree species (*C. calophylla* and *E. wandoo*) have been observed recruiting into riparian areas following fire. In an Australian context, climate change is expected to have the most significant consequences for fire regimes and the sclerophyll-dominated vegetation of SWA (Williams *et al.* 2009).

Drought-induced tree mortality will also be an important mechanism in the creation (and loss) of potential hydrologic refugia. Globally, examples of drying-induced tree mortality are well-recorded (Allen *et al.* 2010; Dillon *et al.* 2015; Stephens *et al.* 2018; Stralberg *et al.* 2020a). In the NJF, an extremely hot 2010/11 summer caused crown dieback across 1,350 ha of the forest (Matusick *et al.* 2013). An aerial survey identified correlates, including rocky soils with low water holding capacity and the probability of dieback presence (Brouwers *et al.* 2013). The causes of mortality are linked to carbon starvation, hydraulic failure and insect/pest damage (McDowell *et al.* 2008; Poot and Veneklaas 2013). For several species of Eucalypt, water relation measurements over a gradient of drying showed lower summer leaf water potentials, osmotic potential, and vulnerability to cavitation; and higher stomatal conductance and relative sapflow velocity in drier-adapted taxa (Poot and Veneklaas 2013). *E. marginata* exerts stomatal control when drought-exposed and leaves of lower-rainfall populations have been shown to experience stronger drought stress than those from higher-rainfall populations, especially during late summer (Poot and Veneklaas 2013). These factors will interact to also influence refugia dynamics.

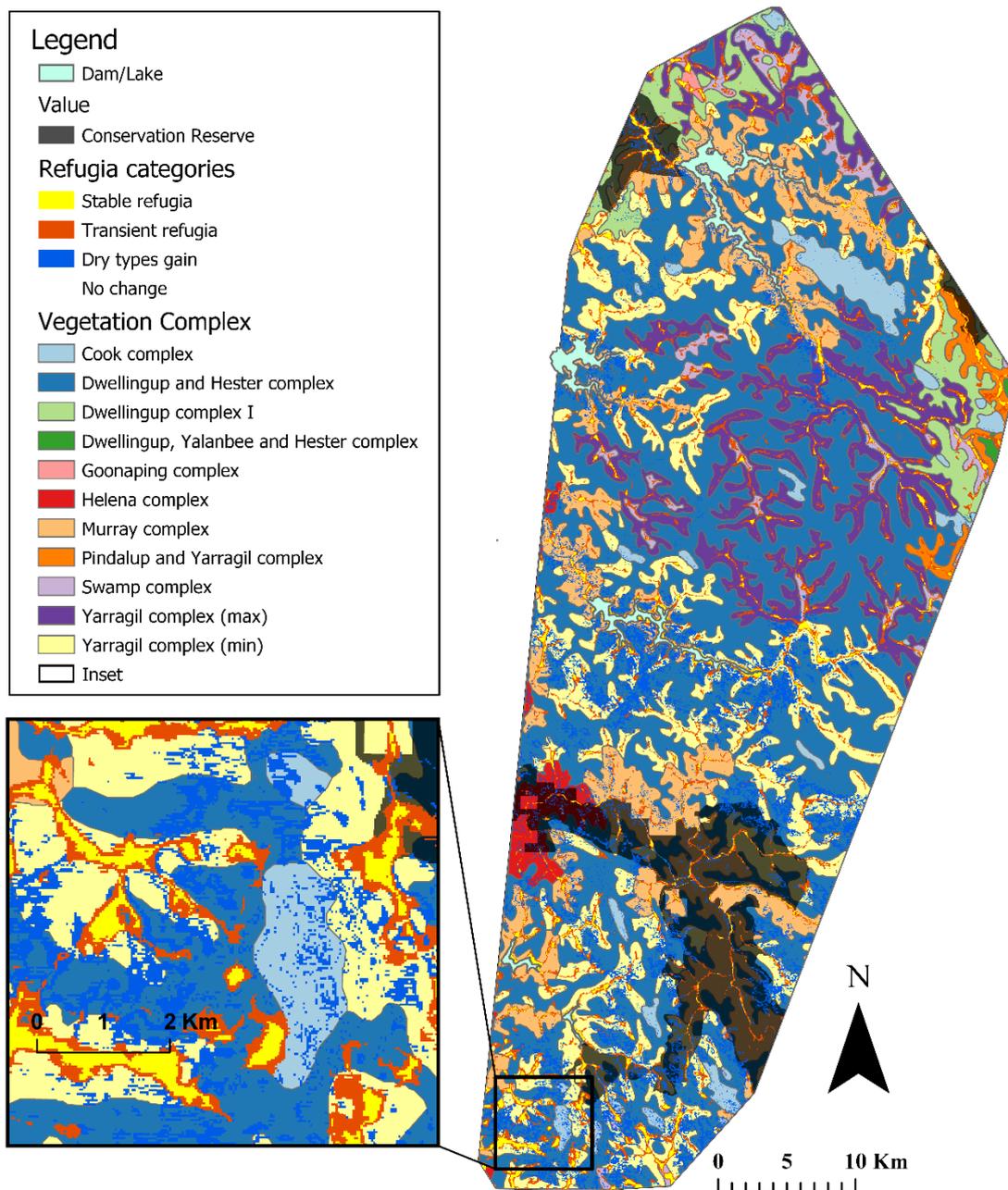


Figure 4.6 Vegetation complexes are a broader unit than the site vegetation types and provide a basis for interpreting hydrologic refugia modelling results in the forest. Refugia are well-represented by the swamp complexes (Swamp and Yarrigal), with very few refugial sites found in the main upland unit (Dwellingup, dark blue).

4.5.3 Hydrologic refugia in an SCP framework

The NJF's current reserve network meets national CAR targets of 10% however, it falls short of Aichi's Target 11 of 17% and was developed prior to widespread recognition of how anthropogenic climate change may impact ecosystems (Figure 4.5) (Groves *et al.* 2012). Extending the protected area network to cover the mesic zones highlighted by modelling is a small increase spatially (from 10 to 13%) and officially recognises the importance of riparian vegetation and its vulnerability to drying. While a

vegetation classification approach enables the delineation of units and percent-based measures for management, this work shows that it isn't adequately capturing the process of drying and its impacts on the restricted riparian zone vegetation. Given the difficulties in classifying NJF vegetation (Chapter 3) (Havel 1975a; 1975b), incorporating alternative approaches like hydrologic refugia modelling into conservation decision-making is important for a truly CAR system.

The results presented here provide a case-study for how climate change impacts can be incorporated into conservation planning through hydrological refugia modelling (Keppel *et al.* 2015; Morelli *et al.* 2016; Michalak *et al.* 2020). The global importance of Mediterranean systems, with lower than average levels of protection, provides additional weight to incorporating refugia into conservation planning in these areas. While summer drought and fire are integral to Mediterranean systems, changes the timing of precipitation, streamflow and groundwater hydrology will impact forests (Wardell-Johnson *et al.* 2015). The structure that SCP and CAR provides administration and governments can easily be extended to incorporate hydrological refugia projections in time and space.

4.6 Conclusion

The expansion in access to high-quality topographic (SRTM DEM) and global vegetation mapping data (the Global Index of Vegetation-Plot Databases; the European Vegetation Archive) means that hydrological refugia models can be applied in many systems worldwide. Streamflow-rainfall elasticity estimations are also available for over 500 catchments worldwide, enabling refugia modelling in landscapes when rainfall and temperature data may only exist at broader resolutions (≥ 1 km). The approach used here provides a means to harness the power of these datasets and determine priorities areas for monitoring and conservation management across a landscape. Refugia maps also provide a practical way to incorporate climate change into current protected area networks. Developments in the availability of data and sophistication of models will improve accuracy, especially if biotic factors and interactions with fire and other disturbances can be incorporated. Additional modelling strategies (e.g. species or dissimilarity modelling) may also support preliminary results, improving confidence in, and uptake of outputs.

Chapter 5 An indicator species approach to defining hydrologic refugia in a drying Mediterranean forest.

5.1 Abstract

Multiple approaches can be used to define hydrological refugia, enabling comparison between different datatypes and modelling methods. Consistency in results will improve confidence in model outputs and increase the likelihood of use in practical applications. Topographic variables derived from 2 m LiDAR data were used to model the current and 2090 distribution of indicator species for riparian plant communities in the northern jarrah forest of southwestern Australia. Results of species and community modelling were then compared and combined to describe zones of stable and transient hydrologic refugia. Maxent models were fitted to 25 species using eight topographic variables. Variable contribution and permutation importance were the highest for SAGA wetness and convexity, with aspect (eastness and northness) also significant. Mean accuracy of the models was 0.62, ranging from 0.5 for dominant, widespread tree species (*Eucalyptus marginata*, and *Corymbia calophylla*) to 0.8 for *Melaleuca preissiana* and *Hakea varia*. Percent change ranged from 80% contraction in *Acacia divergens*, a shrub associated with watercourses and swamps, to 40% expansion of *Trymalium ledifolium*, likely due to its widespread occurrence from creeks to ridgetops. Several species did not perform as thought (*Lepidosperma squamatum* and *E. megacarpa*), likely due to having larger environmental tolerances than expected. Overall, species results support community modelling. Site vegetation type 'C' shows the greatest loss in suitable range (-99%) and its indicator species are predicted to contract the most by 2090 (-63 to -80%). Types A and D follow this trend. Species-results, however, also point to the importance of micro-topography in defining hydrological refugia in this landscape. Within the blocks identified by community modelling, a patchy mosaic of stable refugia were found, varying subtly with convexity, aspect and wetness; and reflecting species different tolerances to drying. These differences provide a challenge to distinguishing between areas of stable and transient refugia, with the possibility that no areas will be stable under projected drying. Our definition of plant-hydrology-terrain relationships is the first attempt at developing indices for climate-ready conservation planning in the northern jarrah forest. Further work investigating the causal relationships underlying the patterns and thresholds found is recommended, including ecophysiological studies and monitoring to measure interactions with fire.

5.2 Introduction

Mediterranean-climate ecosystems (MCE's) house 20% of known vascular plant species in 2% of the world's land area and are one of the biomes predicted to experience greatest biodiversity change over the next 80 years (Sala *et al.* 2000). However, protection is low, with 5% of MCE land-area protected, compared to an average of 12% in other biomes (Hoekstra *et al.* 2005). Of the world's five MCE's, Australia's south-west (SWA) is drying strongly, with a 12-15% reduction in rainfall and 40-100% decline streamflow from pre-1975 levels, and a further 6% decline predicted by 2090 (Petrone *et al.* 2010; CSIRO and Bureau of Meteorology 2015). The area is also one of 34 international biodiversity hotspots (high endemism and >70% of original vegetation cleared) (Myers *et al.* 2000). With only 10% of current

protected areas expected to remain stable to 2100, identifying potential hydrologic refugia is a critical conservation task in this region (Klausmeyer and Shaw 2009).

Hydrologic refugia are sites of elevated or continuous water availability, providing a mesic microenvironment that persists in the face of landscape-scale drying (McLaughlin *et al.* 2017). The capacity of hydrologic refugia to protect species will depend on the physical characteristics of the site and their continuity under further drying. An organism's physiological tolerances and biotic interactions must also continue to be supported (e.g. pollination, dispersal) and refugia may be stable, relative or transient (Keppel *et al.* 2015; McLaughlin *et al.* 2017). Stable refugia include mechanisms that de-couple regional and local climate, while relative refugia track regional changes in climate, becoming transient as critical thresholds are crossed (McLaughlin *et al.* 2017). Physical characteristics of potential hydrological refugia include: topographic heterogeneity (drainage, channels, slope variation, fractures/faults, flood plains), shallow/perched aquifers due to geological and edaphic variation (aquicludes, aquitards, clay/caliche lenses), deep soils or regolith acting as water-storage (Dawson and Pate 1996), and seeps and springs (McLaughlin *et al.* 2017). In SWA, outcropping of springs, the fringes of granite outcrops and the edges of creeks and streams have been identified as potential hydrologic refugia (Wardell-Johnson and Roberts 1993; Wardell-Johnson and Horwitz 2000).

Species distribution modelling (SDM) parameterises a species environmental niche and has been used widely to predict range-shifts under climate change (Guisan and Zimmermann 2000; Elith *et al.* 2006). While popular, criticisms of SDM's include that the resolution of spatial data used for prediction is often too coarse, limiting the characterisation of micro-topography (Robinson *et al.* 2019). Micro-topographic patterns are predicted to be important in determining refugia by providing habitat conditions that are different from the general landscape and therefore buffer regional changes in climate (Keppel *et al.* 2012). In subdued landscapes like SWA, the persistence of moisture conditions (including continuous flow and water-storage) may be entirely dependent on particular micro-topographic conditions. For example, subtly concave locations on flat terrain underlain by an impermeable horizon may store the surrounding region's water in winter (Havel 1975a) and incised areas (small creeks) be the only areas with sufficient moisture for riparian species when the larger landscape dries. Additionally, the capacity of these locations to retain moisture is finite and after thresholds are crossed, will no longer be sufficient. LiDAR data at 1-2 m resolution has recently been used with great success to delimit zones of habitat suitability for rare and threatened banded ironstone taxa (Robinson *et al.* 2019) and the rare mountain endemic *Ornduffia calthifolia* (Keppel *et al.* 2017).

Indicator species show high fidelity to particular environmental conditions, providing a tool for vegetation monitoring and mapping (Dufrêne and Legendre 1997; De Cáceres and Legendre 2009). Fidelity is determined by analysing the concentration of species to a group of sites, with a range of methods available (De Cáceres and Legendre 2009; Ricotta *et al.* 2015). In the northern jarrah forest (NJF), vegetation mapping at local scales has been done using indicator species for 'site vegetation types' (SVTs). SVTs were developed in the 1970's to help identify sites of high productivity for forestry and later adapted for ecological management (Havel 1975a; 1975b; Mattiske 2012). These indicator species provide a means to verify the results of community-level refugia modelling. Individual responses may

mirror, or provide nuances to, predicted community-level change and provide hypotheses for how communities versus species will respond to drying.

Lidar data at two-meter resolution is available for topographic niche modelling for a large extent of the NJF. Indicator species and SVTs have been recorded for 30,000 plot locations within the area, providing the opportunity to analyse hydrologically driven responses at two scales. Community-level modelling of SVTs has shown strong contraction of riparian groups to the edges of waterways, larger tributaries and swampy valley-floors in the forest (Chapter 4). Here, we investigate how key indicator species for these riparian SVTs will respond to the hydrological changes predicted in 2090. The key aims and expectations are:

1. To model the distribution of key indicator species and define the topographic niche for each. Based on the results of Chapter 3 and 4, aspect and wetness are expected to be the most important explanatory variables.
2. To determine which species will provide the most accurate and therefore reliable understanding of how drying will affect distributions. Taxa with the highest dependence on moisture are expected to be the most accurate, due to occupying a narrower niche and more distinct topographic conditions (e.g. valley bottoms only).
3. To describe the difference between community and species responses to drying. We expect predictions to be correlated, but that species-level variation will be more nuanced due to the higher resolution of data used in modelling and differences between species in their tolerance to drying.
4. To use community and species responses to detect and describe the difference between areas of stable and transient refugia. Thresholds of change are expected, based on key topographic features like elevation (e.g. valley bottoms vs up-slope areas) and aspect (e.g. north vs south facing slopes).

5.3 Methods

5.3.1 Study site

The NJF has a Mediterranean climate, with an average annual rainfall of 600-1300 mm that falls predominantly in winter (Havel 1975b; Gentilli 1989). The vegetation is characterised by subdued topography, an east-west rainfall gradient and regular fire incidence. The forest also has a history of logging and mining (1834-current), that has significantly impacted structure (Churchward and McArthur 1980; Wardell-Johnson *et al.* 2015). Floristically, the forest overstorey is dominated by two tree species (*Eucalyptus marginata*, (jarrah) and *Corymbia calophylla* (marri)) (Figure 5.1 i, j) but has a complex and diverse understorey (3096 taxa) (Dell *et al.* 1989; Western Australian Herbarium 1998-). The climatic and hydrological regimes of the forest are shifting; rainfall has decreased 12-15% and temperature increased 0.4% since the mid-1970's (Bates *et al.* 2008). This has led to decreased surface run-off (20-100%), falling water-tables and shifts from perennial to ephemeral streams (Petrone *et al.* 2010; Hughes *et al.* 2012). High summer evapotranspiration offsets winter rainfall, which is predicted to further decrease under climate change.

5.3.2 Study species

To enable comparison to Chapter 4's community modelling results, each species met the requirement of being an indicator species for a riparian SVT that had >200 location records. Riparian SVTs were defined as having a strong, positive correlation with topographic wetness (A, C, D, E, W, Table 4.3, Chapter 4). Despite this requirement, indicator species included both the wide-spread overstory trees jarrah and marri, and riparian-restricted vegetation (e.g. *Melaleuca preissiana* and *Taxandria linearifolia*) (Figure 5.1). In total, 25 species were selected for analysis and general site and soil-preferences noted (Appendix E1). Data was sourced from Alcoa of Australia Ltd (Alcoa), a bauxite mining company that operates in the NJF. Since 1991, Alcoa have been recording floristic and vegetation data at sites based on a ore exploration grid (120 m resolution), with over 31,000 plots collected to date. Species records used in this study were extracted from plot data (20 x 20 m), with points recorded using a Garmin handheld GPS (GDA94 and WGS84).

5.3.3 Environmental variables

Light Detection and Ranging (LiDAR) data collected by Alcoa was used to generate a 1 m digital elevation model (DEM), which was subsequently mosaicked into a single 2 m DEM for the study area (Figure 5.2). Based on (Robinson *et al.* 2019) and logistic regression modelling (Chapter 4), ten morphometric and hydrological predictors were selected. (Robinson *et al.* 2019) found five uncorrelated variables were important for predicting rare species occurrence: Topographic Wetness Index (TWI), Saga Wetness Index (SWI), Annual Solar Radiation, Topographic Position Index (TPI) and Topographic Roughness Index (TRI). Chapter 4, this thesis, found TWI, distance to channel base network, latitude, longitude, plan convexity (CONVEX), distance to valley depth (VALDEP), LS-factor and longitudinal/profile curvature (LONGC) to be significant predictors of community vegetation variation. Aspect was used instead of Annual Solar Radiation due to much lower computing requirements but a high correlation (0.8) (Chapter 4). Note that for convexity positive values indicate convex-upward areas, negative values concave areas and zero flat surfaces (Iwahashi and Pike 2007).

All variables were generated by the System for Automated Geoscientific Analyses (SAGA) (Conrad *et al.* 2015). Due to the large size of the original DEM (15 GB), a virtual machine (48 GB RAM) on Nimbus, Pawsey's cloud service was used to generate each layer. Variables were tested for multicollinearity and values with a correlation coefficient of >0.5 excluded. Aspect was converted to northness (NORTH) and eastness (EAST) in R, leaving eight final variables for modelling (CONVEX, EAST, LONGC, NORTH, SWI, TPI, TRI and VALDEP).

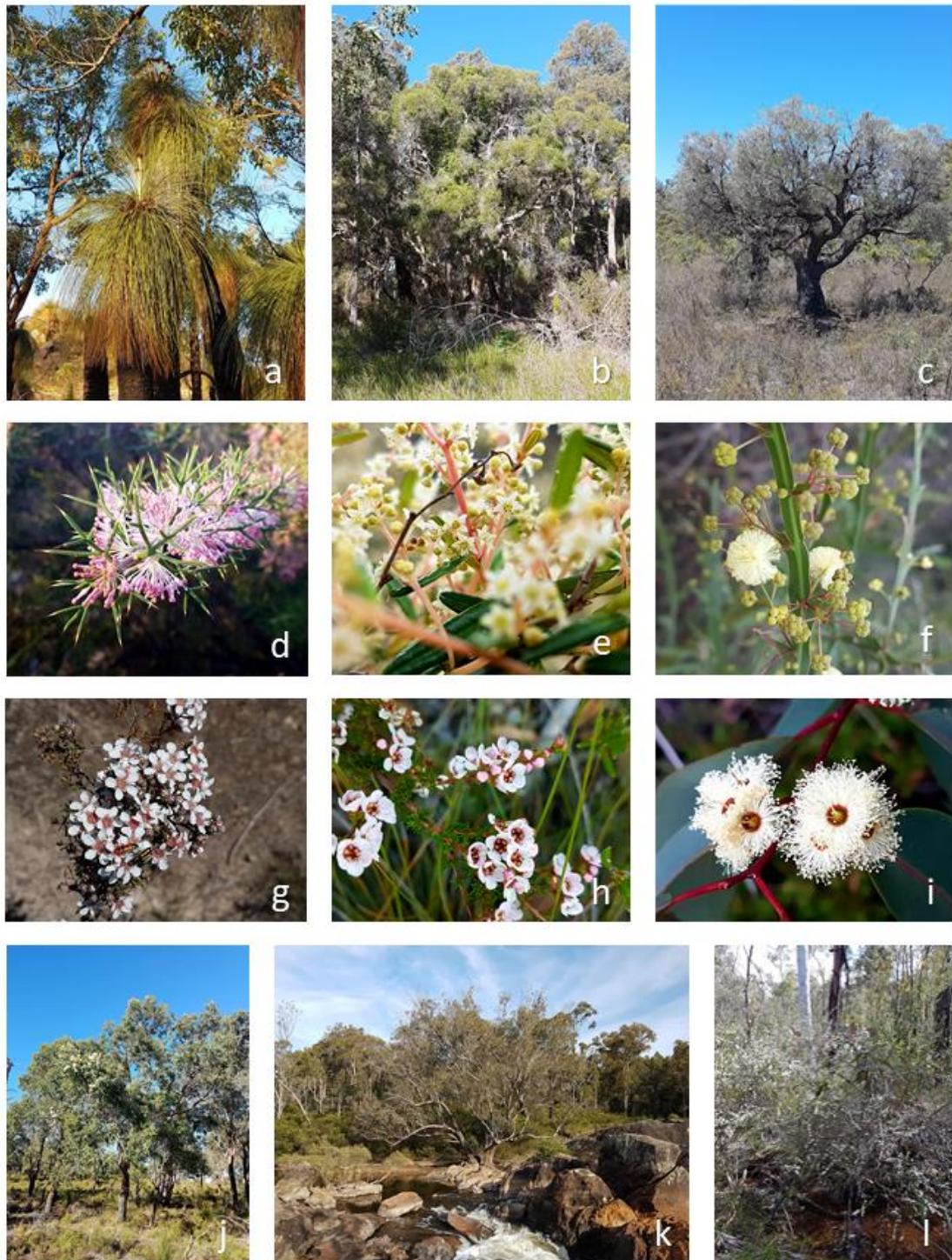


Figure 5.1 Selection of species used in hydrologic refugia modelling: the tree-like monocot a) *Kingia australis*; key swamp-related trees b) *Melaleuca preissiana* and c) *Banksia littoralis*. d) *Hakea lissocarpa*, e) *Trymalium ledifolium* and (f) *Acacia alata* occupy a wide range of conditions. (g) *Pericalymma ellipticum* is found in elevated areas on seasonally swampy platforms and (h) *Babingtonia camphorosmae* occurs mainly on heavy soils. The dominant upland overstorey trees in the northern jarrah forest are i) *Eucalyptus marginata* and j) *Corymbia calophylla*. k) *E. rudis* is associated with the wetter parts of south western Australia, but it can also be found on hillsides, while l) *Taxandria linearifolia* borders swamps and watercourses. Photos, S. Luxton.

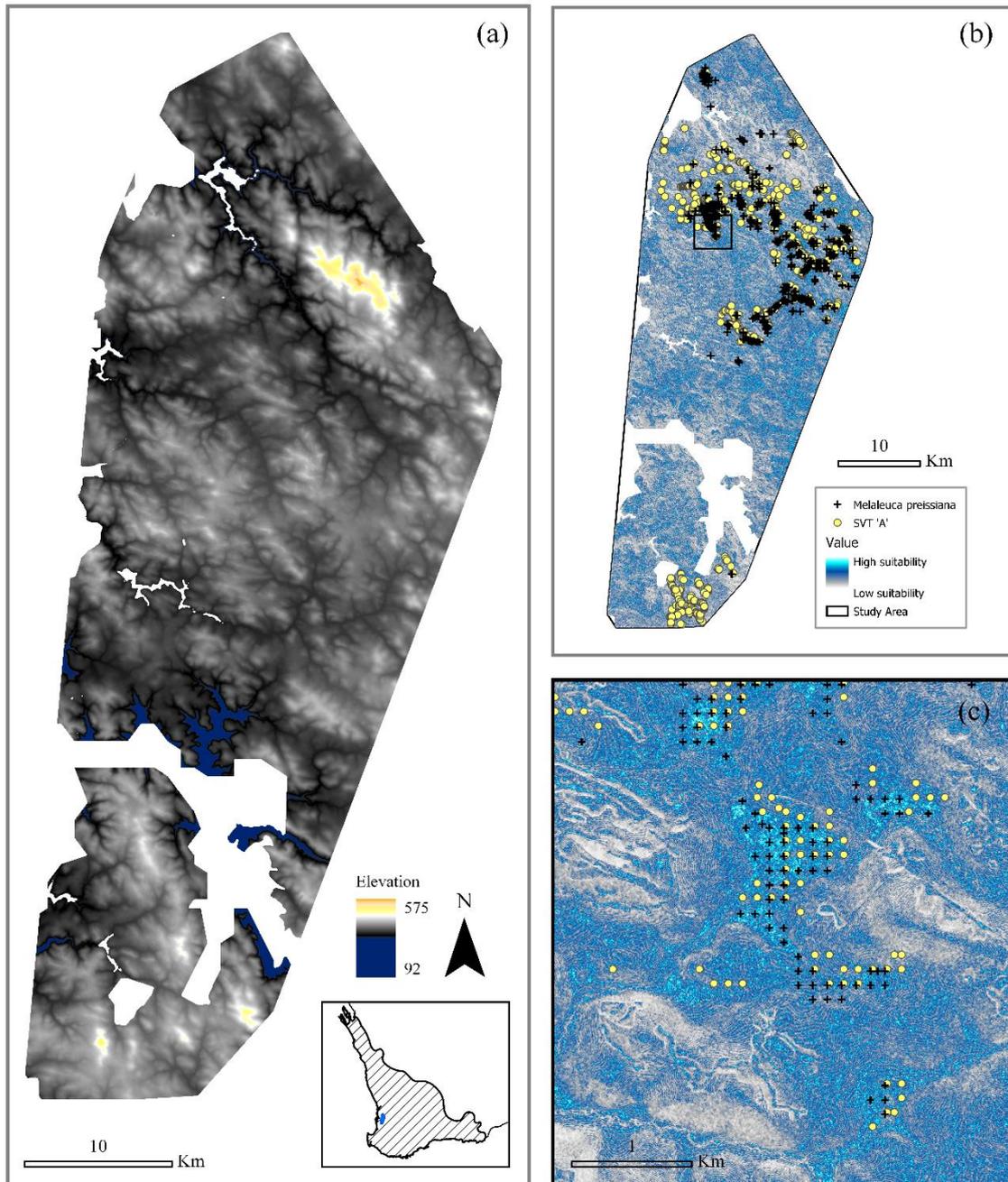


Figure 5.2 (a) Lidar data is available at 2 m resolution for a large section of the study area, enabling elevation and other topographic variables to be derived; (b) exemplar showing *Melaleuca preissiana*, a riparian shrub or tree and its predicted habitat suitability in 2090. Differences in the species presence (cross) and the site vegetation type it represents (yellow, SVT 'A') are evident; (c) a close-up of habitat suitability in 2090. *M. preissiana* prefers sandy soils and swamps, forming open stands in shallow depressions.

5.3.4 Streamflow elasticity estimator and rainfall

Sankarasubramanian's (2001) non-parametric precipitation elasticity of streamflow estimator (ϵ_p) was used to relate current and future streamflow and rainfall [1]:

$$[1] \quad Q_t = \bar{Q} + (\epsilon_p \cdot \bar{Q} \cdot (P_t - \bar{P})) / \bar{P}$$

where Q_t = change in wetness index (site); \bar{Q} = current wetness index (site); ϵ_p = streamflow-rainfall elasticity estimator; P_t = predicted mean annual rainfall (region); \bar{P} = current mean annual rainfall (region). ϵ_p provides a measure of the sensitivity of streamflow to changes in rainfall and has been calculated for 219 catchments in Australia (Chiew 2006) and over 500 catchments worldwide (Chiew *et al.* 2006). SWA has a mean ϵ_p of 3.0 (a 10% decrease in rainfall corresponds to a 30% decrease in streamflow). SWI represents water flow and accumulation across a landscape and provides a surrogate for streamflow data across the study area.

Rainfall actual values for 1975-2006 were based on gridded climate data from Commonwealth Scientific and Industrial Research Organisation's (CSIRO) Representative Climate Futures Framework (Clarke *et al.* 2011; Whetton *et al.* 2012; Williams *et al.* 2018). Projected values were developed using the Australian Governments 'Climate Change in Australia' 'Climate Futures' tool, using representative concentration pathway (RCP) scenario 4.5. This tool compares the results of 60+ models (for SWA) and stratifies predictions (e.g. % decline in annual rainfall) by the number of models that agree with an outcome (e.g. 17/68 models predict a 'much drier' (15% decrease) annual rainfall under RCP 4.5 in 2090 (see <https://www.climatechangeinaustralia.gov.au/en/climate-projections/climate-futures-tool/projections/> and Appendix E11 for full details).

Additional projections were obtained from the CSIRO Sustainable Yield Project (CSIRO 2012) and previous work on streamflow decline in SWA (Petroni *et al.* 2010; Hughes *et al.* 2012) (Appendix E2). Both the Climate Futures tool and Sustainable Yield Project report a 12.5% decrease in rainfall from historic conditions (1986-2005 and 1975-2007 respectively), which was used as the baseline value to model the "current" (2006) distribution of vegetation (Appendix E2).

5.3.5 Modelling & accuracy assessment

Maximum Entropy (MaxEnt) species distribution models (SDM) in R were used to model the current (2006) and future (2090) distribution of each species (package 'dismo') (Hijmans *et al.* 2017) (Phillips *et al.* 2006). MaxEnt identifies the maximum entropy between two probability distributions (background points and presence data) and fits a set of 'features' equivalent to coefficients to predict habitat suitability (for a full explanation see (Elith *et al.* 2011)). While the data used in this study could be described as presence-absence for perennial species (Chapter 2), the use of MaxEnt accounts for biased or incomplete absence data (Hijmans and Elith 2015; 2019). Permutation importance for each predictor was recorded and a response curve based on a sample of 100 points drawn.

Model accuracy was determined using inbuilt Maxent statistics, including area under the curve (AUC), correlation (cor) and balanced accuracy. Models were then binarised based on the inflection point: maximum of true positive rate and false positive rate, using independently calculated receiver operating characteristic (ROC) curves (Lippitt *et al.* 2008). Data was split into training (80%) and test sets (20%) prior to modelling (random sample, R) and accuracy, kappa and significance independently calculated based on the binarised layers (ROC, confusion matrix). K-fold cross-validation was not used due to the high resolution of the environmental data and time taken for each model to run (6-12 hours). The area of presence in 2006 and 2090 was calculated and percent change recorded (R script, Appendix E13).

5.3.6 *Defining species vs community refugia, variability and thresholds*

Potential hydrologic refugia were defined based on stacked SDMs for taxa that were positively correlated with SWI ('negative percent change', Table 5.2). Each species was binarised and assigned '1' for presence and '0' for absence. Layers were then summed and split into three refugia categories: 1-7 species (non-refugia); 8-12 species (transient refugia, 50-79% of species present in 2090); and 13-17 (stable refugia, 75-100% of species present in 2090). Cells with seven species present or less were not classed as potential refugia (<50%). The proportion of species represented by each refugia category was assessed using the 'Combine' tool in ArcGIS and summary statistics calculated.

To capture variability in responses to drying and attempt to define potential thresholds of change, [1] percent and [2] raw change in habitat suitability (px) was also calculated per cell and plotted for the original data points for each riparian species. Delta(px) is the overall change in habitat suitability as a percentage from 2006 (px06) to 2090 (px90), while 'raw(px)' is total change.

$$[1] \quad \text{delta(px)} = ((\text{px90} - \text{px06}) / \text{px06}) * 100$$

$$[2] \quad \text{raw(px)} = \text{px06} - \text{px90}$$

Finally, species and community refugia maps were overlain in ArcGIS Pro (version 2.2.0) and the different approaches visualised graphically.

5.4 Results

5.4.1 *Species distribution models*

The 25 indicator species modelled provide insight into how topography is related to habitat suitability across the study area at 2 m resolution (Figure 5.2 b,c). Five riparian SVTs are represented by the species, with discernible patterns in variable importance between the groups. Overall, convexity and ruggedness had the highest influence in drier species and types (E, D, W), with SWI having a higher variable contribution in the wettest types (A, C) (Figure 5.3). Convexity was greatest for the two dominant tree species *Eucalyptus marginata* (76%) and *Corymbia calophylla* (72%) (Appendix E3), but also high for the shrubs *Trymalium ledifolium* and *Pericalymma ellipticum*. Eastness was consistently more important across the wettest SVTs (11-26% vs 3-10%), although *E. megacarpa* was an exception (21%). Northness

made the greatest contribution to *Hakea lissocarpa*, a non-riparian shrub, but otherwise does not show a clear pattern between species or SVT groups. Overall, SWI contributed the greatest amount in species associated with SVT A (mean = 33% vs 13% across all other groups), but unexpectedly was very low in the C-related taxa *Lepidosperma squamatum* (1.7%), *Acacia alata* (2.2%) and *Boronia molloyae* (3.8%) (Appendix E3), which is the type predicted to contract the most by 2090 (Table 4.5, Chapter 4). Topographic position was low across all taxa (0-7.8%) (Figure 5.3).

Modelling provides a basis for defining each taxa's topographic niche (Figure 5.4, Appendix E4). Convexity and SWI varied the most, with upland species associated with more convex surfaces (Figure 5.4 (h) *Corymbia calophylla*, (i) *E. marginata*, (y) *Trymalium ledifolium*). Riparian taxa were not associated with flat or concave terrain (represented by zero or negative values), but convexity was generally lower (e.g. Figure 5.4 (l) *E. rudis*, Appendix E4). Roughness and slope curvature did not discernibly vary for most species (Figure 5.4). Wetness has a distinct window of influence, sharply increasing above 0.75, then plateauing or declining above 0.85 (Figure 5.4, Appendix E5). Aspect (both northness and eastness) has a subtle influence unless at the extremes – e.g., distinctly north or east-facing slopes. Figure 5.4 (d) *Astartea scoparia* has a preference for south-west facing slopes, as does (b) *Acacia divergens* and (u) *Lepidosperma tetraquetrum*. For (m) *Gahnia decomposita* a directly east-facing slope is more important. Figure 5.4 (s) *Kingia australis* provides an exception to general patterns, with median levels of convexity and high (>0.8) ruggedness correlated with low suitability.

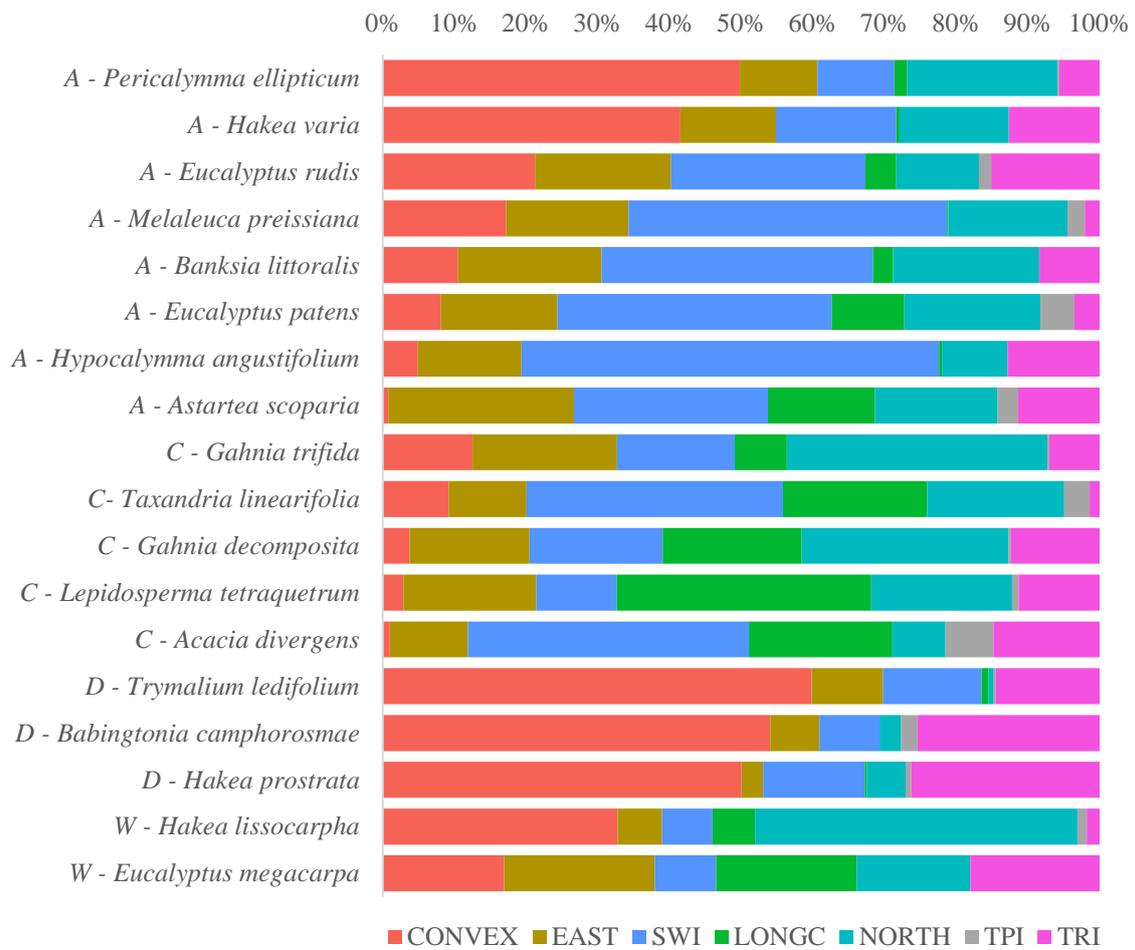


Figure 5.3 Stacked bar chart of permutation importance (%), grouped by site vegetation type (SVT), with convexity showing higher importance in the drier types and saga wetness in the wettest. Abbreviations: CONVEX, convexity; EAST, eastness; SWI, SAGA wetness index; LONGC, longitudinal curvature; NORTH, northness; TPI, topographic position index; TRI, topographic ruggedness index.

— CONVEX — EAST — SWI — LONGC — NORTH — TRI

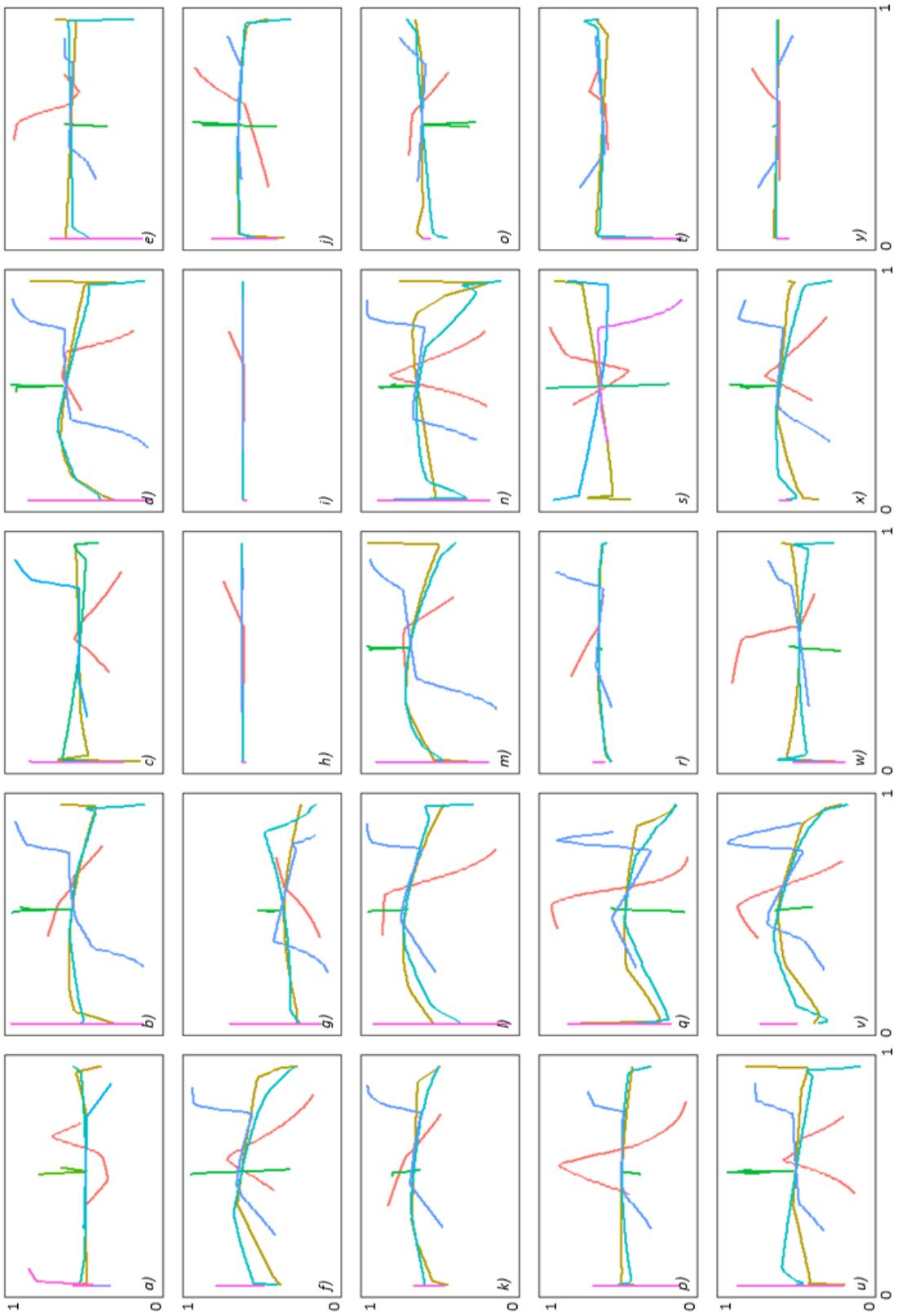


Figure 5.4 Species response curves developed from standardised environmental variables (x-axis) depicting habitat suitability (y-axis) at relative levels for six of the predictors used in modelling. Species are: (a) *Acacia alata*, (b) *Acacia divergens*, (c) *Acacia extensa*, (d) *Astartea scoparia*, (e) *Babingtonia camphorosmae*, (f) *Banksia littoralis*, (g) *Boronia molloyae* (h) *Corymbia calophylla*, (i) *Eucalyptus marginata*, (j) *E. megacarpa*, (k) *E. patens*, (l) *E. rudis*, (m) *Gahnia decomposita*, (n) *Gahnia trifida*, (o) *Hakea lissocarpha*, (p) *Hakea prostrata*, (q) *Hakea varia*, (r) *Hypocalymma angustifolium*, (s) *Kingia australis*, (t) *Lepidosperma squamatum*, (u) *Lepidosperma tetraquetrum*, (v) *Melaleuca preissiana*, (w) *Pericalymma ellipticum*, (x) *Taxandria linearifolia*, (y) *Trymalium ledifolium*. Abbreviations: CONVEX, convexity; EAST, eastness; SWI, SAGA wetness index; LONGC, longitudinal curvature; NORTH, northness; TRI, topographic ruggedness index.

Model accuracy varied from 0.5 for dominant, wide-spread tree species (*Eucalyptus marginata*, and *Corymbia calophylla*) to 0.8 for *Melaleuca preissiana* and *Hakea varia* (mean = 0.62, s.d. = 0.08) (Table 5.1). Kappa was fair to moderate for most riparian-related species, but only slight for upland species (0.01-0.12). Percent change and area lost (km²) was greatest for *A. divergens* (-80%, -287 km²), down to just -7% and 32km² lost for *E. rudis*, among the species predicted to decline. Gains in habitat area are predicted for eight species, but due to poor accuracy and kappa, will not be considered further.

5.4.2 Species vs community responses

Species results show both similarities and differences to community predictions (Figure 5.5). Across the SVTs, community decline tracked species, e.g., negative percent change was greatest for species in ‘C’, which also includes the taxa predicted to contract the most (Table 5.1). Overall, of the 11 species with a mean SWI greater than 6 (the most moisture-dependent species), 64% contracted by more than 50%.

This contrasts community modelling, where C is predicted to lose 99% of its required moisture conditions by 2090 (Table 4.5). This is greater than any of the individual species, where the range in responses is also high (from -80% for *A. divergens* to +40% for *Lepidosperma squamatum*, mean -37%, s.d. 42.7). The indicator species for SVT A are lower overall than community expectations (mean species = -33% vs -60%) but are more consistent than C (s.d. = 17.8, Table 5.1). Across the other ‘riparian’ SVTs (D, E, W), D and W also include species that will both expand and contract, despite strong community predictions of decline (D, -70%; W, -28%).

Finally, while field mapping of SVTs was based on the presence of indicator species, there is a large amount of variability in the SVTs recorded at the same site as each species (Appendix E6). For example, 66% of the *Hakea prostrata* records were scored as SVT D (for which it is an indicator) but it was also recorded at the same location as A, C, P, S, T and W (Appendix E6). Only 50% of sites were scored as C for *Acacia divergens*, and widespread species like *T. ledifolium* (D) and *E. marginata* (E) had very low number of sites that were scored the same (3.6% and 0.3% respectively) (Appendix E6).

Table 5.1 Modelling results and summary statistic for each species, including accuracy and percent change from 2006 to 2090. Site vegetation types (SVT) were used to group data. Grey = the species used to define hydrologic refugia, this chapter.

| Species | Records | SVT | Mean SWI | Habit | AUC | COR | BAL ACC | CI | Kappa | P-val | M P-val | Area 2006 | Area 2090 | Loss (km) | % Change |
|----------------------------------|---------|-----|----------|-------|------|------|---------|-------|-------|--------|---------|-----------|-----------|-----------|----------|
| <i>Astarea scoparia</i> | 471 | A | 6 | S | 0.63 | 0.23 | 0.63 | 0.14 | 0.25 | 0.016 | 0.4 | 538 | 235 | -303 | -56 |
| <i>Eucalyptus patens</i> | 2470 | A | 6 | T | 0.64 | 0.24 | 0.61 | 0.07 | 0.2 | 0.14 | <0.001 | 449 | 197 | -252 | -56 |
| <i>Hypocalymma angustifolium</i> | 4163 | A | 6 | S | 0.57 | 0.12 | 0.57 | 0.04 | 0.14 | 1 | <0.001 | 445 | 265 | -180 | -40 |
| <i>Banksia littoralis</i> * | 1003 | A | 6 | T | 0.68 | 0.31 | 0.66 | 0.1 | 0.3 | <0.001 | 0.026 | 609 | 414 | -195 | -32 |
| <i>Melaleuca preissiana</i> * | 560 | A | 6 | S/T | 0.83 | 0.58 | 0.8 | 0.11 | 0.6 | <0.001 | 0.88 | 444 | 307 | -137 | -31 |
| <i>Pericalymma ellipticum</i> | 371 | A | 2 | S | 0.64 | 0.24 | 0.67 | 0.16 | 0.33 | <0.001 | 0.78 | 595 | 421 | -174 | -29 |
| <i>Haakea varia</i> * | 239 | A | 6 | S | 0.89 | 0.7 | 0.82 | 0.16 | 0.64 | <0.001 | 0.33 | 293 | 253 | -40 | -14 |
| <i>Eucalyptus nuda</i> s | 234 | A | 5 | T | 0.77 | 0.48 | 0.62 | 0.21 | 0.23 | 0.2 | 0.4 | 472 | 440 | -32 | -7 |
| <i>Acacia divergens</i> | 405 | C | 6 | S | 0.64 | 0.24 | 0.64 | 0.15 | 0.27 | 0.24 | 0.04 | 369 | 72 | -297 | -80 |
| <i>Gahnia trifida</i> | 371 | C | 3 | G | 0.67 | 0.3 | 0.6 | 0.17 | 0.16 | 1 | <0.001 | 334.5 | 110 | -225 | -67 |
| <i>Gahnia decomposita</i> | 329 | C | 6 | G | 0.58 | 0.14 | 0.67 | 0.17 | 0.31 | 0.26 | 0.31 | 428 | 148 | -280 | -65 |
| <i>Traxandria linearifolia</i> | 1622 | C | 6 | S/T | 0.67 | 0.28 | 0.65 | 0.07 | 0.3 | <0.001 | 0.2 | 551 | 206 | -345 | -63 |
| <i>Lepidosperma tetraquetrum</i> | 437 | C | 2.5 | G | 0.71 | 0.33 | 0.67 | 0.15 | 0.31 | 0.15 | 0.01 | 338 | 168 | -170 | -50 |
| <i>Boronia mollis</i> øe | 203 | C | 2 | S | 0.79 | 0.52 | 0.72 | 0.21 | 0.43 | 0.008 | 0.21 | 412 | 300 | -112 | -27 |
| <i>Acacia elata</i> | 1516 | C | 2 | S | 0.64 | 0.24 | 0.57 | 0.08 | 0.13 | 0.09 | 0.22 | 487 | 541 | 54 | 11 |
| <i>Lepidosperma squamatum</i> | 1232 | C | 2.5 | G | 0.57 | 0.12 | 0.54 | 0.09 | 0.07 | 0.93 | 0.03 | 501 | 703 | 202 | 40 |
| <i>Haakea prostrata</i> | 378 | D | 6 | S | 0.78 | 0.44 | 0.72 | 0.15 | 0.43 | <0.001 | 1 | 377 | 238 | -139 | -37 |
| <i>Babingtonia camporosmae</i> * | 916 | D | 6 | S | 0.61 | 0.19 | 0.55 | 0.11 | 0.09 | 0.72 | 0.14 | 642 | 482 | -160 | -25 |
| <i>Corymbia calophylla</i> | 25068 | D | 2 | T | 0.57 | 0.12 | 0.52 | 0.02 | 0.04 | 1 | <0.001 | 508 | 519 | 11 | 2 |
| <i>Trymalium ledifolium</i> | 12796 | D | 2.5 | S | 0.56 | 0.11 | 0.52 | 0.03 | 0.03 | 1 | <0.001 | 759 | 1082 | 323 | 43 |
| <i>Eucalyptus marginata</i> | 29415 | E | 2 | T | 0.56 | 0.1 | 0.51 | 0.019 | 0.015 | 1 | <0.001 | 457 | 459 | 2 | 0.4 |
| <i>Kingia australis</i> * | 224 | E | 2 | M | 0.66 | 0.3 | 0.55 | 0.22 | 0.1 | 0.8 | 0.3 | 391 | 416 | 25 | 6 |
| <i>Acacia extensa</i> * | 419 | W | 3 | S | 0.55 | 0.09 | 0.56 | 0.15 | 0.1 | 0.99 | 0.006 | 830 | 611 | -219 | -26 |
| <i>Eucalyptus megacarpa</i> | 1335 | W | 2.5 | T | 0.62 | 0.21 | 0.56 | 0.08 | 0.12 | 0.35 | 0.09 | 603 | 613 | 10 | 1.7 |
| <i>Haakea lissocarpa</i> | 3611 | W | 2 | S | 0.55 | 0.09 | 0.54 | 0.05 | 0.09 | 0.7 | 0.01 | 682 | 729 | 47 | 7 |

5.4.3 Stable vs transient hydrological refugia

Stable refugia were defined during community modelling as pixels where riparian SVTs (A, C, E, D, W) did not change (remained their original SVT), while areas that shifted to a new ('drier') riparian type were classed as transient refugia (Figure 4.5, Table 4.5, Appendix D3). For species modelling, areas that retained 75-100% of taxa predicted to contract were termed stable refugia and 50-75% transient (Figure 5.5). Overlap between the two approaches includes the major creeks, valleys and low-lying swamps highlighted by community modelling (Figure 5.5, Figure 4.6). Thresholds between areas of stable, transient and non-refugia can be seen in these areas – e.g., at valley bottoms and when switching aspect (north-facing slopes have a very low probability of riparian species presence in 2090) (Figure 5.5 d). While consistencies are present, species results are also both patchier within the areas identified by community modelling to be 'stable', and regularly present outside of the major blocks of community refugia (Figure 5.5 c, d, e, Appendix E7).

Fine-scale patterning across the 2 m topographic contours used in species modelling is evident (Figure 5.5, Appendix E7), and differences in stable, transient and non-refugia are reflected by the underlying predictors. Convexity decreases as refugial capacity increases, while there is a marked increase in wetness between non-refugia and refugia areas (Figure 5.6 A). Stable refugia are also generally found lower in valleys, on more south-facing slopes and where the terrain is less rugged (Figure 5.6 A). This corresponds with community findings, where the riparian SVTs (A, C, D, E, W) occur lower in the landscape and at higher wetness (Figure 4.3).

Species contributions to each refugia category vary (Figure 5.6 B, Appendix E1). While all species are found present to approximately the same degree for stable refugia, there's an increase in several species' representation for transient. *Melaleuca preissiana*, *Banksia littoralis*, and *Babingtonia camphorosmae* occur more frequently in this zone, while *Acacia divergens* and *Lepidosperma tetraquetrum* are low. Species that are unexpectedly well-represented outside of the refugial zones include *E. rudis* and *Banksia littoralis* (Figure 5.6 B, Appendix E1).

Defining the threshold between areas of stable and transient refugia is quite subjective and using species frequency to indicate different zones is just one approach (Figure 5.5). Species-level variability in the predictability of suitable habitat may also help to elucidate change-points, with higher variability signalling a shift from stable conditions, or the quality of an indicator for tracking change (Figure 5.8). For a few species, a decline in percent suitability is matched by raw values (e.g. *A. extensa*, -30% = -0.3, but for others, a drop-in percent suitability corresponds to both small and large changes in raw values (*Melaleuca preissiana*, *Boronia molloyae*) (Figure 5.8 b).

Finally, the streamflow elasticity constant used to adjust SWI results in a maximum value of 0.72 in 2090. Individual species response plots show optimal habitat suitability at 0.75-0.85 SWI (Appendix E5), opening the possibility that the required moisture conditions for stable refugia will not exist for the study area in 2090 (Appendix E8).

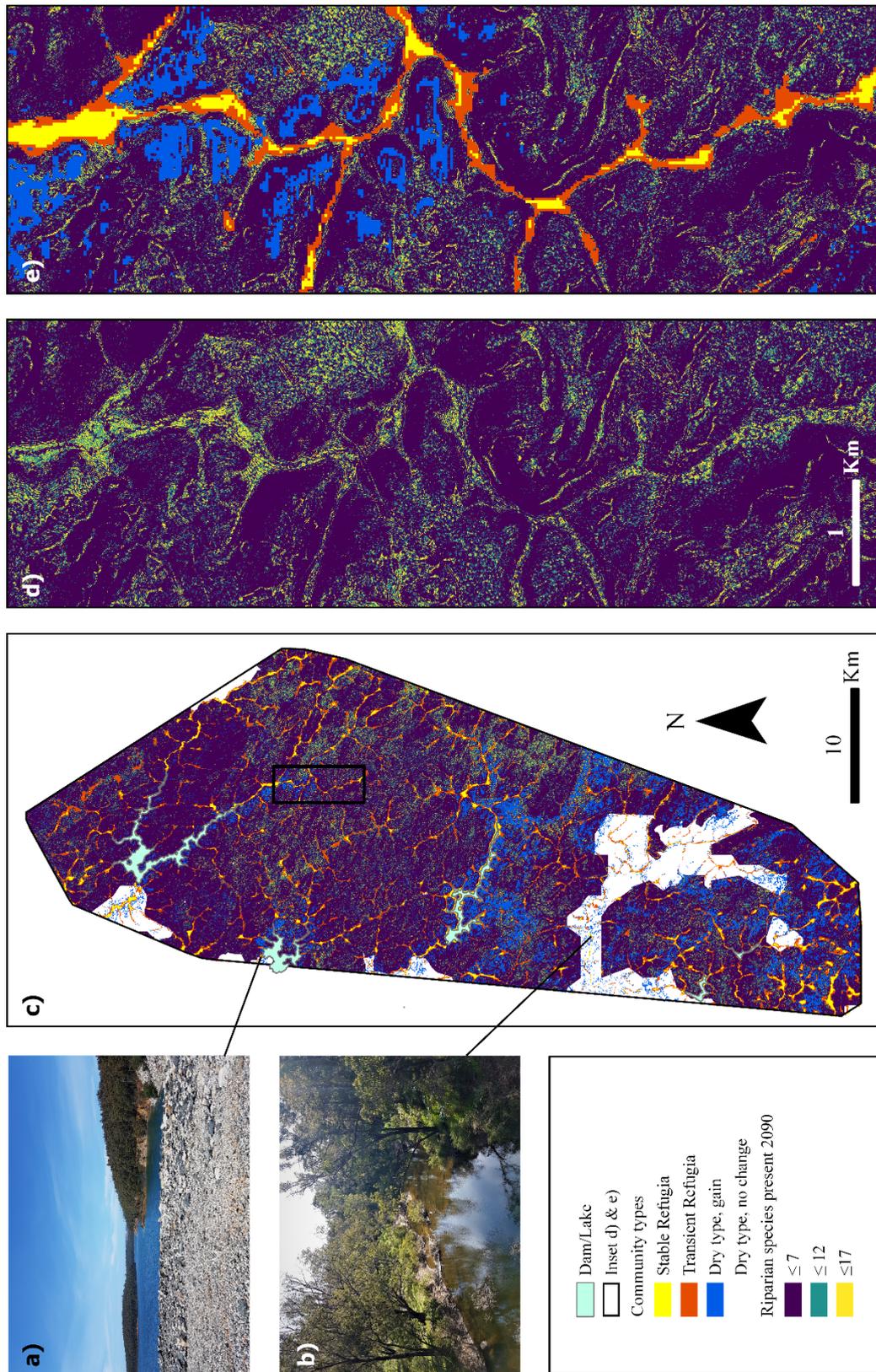
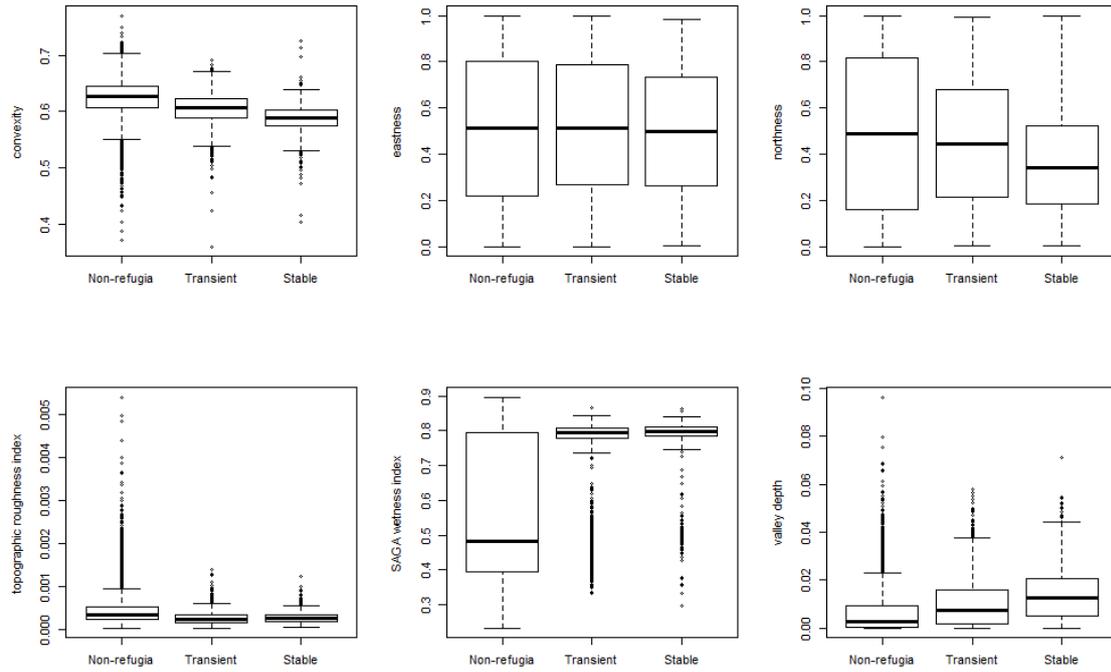
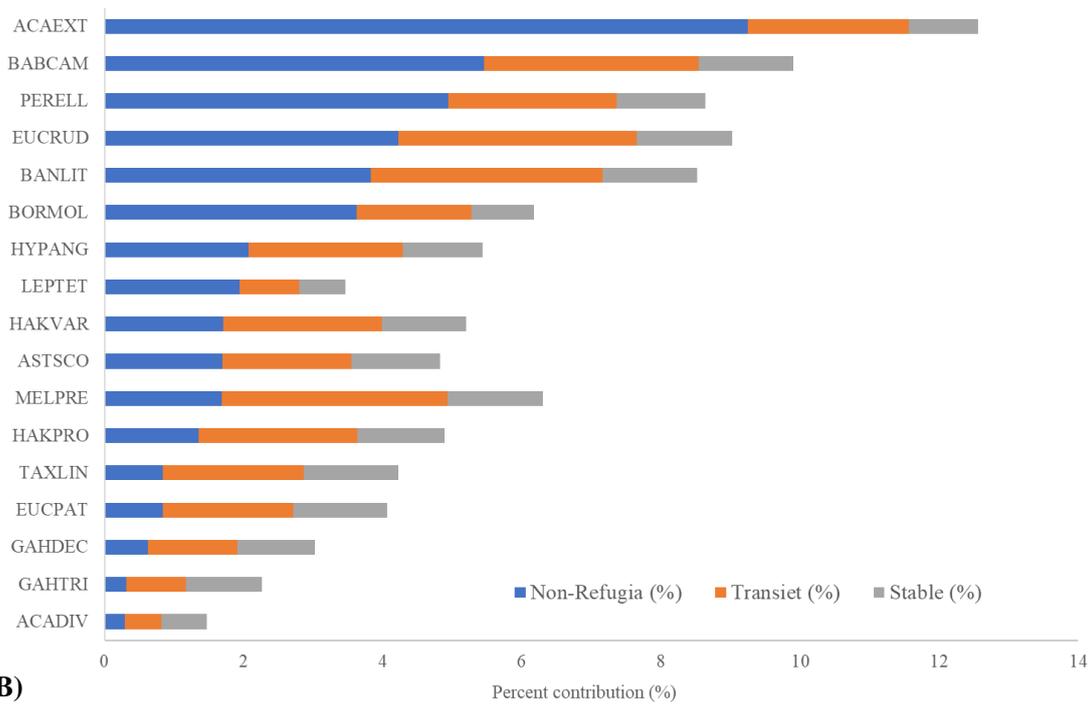


Figure 5.5 Species modelling provides further insight into the dynamics of potential stable and transient hydrologic refugia in the study area: a) the areas around dams include refugia but may not be suitable due to human influences like clearing to enhance run-off and manipulation of water-levels; b) the Murray River flows through Lane Pool Conservation Park and represents a major tributary in the study area; c) stacked species and community models of potential hydrologic refugia showing how d) species results both correspond with and show a more nuanced perspective to the (e) community model.



(A)



(B)

Figure 5.6 (A) Stable, transient and non-refugia areas show differences in the topographic predictors used in modelling. Convexity decreases as refugial capacity increases, while there is a marked increase in SAGA wetness between non-refugia and refugia areas. (B) Species contributions to stable, transient and non-refugia – all species contribute similarly to stable refugia (0.7-1.4%), but not non-refugial areas (0.3-9%).



Figure 5.7 (a) The subdued topography of the northern jarrah forest lends itself to complex patterns of potential hydrologic refugia; (b, c) small creeks include many riparian species but are identified to lose taxa with continued drying; (d) some upland creeks already show evidence of encroachment by species associated with dry conditions like (e) *Banksia grandis*; and (f) particularly following fire (*Corymbia calopylla*); (g) The banks of larger streams will likely hold the moisture conditions required for riparian species to 2090 and are the obvious location of potential hydrologic refugia. However, how drying will actually impact more marginal sites (e.g. (h) wet, leached acid sands, underlain by an impermeable horizon and waterlogged in winter) that include mixed suits of upland and riparian taxa requires further work. The areas identified by modelling provide a framework for monitoring, decision-making and further work in the forest. Photos, S. Luxton.

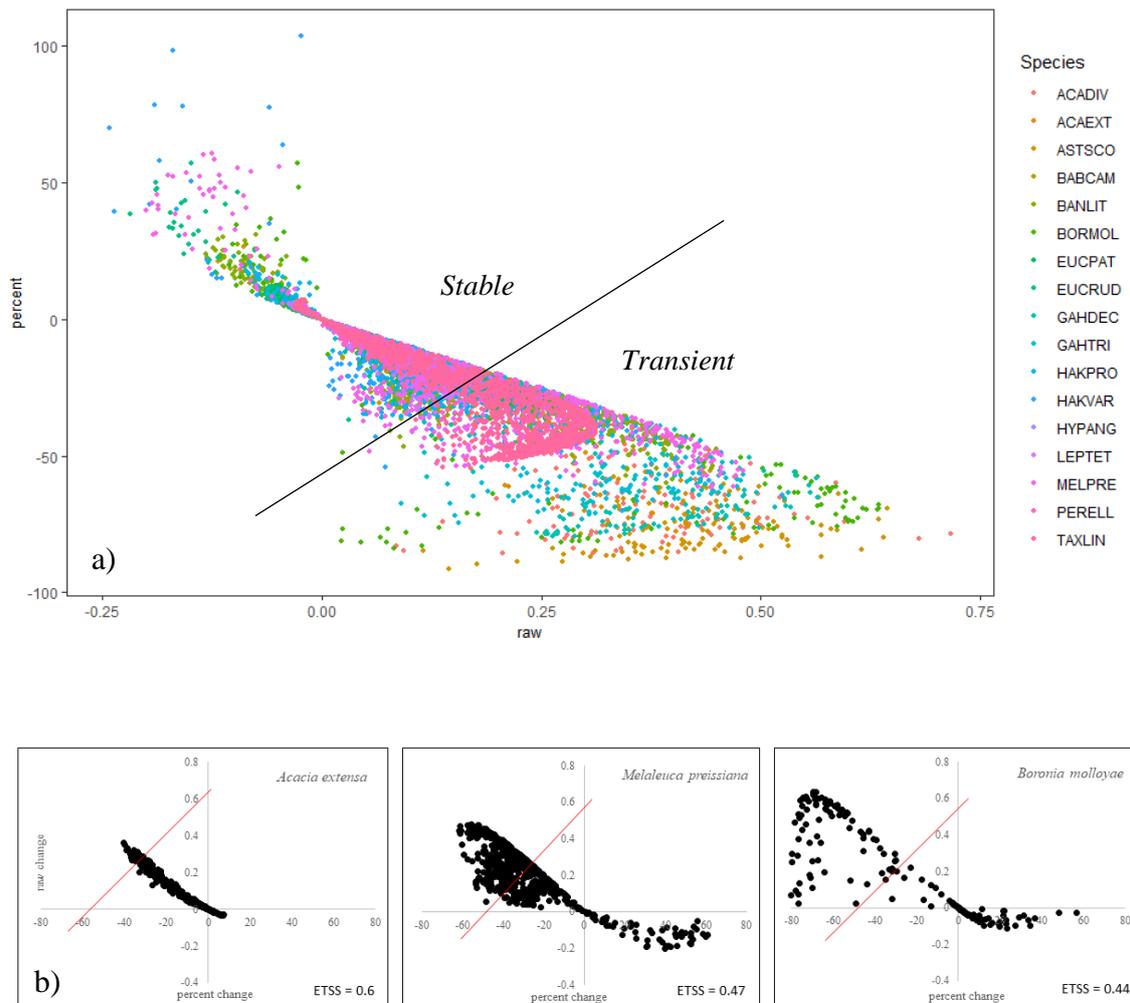


Figure 5.8 a) Changes in percent and raw habitat suitability across the range-contracting species (2006 to 2090) may assist with defining the threshold between stable and transient refugia; b) in some taxa (e.g. *Acacia extensa*) raw and percent change are consistent but in others (*Melaleuca preissiana*, *Boronia molloyae*) a percent drop in suitability corresponds to both small and large decreases in raw values. See Appendix E10 for all species.

5.5 Discussion

Anthropogenic climate change has led to declining rainfall and streamflow across southwestern Australia. Riparian plant species are predicted to contract into areas that continue to support their moisture requirements despite drying – hydrologic refugia. These areas may be stable (persistent) or transient, disappearing as certain thresholds of water-availability are crossed. The effects of drying were modelled for 25 taxa that are indicator species for riparian plant communities in the forest. Topographic variables generated from high-resolution Lidar data (2 m) were used to parameterise the models. An established relationship between changes in regional rainfall and streamflow decline was used to project to 2090 conditions. Finally, results were compared to community-level modelling for the same area. Results are complementary, with both sets of models being driven by convexity, aspect and wetness, and contain areas of similarity for where stable refugia will occur. However, variance at the species-level and the high

resolution of species data (2 m), also show potential hydrologic refugia as more dispersed and patchier than community results alone. This definition of refugia at two scales provides insight into patterns not evident by a single-method approach, while fine-scale patterning across the 2 m topographic contours illustrate how microtopography may mediate regional changes in climate. Results form a basis for conservation management under drying and highlight areas for experimental and monitoring work in the forest.

5.5.1 *Species distribution models*

SDMs define the realised environmental niche by correlating available predictors with a species occurrence. We sought to define the topographic dimension of potential hydrologic refugia at two scales, with the expectation that consistency between model outputs would improve confidence in results. In both the community and species models, aspect, convexity and SWI were key factors, despite the differences in data resolution (2 m vs 30 m). Aspect has also been shown to be the critical factor in separating broad community types (Figure 3.6, Chapter 3), as well as an important gradient underlying SVTs (Figure 4.4, Chapter 4, Figure 5.4). Functionally, aspect is an indirect predictor of plant dynamics, moderating direct and resource variables like solar radiation and temperature (Guisan and Zimmermann 2000). These in turn influence soil moisture levels, heat stress and rates of evapotranspiration, creating the moisture conditions that riparian species require on south and southwest facing slopes (Figure 5.4) (Bennie *et al.* 2006).

Variables differ in their capacity to elucidate the topographic niche of species (Appendix E4). While important, the subtlety of convexity makes it difficult to distinguish clear differences and correlate it to particular terrain or landform types (Figure 5.4, Appendix E4). A separate analysis that classifies landform types relative to remotely sensed convexity may assist (e.g. Iwahashi and Pike 2007). Conversely, there is a clear gradient in response to SWI in both the species and community results (Table 5.2, Figure 5.5). SWI models wetness across a landscape but doesn't see flow as a film, rather, values are calculated to account for "filling", and it is thought to be a more realistic indicator of potential soil moisture. Most species predicted to decline show a peak in habitat suitability at 0.75-0.9 SWI, (Figure 5.4, Appendix E5), which is also the portion of the wetness spectrum predicted to disappear by 2090. Results provide a clear description of patterns of soil moisture across the landscape and a measurable signal for change.

5.5.2 *Reliability of species predictions*

Taxa with the highest dependence on moisture were expected to be the most accurate in modelling, due to occupying a narrower, more distinct topographic niche (e.g. valley bottoms only). This was generally found to be the case (e.g. *A. divergens*, -80%, median SWI = 6.1, Table 5.2, Appendix E1). Exceptions include *Gahnia trifida* and *Lepidosperma tetraquetrum*, that were predicted to contract by over 50% but had a mean SWI of only 2.8-2.9 (Appendix E1). Variation in soils across the topographic gradients of the forest and difference in ecophysiological capacity are likely the cause of discrepancies. For example, *Pericalymma ellipticum* (-29%, median SWI = 2.3) occurs in elevated areas on seasonally swampy platforms and is a well-recognised riparian species (also known as "Swamp Teatree") (Western Australian Herbarium 1998-). Conversely *E. marginata*, which is known as an upland species, had a

median SWI of 2.4, but was also recorded in some of the wettest areas of the forest (max SWI = 7.6). Jarrah has a high degree of phenotypic plasticity, ranging in form from mallee's in the dry northern end of its range to large forest trees in the higher-rainfall south (Dell *et al.* 1989). An unexpected result was that for *E. megacarpa*, which was classified by Havel as a core indicator species for riparian type C (the type predicted to decline the most in the forest) and is often found near swamps and streams (Western Australian Herbarium 1998-). *E. megacarpa*'s modelled distribution had both a low accuracy (0.56) and it was predicted to slightly expand (+1.7%) (Table 5.1), in contrast to its association with a riparian community.

Overall, the most widely distributed species had the lowest SWI and accuracy. They are found across the elevational and spatial gradients of the study area, tolerating a large range of environmental conditions. For these species, making predictions for how they may expand into areas of potential refugia is difficult. Therefore, the focus of further work is on the responses of species predicted to decline.

5.5.3 *Species vs community responses*

Community responses to drying create a distinct network of refugia in the forest, tracking the upper reaches of streams, and flats and basins of larger valleys (Figure 4.5, Figure 4.6). Conversely, species-mapping of refugia tracks the key habitat features of community modelling, but also highlights a patchy mosaic of potential microsites outside of these major topographic features (Figure 5.5, Appendix E7). Both the similarities and differences observed between these results are in-part explained by the different resolutions of the datasets. The 30 m grid-cell of the SRTM dataset averages variation across the study area, estimating changes in hydrology more broadly. For planning purposes – this may be a more useable scale at which to work. However, the results from 2 m modelling provide deeper insight to topographically driven change and indicate that micro-topography may also have a role to play in retaining the moisture conditions that species will require. Microtopography is well-recognised as a way that plants can compensate for regional differences in climatic conditions (Guisan and Zimmermann 2000). This may include slight changes in aspect that provide more shading, or concave sites at the meter scale that collect more water than the surrounding area (Robinson *et al.* 2019). Differences between model methods (multinomial logistic regression vs stacked SDMs) and model accuracy (Table 4.4 vs Table 5.1) also contribute to differences in results.

Biotic factors operating at the community vs species level may also be important in determining the differences between modelled results (Walther 2010). Emergent community properties include the timing of phenological events (Walther 2010), below-ground facilitative relationships with soil mycorrhiza (Bunn *et al.* 2009; Albornoz *et al.* 2016) and local species interactions (le Roux *et al.* 2014). Community integrity may contribute to the persistence of more vulnerable species through positive feedback with soil moisture conditions (e.g. increase of infiltration) (Dekker *et al.* 2007).

Alternatively, the ability to describe in detail the habitat preferences for an individual species enables the capture of fine-scale details including: soil preferences (Orians and Milewski 2007); patterns of dispersal (Bennie *et al.* 2006); and seed and seedling dynamics (e.g. Pate *et al.* 1990) as well as fine-scale hydrological patterns persisting at a local level (Wardell-Johnson and Horwitz 1996; Sander and Wardell-Johnson 2011).

5.5.4 Stable vs transient hydrological refugia

Based on community results, clear thresholds of change between stable, transient and non-refugia were expected. These boundaries were confirmed, but a patchy mosaic of suitability was also found (Figure 5.5). The features associated with distinct areas of stable hydrologic refugia mirror community results - including the edges of major water bodies and creeks in conjunction with aspect (south-west facing slopes), and swampy valley floors (Figure 5.5). Creek and stream zones are well-recognised as important areas in the forest. They are mapped as distinct units, along with swamps and categorised as 'informal reserves' in the forests reserve system, within a 150 m buffer zone (The State of Western Australia 1999) (Appendix A1, Appendix D9). While dam edges show areas of potential hydrologic refugia, they are not likely to be suitable (Figure 5.5 a). Due to the damming process, surrounding vegetation includes predominantly upland species, while silviculture treatments are used to enhance runoff and inflow (Figure 5.5 a) (Corporation 2015). As a conservation prioritisation tool, the confirmation of community results by species modelling provides an additional layer of evidence in support of potential hydrological refugia areas, and of the value of formally including them in the current reserve network.

McLaughlin *et al.* (2017) define relative/transient refugia as sites that have the capacity to preserve differences in relative moisture availability, despite experiencing drying in absolute terms (i.e. they are not de-coupled from general drying, as stable refugia are). The fine-scale patterning across the 2 m topographic contours shown by species models show a complex mosaic of transient (or relative) refugia, not captured by broader-scale modelling (Figure 5.5). These results provide evidence for, and insight into, the role of microtopography in providing and maintaining hydrological diversity at the site scale (Sander and Wardell-Johnson 2011; Schut *et al.* 2014). Characteristics include a similar average wetness to stable refugia and a slight increase in convexity and northness (Figure 5.6). The importance and relevance of these sites is reflected in the species that occupy the transient zone (Table 5.2). *Eucalyptus rudis*, *Banksia littoralis* and *Melaleuca preissiana* are the most abundant taxa, with *M. preissiana* occurring at a higher frequency there than in stable, or areas not identified as refugia (Table 5.2).

Understanding mechanisms of change is important if modelled results are to have on-ground implications. Differences in life history strategy, including the seed and seedling biology of riparian vs upland taxa will likely play a critical role. Riparian plants have physiological and morphological adaptations that enable them to survive flooding and fluvial disturbance (Capon and Dowe 2007). The riparian tree *B. littoralis* (Swamp Banksia) grows in low-lying winter-wet locations and wetland fringes (Western Australian Herbarium 1998-; Groom 2004). Seedlings are flood tolerant relative to species that occupy well-drained sandy soils, enabling their survival during the wet winter and spring months (Groom 2004). Groom (2002) also found that *B. littoralis* is less water-use efficient than xeric congeners (Groom 2002). Ecophysiological factors such as these, the fire-sensitivity of many riparian taxa and their slower time to maturity, in conjunction with the forecast increase in the frequency and severity of fires will be important drivers of vegetation change in the forest (Williams *et al.* 2009). Shifts in species distributions are already starting to be seen, with some upland creeks showing evidence of encroachment by species associated with dry conditions, particularly after fire (Figure 5.7).

5.5.5 Defining thresholds

Multiple approaches to detecting and understanding potential thresholds of change are needed, due to ecosystem complexity and the range of processes interacting that will determine change (Lindegren *et al.* 2012). Analysing variability in different species responses to changes in habitat suitability provides an additional way to understand how drying in the landscape will impact distributions (Figure 5.8). The receiver operating curve (ROC) method used to delineate suitable vs unsuitable habitat is based on how well the model detects actual presences (and absences) vs ‘false’ positives/absences (Lippitt *et al.* 2008). However, species responses to habitat suitability vary both discretely and continuously (Austin 2013). Species-level variability in the predictability of suitable habitat may help to elucidate change-points, by signalling a shift from stable conditions (lower variability) to less suitable and drying conditions (higher variability) (Figure 5.8).

A larger range of potential 2090 values may also signal that a species is losing the capacity to track available moisture, or that it has high plasticity, and therefore is not a good indicator of change. For example both *Acacia extensa* and *Boronia molloyae* are predicted to decline by the same amount (-26% vs -27%), but have very different patterns in changes of variability (Table 5.1, Figure 5.8 b). These differences are also likely to contribute to the dispersed, patchy patterns of refugia – with overlapping variation in species adaption to both the current and future moisture profile. Punctuations at certain thresholds in environmental predictors (Figure 5.4) or change (Figure 5.6, Figure 5.8) may occur, but patterns may also be mixed and difficult to define (Figure 4.4, Figure 5.5).

To adequately understand change and where thresholds may lie in refugia dynamics, linking tests of remotely sensed results with ground-based measurements are needed. Potential approaches include: correlating SWI with soil moisture and rates of transpiration (or thresholds of transpiration shutdown in species like jarrah). Alternatively, predicted habitat suitability could be used as a basis for a multi-factorial experiment investigating seedling germination and establishment rates across a fire x drying gradients. Linking results from modelling projects with on-ground experimental work would be a powerful approach to informing conservation and resource plans.

5.6 Conclusion

Species distribution modelling of indicator species for riparian plant communities predicted to contract under climate drying provides insight into broad and micro-topographic patterns of potential hydrological refugia. While key areas, including the upper reaches of streams and valley floors, have been identified as sites of moisture retention, small-scale refugia also persist. The patchy distribution of these microsites exhibits subtle changes in convexity and aspect from the surrounding matrix, with particular soil or geological elements likely. This patterning of local-scale refugia sits in contrast to the percent-area based units that form the current conservation reserve network in the forest. With evidence of species range changes in response to drying in higher-elevation creeks, and post-fire recruitment of upland taxa into stream zones, inclusion of potential refugia into the reserve network is urgent. This approach provides a flexible and data-driven means to link current reserve strategies with climate change adaption, an essential step-forward in the management and conservation of this drying Mediterranean forest.

Chapter 6 General discussion

The aim of this thesis was to model potential hydrological refugia to determine priority areas for conservation action under climate change. By parametrising the data to be used in modelling within the wider floristic and spatial context of the region, **Chapter 2** showed that a best-practice modelling approach should focus on shrub and tree species and limit model projections to the study area. **Chapter 3** found that the current units used for forest mapping at the community scale are the best-available for further modelling work, after classification of the dataset did not result in clear community groupings. Potential hydrological refugia were then modelled at the community (30 m) scale and projected to a 2090 drying scenario (**Chapter 4**). Key riparian communities were shown to contract significantly, including 99% loss of the wettest type. Hydrological refugia were predicted to occur along the edges of major streams and in broad winter-wet valleys, forming a process-based picture of important conservation areas that lie outside of the current reserve system. In **Chapter 5**, maxent modelling of indicator species provided a more in-depth analysis of potential change. Species responses were more variable than communities, both confirming the location of potential hydrologic refugia mapped in Chapter 4 and highlighting the role that microtopography may have in retaining mesic sites within the larger, drying, forest matrix. Habitat suitability curves were also used to estimate thresholds between areas of stable, transient and non-refugia. These curves are a first step towards a more in-depth analysis of the drivers that form potential hydrologic refugia and a way to identify signals of change. Final maps reflect the topology of the forest, its complex relationship with vegetation and patterning at two scales – with both a patchy mosaic and areas of distinct hydrological refugia found. In this final chapter, I discuss the value and limitations of the work in the context of the literature; and argue for the inclusion of refugia in the reserve system, as a mechanism for bringing climate change adaptation into conservation policy in the forest.

6.1 The value and limitations of a hydrologic refugia approach

Refugia theory provides a way to distinguish areas of stability within larger bioclimatic envelopes of change and relate quantitative results to policy and decision-making (Keppel *et al.* 2012; Keppel *et al.* 2015; Morelli *et al.* 2020). After over a decade of scientific discussion and work to conceptualise, define and quantify refugia; uptake by practitioners is gaining momentum (Morelli *et al.* 2020). The basis for identifying refugia is now well-developed, including that the type of refugia must be pre-identified relative to the study system and threatening processes of interest (e.g. refugia from disturbance, drought, heat or cold) (Keppel *et al.* 2012; Keppel *et al.* 2015). This also extends to whether refugia are in-situ or ex-situ (Ashcroft 2010); macro or micro (Ashcroft *et al.* 2012; Stralberg *et al.* 2020a); stable, transient or relative (McLaughlin *et al.* 2017; Cartwright *et al.* 2020) and occur as holdouts or a larger network across a landscape (Hannah *et al.* 2014; Keppel and Wardell-Johnson 2015; Morelli *et al.* 2020).

This study identified a network of hydrological refugia driven by landscape position, topography and soil moisture (Figure 4.5, Figure 5.5). Within this network a gradient of refugia likelihood occurs (Figure 5.5, Figure 5.8); from stable (the retention of the most moisture dependent communities and species) to varying degrees of transience. Transient areas hold different combinations of species and communities (Table 4.5, Figure 5.8) and occur both at the edges of stable refugia and as ‘micro-holdouts’ across the landscape (Figure 5.5). While predicted to lose their original species compositions, transient refugia will

act as relative zones or ‘ex-situ’ refugia (Ashcroft 2010), between areas of stable and non-refugia for less moisture dependent riparian species (Table 4.5). They will provide a ‘slow lane’ for biodiversity under drying (Morelli *et al.* 2020).

A key strength of this thesis is using both community (which captures interactions, shared responses and assembly processes) and species modelling (direct gradient responses) (Austin *et al.* 2009; Austin 2013; Mokany *et al.* 2019). This contrasts to other climate-change related studies in SWA, which have focused on: specific taxa (e.g. *Banksia*, *Ornduffia*) (Fitzpatrick *et al.* 2008; Yates *et al.* 2010; Keppel *et al.* 2017); assessing model metrics (Molloy *et al.* 2016); modelling vegetation structure and cover at fine scales to identify refugia (Schut *et al.* 2014); and the development of a conceptual framework to facilitate biodiversity adaptation that prioritises actions against risk (e.g. preventing anthropogenic degradation (low risk) vs fire management and the protection of refugia) (Prober *et al.* 2012).

More broadly, this thesis fits into a strong literature of refugia modelling in Australia. Nationally, an assessment of refugia has led to the call for modelling at fine scales, to complement and provide more realistic assessments of change (Reside *et al.* 2013; Reside *et al.* 2014). Multi-species modelling has been done for Tasmania's palaeoendemic flora to identify past, present and future refugia, and connectivity between them (Mokany *et al.* 2016); while phylogeographic work has provided evidence for multiple localized refugia during Pleistocene climatic oscillations across southern Australia (Byrne 2008). Topoclimatic grids have been used to improve predictions of fern and grass occurrence in the Hunter Valley of New South Wales (Slavich *et al.* 2014). Multiple climate scenarios gave insight into distributional decline for Australian butterfly species, including that 88% of species modelled (n=24) would decline under a conservative scenario (temperature increase of 0.8–1.4°C by 2050) (Beaumont and Hughes 2002). Historic work has set the foundation for refugia in Australia, including the importance of fine-scale hydrological patterns for sustaining refugia (Morton *et al.* 1995; Wardell-Johnson and Horwitz 1996). Finally, centers of endemism, with historical stability, topographic complexity and dispersibility have been used to identify refugia in subtropical Australia (Weber *et al.* 2014).

Worldwide, refugia studies span data types, modelling methods and timeframes (Stewart *et al.* 2009; Morelli *et al.* 2016). While not explicitly related to refugia, forecasting species changes in response to global warming is a vast area of research (Araújo and Luoto 2007; Butt *et al.* 2016; Araújo *et al.* 2019; Michalak *et al.* 2020). Most directly relevant to this study are hydrological and mesic refugia (Davis *et al.* 2002; Byrne *et al.* 2011; Davis *et al.* 2013; Weber *et al.* 2014; Dillon *et al.* 2015; McLaughlin *et al.* 2017; Cartwright *et al.* 2020). The definition of hydrological refugia by McLaughlin *et al.* (2017) and their description of stable vs transient and relative refugia became a cornerstone of this thesis. The concept of zones of refugia was the catalyst for Chapter 5, and the attempt to measure and map the drivers and location of thresholds of change (Figure 5.6, Figure 5.8, Appendix E5, Appendix E10). It is a more nuanced view of refugia and landscape change, and when ground-truthed, a valuable tool for management (McLaughlin *et al.* 2017; Barrows *et al.* 2020; Cartwright *et al.* 2020).

The major limitation of a hydrologic refugia approach (and other forms of refugia and climate change modelling more generally) is predicting future ecosystem states based on correlative modelling of current conditions (Elith and Graham 2009; Yates *et al.* 2018; Briscoe *et al.* 2019). The structure of environment-

to-species relationships may not hold under future conditions, due to processes interacting non-linearly to change (Butt *et al.* 2016; Briscoe *et al.* 2019). Additionally, climate may not change uniformly (Beaumont *et al.* 2008) and extreme events increase in intensity and frequency (Harris *et al.* 2018), which are not well-accounted for in correlative models. Multiple scales and types of analysis are recommended (Araújo and Peterson 2012; Araújo *et al.* 2019). However, standard protocols to assess model uncertainty are limited (Yates *et al.* 2018; Araújo *et al.* 2019; Thuiller *et al.* 2019; Zurell *et al.* 2020); and there are difficulties getting the required data for more complex process-explicit modelling (Briscoe *et al.* 2019). Additionally, the full distribution of species may be unknown, making the validation of thresholds and extremes difficult (Barrows *et al.* 2020).

Refugia can be modelled using a number of approaches and no 'best' method exists (e.g. Anderson *et al.* 2000; Ashcroft *et al.* 2012; Gavin *et al.* 2014; Guillera-Arroita *et al.* 2015; Mokany *et al.* 2016; Michalak *et al.* 2020; Stralberg *et al.* 2020b). Methods should be matched to the question, data-type and end-use (Guillera-Arroita *et al.* 2015), while standards to assess model choice are in development (Araújo *et al.* 2019; Elith *et al.* 2020; Zurell *et al.* 2020). The limitations of the analyses used in this thesis are three-fold: statistical, data-input and lack of biotic-interactions. Statistical issues include the limitations of correlative modelling mentioned above, that other models may have improved predictive performance (e.g., machine learning), and a lack of uncertainty testing and external validation of results. Data-inputs were only topographic, due to a mismatch in resolution between climate (1 km) and topography (30 m), however, other methods (e.g. topoclimate grids) have been successful in modelling climate at fine scales (Ashcroft 2006; Ashcroft and Gollan 2012; Slavich *et al.* 2014). Finally, incorporating biotic interactions is a developing frontier that has been shown to improve the predictive performance of ecological models (Bardgett and Wardle 2010; Van der Putten *et al.* 2010; Wisz *et al.* 2013; Morales-Castilla *et al.* 2015).

6.1.1 Statistical limitations

The strength of multinomial linear regression is that it enables the relationship between predictors and response variables to be parameterized. However, other algorithms like neural networks, random forest trees and generalised dissimilarity modelling (GDM) may have had better predictive power (Breiman 2001; Elith *et al.* 2006; Ferrier and Guisan 2006; Ferrier *et al.* 2007; Bourel and Segura 2018). Machine learning algorithms maximize predictive power using unsupervised techniques that search the natural structures in data and build relationships for prediction from this. However, they are often a black box where it is difficult to interpret drivers (e.g., no predictor-response relationship is described). Random forest models have gained enormous traction over the past 20 years (Couronné *et al.* 2018), however, while Couronné *et al.* (2018) found that they performed significantly better than logistic regression, the gain was only marginal (an increase in accuracy of 0.029%). GDM has a wide range of applications, including reserve network selection and incorporates both environmental and species data to predict beta-diversity (House *et al.* 2012; Molloy *et al.* 2016; Di Marco *et al.* 2019). The use of GDM with soil data (e.g. the soil and landscape grid of Australia (Grundy *et al.* 2015)) would likely improve predictive performance.

After further reading of the SDM literature other methods are likely better suited to a dataset of this size and structure than Maxent (Guillera-Arroita *et al.* 2014; Thibaud *et al.* 2014). For example, Thibaud *et al.*

(2014) compared Maxent to three presence-absence methods and found a generalized linear model (GLM) with the probit link function, quadratic terms for each predictor and no interactions to have the best accuracy with a large sample size ($n=500$). Maxent outperformed other models in smaller sample sizes ($n=100$), but Guillera-Arroita *et al.* (2014) produced results that disagreed with this, and they argue for the use of GLM's in small datasets with presence-absence data as well. In hindsight, a GLM, or other model that takes advantage of well-structured presence-absence data may have led to better predictive performance. Additionally, final SDM results should have been aggregated up to 20 m resolution (e.g. mean value for a 20 m² grid cell), to mirror the scale that species data was collected at (plot size) – rather than producing maps at the scale of the LiDAR data (2 m).

Some potentially meaningful environmental variables could also have been dropped from modelling due to statistical requirements. For example, rainfall is likely an important driver of vegetation patterns in this system, particularly at broader scales and across the east-west gradient of the forest, which dries as you move east. However, it was not directly used in the analysis because the data resolution was coarse relative to other environmental and floristic data (1 km vs 30 m and 120 m). This coarseness equates to a lower correlation with, and a poorer signal of, vegetation change; whereas fine-scale variables will track the vegetation better statistically, but potentially not ecologically (i.e., correlation doesn't equal causation). This potential shortcoming means that the rainfall-vegetation relationship was not captured by modelling.

Uncertainty testing and external validation have been identified as two areas in ecological forecasting that require improvement (Beaumont *et al.* 2007; Beaumont *et al.* 2008; Harris *et al.* 2018; Araújo *et al.* 2019). Beaumont *et al.* (2007) found that within model variability (e.g., changing the starting conditions or internal parameters) can have a larger influence on a projected species occurrence than different climate models do. In contrast, Thuiller *et al.* (2004) found that the type of SDM used had a greater influence. Incorporating variability across SDMs, dispersal strategies, global circulation models and representative concentration pathways (RCP), (Thuiller *et al.* 2019) assessed uncertainty for ~11,500 amphibian, bird and mammal species. They found that multiple SDMs, RCPs, dispersal assumptions and GCMs are needed to adequately assess and report uncertainties. While the majority of modelling subsets data into training and testing sets to build and validate models (Rykiel Jr 1996), Barrows *et al.* (2020) argue for different kinds of data to ground-truth and validate models. For example, genetic traits were used to validate refugia for Whitebark pine (*Pinus albicaulis*) in the western US. Only a small portion of protected refugia were predicted to protect the genetic diversity of the pine as climate change intensified, clarifying conservation and management needs (Barrows *et al.* 2020).

Ground-based studies are needed in the NJF to understand how forecast change will affect physiological (e.g., drought tolerance and mortality thresholds) and ecological processes in the forest (e.g., recruitment and establishment dynamics and their interaction with drying and fire). In the Mojave Desert, demographic data (recruitment abundance relative to distance from refugia and patterns of pollination and viable seed set) have been used to field-verify modelling (Sweet *et al.* 2019; Barrows *et al.* 2020). The findings of this thesis (Figure 5.5, Figure 5.8) could also be validated using field data, e.g. by measuring seedling germination and establishment, and adult mortality, in riparian species across gradients of drying.

6.1.2 Data limitations

Two key technical data-limitations encountered during this thesis (among others, Appendix F1), was the use of the topographic wetness index (TWI) and streamflow elasticity equation as the basis for climate projections.

TWI is designed to be a surrogate for complex hydrological processes, providing a spatial description in lieu of direct measurements like soil moisture (Moore *et al.* 1991; Conrad *et al.* 2015). When paired with species data it is a powerful way to understand and scale plant-topography-hydrological relationships across large areas. However, it was not partitioned into component parts and cannot account for how solar radiation and soil physical and chemical properties affect runoff and soil moisture (Ma *et al.* 2010; Kopecký *et al.* 2021). Also, no tests of alternative TWI algorithms were done, which can be a source of significant variation (Kopecký and Čížková 2010; Pei *et al.* 2010; Kopecký *et al.* 2021). Additional modelling would be required to calculate more complex aspects of hydrology like soil moisture storage and groundwater recharge. An example is TOPMODEL; a physically based, distributed model that simulates hydrologic fluxes of water through a watershed, including explicit ground and surface water interactions (Beven and Kirkby 1979; Beven *et al.* 1984; Wolock and McCabe Jr 1995). This would introduce challenges, as hydrological models, like global climate models are sources of significant uncertainty in ecological forecasting (Feng and Beighley 2020).

Streamflow elasticity was treated as a constant over the whole study area, which has implications for results. The original work that defined the rainfall-streamflow runoff relationship for southwestern Australia ('Elasticity' $\epsilon_p = 3$, where a 10% reduction in rainfall equates to a 30% decline in streamflow/runoff) was done at the catchment scale (Chiew 2006). However, raw data from the West Australian Water Corporation shows considerable variation in streamflow (discharge rate: m^3/sec) from 1970 to 2010 (20-80% drop) (Figure 1.2). As predictions did not incorporate variation in the rainfall-streamflow relationship, results will be incorrect at some locations (e.g., streams with especially high or low flow rates). As mentioned above, ideally a hydrological model based on the raw streamflow data would be used to calculate discharge rates for ungauged streams and percent decline interpolated across the study area at the stream scale. Hydrological modelling, however, is a complex and was outside of the scope of this study (Beven and Kirkby 1979; Devia *et al.* 2015; Gemitzi *et al.* 2017; Singh 2018; Visser *et al.* 2019; Feng and Beighley 2020).

Topoclimatic grids provide another way to overcome issues with broad scale climate data (Ashcroft 2006; Ashcroft and Gollan 2012; Gollan *et al.* 2013). They use forcing-factors to relate macro-climate with topographic and other factors to model climate at finer scales (e.g. 25 m). For example, Slavich *et al.* (2014) compared climate variables derived from weather stations (interpolated using elevation and location only), to topoclimate variables generated using field measurements, topography and canopy cover. One hundred and twenty-seven iButtons placed 5 cm beneath the ground were used to record hourly, near-surface temperature and humidity over 340 days (Ashcroft and Gollan 2012). They were placed across environmental gradients, and with other factors (topographic exposure, relative altitude, remotely sensed canopy cover, Euclidean distance to coast) used to model climate at 25 m resolution.

This could be very effective in the NJF (quantifying how topography mediates climate for microsites) and improve our understanding of species predictions (Figure 5.5). A few other potential approaches in the NJF include correlating TWI with field measurements of soil moisture (Mohamedou *et al.* 2019; Kopecký *et al.* 2021) or streamflow (Nruthya and Srinivas 2015); fine-scale measurements of rainfall; and remotely sensed indices of vegetation health and function (e.g. Sentinel-2, 10-20 m resolution) (Frampton *et al.* 2013).

6.1.3 Biotic interactions

Capturing biotic interactions is a developing frontier for ecological modelling (Wisn *et al.* 2013). While climate and topography are essential in structuring vegetation patterns at broad-to-medium scales, biotic factors can be important at fine scales (although soil, topography and fire are considered the critical factors in Australia due to its ancient, weathered landscape (Oriens and Milewski 2007)). Van der Putten *et al.* (2010) reviewed the biotic interactions literature and argue that selection direction during climate change will vary based on how trophic interactions are disrupted (Van der Putten *et al.* 2010). This includes controls on plant abundance by aboveground and belowground multitrophic interactions with symbionts, pathogens and herbivores. Briscoe *et al.* (2019) show how process-explicit models can be used to do this, including eco-physiological and coupled-SDM-population models. Bush *et al.* (2018) used thermal extremes to adjust GDMs and (Armitage and Jones 2020) incorporated competition terms in a mechanistic niche model to predict the poleward range boundary of two aquatic plants. Improvements in model performance have also been found when the cover of dominants were used as a proxy for inter-specific interactions between plants (Roux *et al.* 2012; Le Roux *et al.* 2013; le Roux *et al.* 2014). Rooting strategies (including mycorrhizal associations) and the effects of fire and drying on population responses (survival and recruitment) are two biotic factors that warrant further investigation in the NJF (Lamont 1982; Brundrett and Abbott 1991; 1994; Enright *et al.* 2014; Brundrett 2017).

Attempting to predict biological responses to climate change is difficult. In light of this, and the matters discussed above, the realism and reliability of results in this thesis are difficult to assess. This is despite efforts to follow best-practice (Chapter 2, Chapter 3) and develop models that are relevant for practitioners (Addison *et al.* 2013; Guisan *et al.* 2013). Overall, while limited, the work provides a foundation that can be improved as additional data and modelling tools become available (Moore *et al.* 1991), and as a basis for further field-based hypothesis testing and validation.

6.2 Fire and weed management

Hydrological refugia are areas that require investment into targeted fire and weed management (Barrows *et al.* 2020). Unpublished observations by Alcoa's environmental staff and this thesis show that after fire, upland taxa are recruiting into areas previously dominated by riparian species (Figure 5.7 d-f). Fire is a contentious issue in SWA and prescribed burning regimes are yet to be optimised for biodiversity outcomes in the forest (Burrows and McCaw 2013; Ruane 2018; Howard *et al.* 2020). Community shifts in response to drying are being recorded worldwide. In Portugal, climatic extremes are affecting the evergreen and deciduous oak forests, which are declining and/or being replaced by more xeric species

(Acácio *et al.* 2017). In boreal North America rapid forest change is occurring due to stand-replacing fires that are exacerbated by drying (Stralberg *et al.* 2018; Stralberg *et al.* 2020b).

Management strategies are being implemented to try and protect refugia (Morelli *et al.* 2016; Morelli *et al.* 2020; Stralberg *et al.* 2020a). In Nevada, prescribed burning around areas of refugia for Joshua trees successfully protected them over a fire season (Barrows *et al.* 2020) and could be used as a strategy in the NJF. In cases where fire does burn hydrological refugia, post-fire monitoring should be done to track the germination of upland taxa and any shifts in species composition. Given observed rainfall reduction (-12.5% from pre-1970 averages (Hughes *et al.* 2012)) and changes in species recruitment recorded for the forest, the NJF also provides a test-system for ecological climate change hypotheses more broadly. Finally, weed invasion along riparian habitats is also of concern (Williams and Mitchell 2001). Riparian systems are vulnerable to invasion by exotics (Hancock *et al.* 1996) and flow reduction can facilitate this (Catford *et al.* 2011). Weed management of refugia will be an important part of their protection and persistence.

6.3 Improved conservation outcomes in the northern jarrah forest.

Conservation decision-making is often underpinned by area-based numeric targets. Internationally, Aichi Target 11 aims to protect 17% of terrestrial land cover and in Australia, the Commonwealth Government has committed to reserving 10% of key land-units by 2030 (Commonwealth of Australia 2010). While providing guidance, the limits of percent-based targets are well recognised (Carwardine *et al.* 2009) and efforts are being made to improve the realism of conservation prioritisation models (e.g. Runge *et al.* 2016). In SWA, the current conservation reserve system meets minimum requirements (10%) under the national CAR criteria (comprehensive, adequate and representative) (Janis 1997). The underlying principles and data that both form, and meet, these criteria are complex (Regional Forest Agreement Steering Committee 1998; Commonwealth of Australia 2010; Conservation Commission of Western Australia 2013; Kukkala and Moilanen 2013). To clarify why refugia are an important adaptive strategy for climate change management, I examine results in light of the CAR principles of SCP. Based on this, I then argue for their inclusion in SWA's formal conservation reserve network.

The principles of CAR are part of the larger system of systematic conservation planning (SCP). SCP uses biogeographic-economic analysis to identify areas for conservation protection, supporting policy decision-making under ecological and socio-political uncertainty (Margules and Pressey 2000; Kukkala and Moilanen 2013). SCP has a strong mathematical focus, using decision theory and optimisation tools to guide protected area identification (Marxan, ConsNet, C-Plan) (Moilanen *et al.* 2009). While this transitioned reserve selection from ad hoc to a quantitative, repeatable and transparent decision-making process (Pressey and Bottrill 2009), it is critiqued as over-simplifying real-world problems (Langford *et al.* 2011). The process-based description of important conservation areas identified by modelling refugia in this thesis, sits in contrast to current percent-area, CAR-defined reserves (Figure 6.1). But it is also complementary, providing a way to incorporate climate adaptation into the current reserve system.

In addition to data tools, 12 principles guide SCP at the conceptual level (Kukkala and Moilanen 2013). Within these, the CAR criteria are 'comprehensiveness' (the spatial objective), 'adequacy' the (likelihood

of species persistence in space and time) and ‘representativeness’ (the spatial solution) (Kukkala and Moilanen 2013). Again, while widely used, limitations include that the concepts are subjective (i.e. set on arbitrary target areas) and difficult to define and assess (e.g. predicting adequacy (Possingham *et al.* 2005; Wilson *et al.* 2009)).

In Australia, comprehensiveness dictates that the full range of regional ecosystems be included in the national reserve system (Thackway and Cresswell 1995; Commonwealth of Australia 2010). However, spatial data on biodiversity is limited (Wilson *et al.* 2009) and Australia lacks a consistent national, vegetation classification system based on quantitative floristics to underpin the criteria (Faber-Langendoen *et al.* 2018; Gellie *et al.* 2018). In SWA, forest ecosystem units are used to justify comprehensiveness, despite being largely based on structural forest attributes and overlooking significant floristic heterogeneity (Appendix C5). Additionally, heterogeneity is also acknowledged as not being dealt with well in the SCP literature (Possingham *et al.* 2005; Kukkala and Moilanen 2013). Given these issues and limitations, adaptive conservation policy is needed, with the flexibility to incorporate scientific developments such as refugia (Reside *et al.* 2018).

The capacity to identify refugia also informs the adequacy component of CAR (the likelihood of species persistence in space and time). Adequacy is regarded as difficult to assess and define, due to the challenges in modelling the ecological and evolutionary processes that determine persistence (Possingham *et al.* 2005; Klein *et al.* 2009; Kukkala and Moilanen 2013). The results presented in this thesis predict the likelihood of species persistence in space and time, providing a method and approach to improve our assessment of adequacy in the NJF.

Accounting for ecological processes and protecting climate refugia was identified by Groves *et al.* (2012) as two ways that SCP could be expanded to include climate change impacts. The refugia modelling done in this thesis provides an avenue to do this, synthesising complex plant-topographic-hydrological relationships across a large area of the forest. The addition of these zones to the reserve network is not large, accounting for only a 3% increase in area. Given the conceptual and methodological issues in defining CAR, and the significant drying recorded across SWA, refugia modelling provides an effective way to place the current reserve system into a climate adaptation framework.

6.4 Concluding remarks

This thesis models and maps areas of potential hydrological refugia in the northern jarrah forest of southwestern Australia, providing a basis for adaptive conservation management under drying. By performing analyses at two topographic resolutions (30 m vs 2 m) and scales of vegetation response (community and species), results offer insight into both the broad and subtle implications of drying. Hydrologically important areas for vegetation conservation include major stream tributaries and broad winter-wet valley floors, and a patchy mosaic of moisture-holding microsites in the general forest matrix. Mapping provides a basis for understanding how these zones are different from the surrounding landscape, and a framework to prioritise further research and management objectives. Finally, this thesis provides evidence for areas of conservation concern under climate change and justification for their inclusion in the current reserve network.

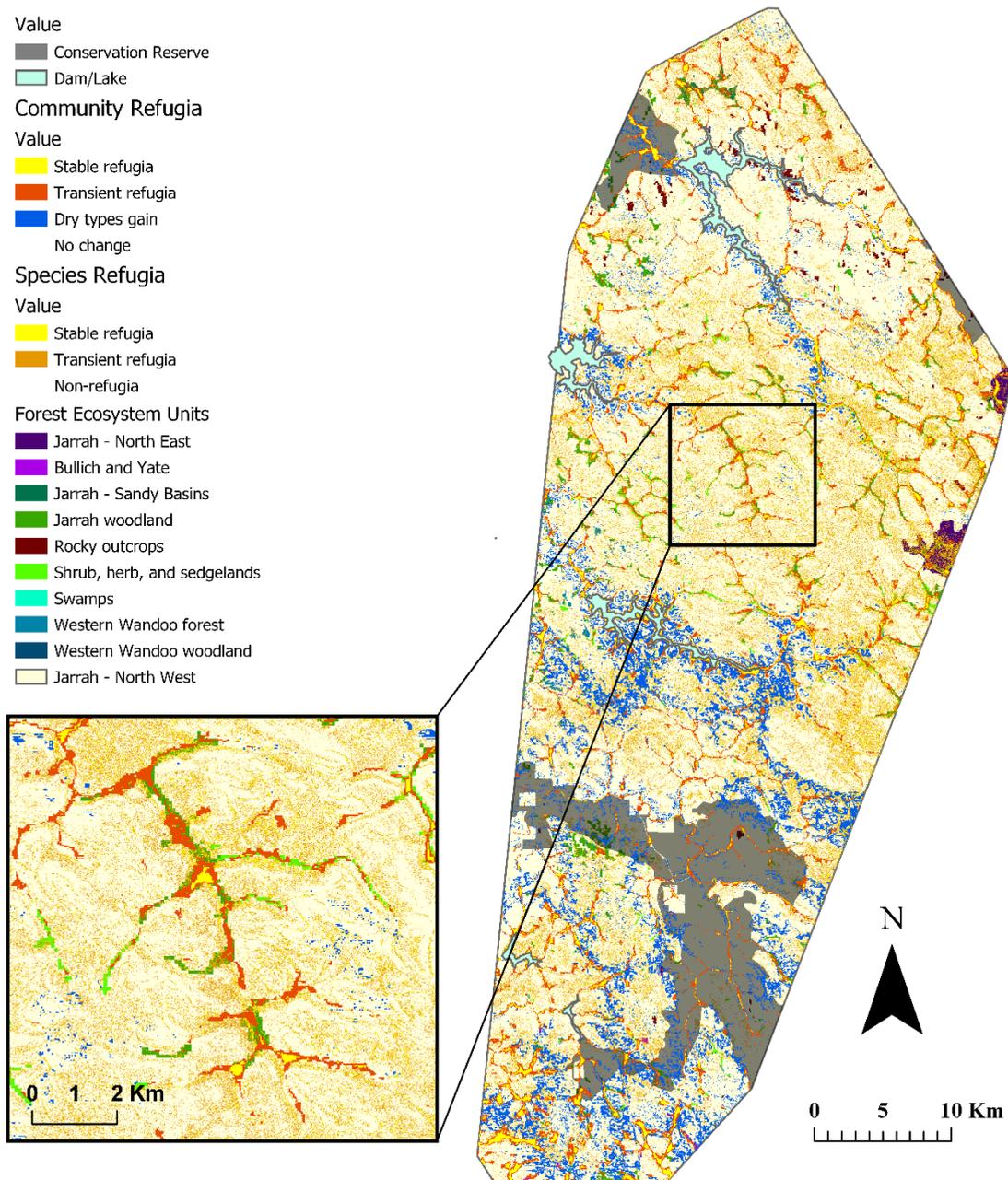


Figure 6.1 Potential hydrologic refugia at the landscape scale show distinct zones of refugia coupled with a complex pattern of variation in water availability and riparian vegetation throughout the forest matrix. This process-based description of important conservation areas sits in contrast to current reserves (dark grey), which are based on the percent-area systematic conservation planning criteria ‘CAR’ (comprehensive, adequate and representative). An increase in reserve area of only 3% is required to formally protect potential zones of refugia.

References

- Acácio, V., Dias, F. S., Catry, F. X., Rocha, M. & Moreira, F. 2017. Landscape dynamics in Mediterranean oak forests under global change: understanding the role of anthropogenic and environmental drivers across forest types. *Global change biology*, 23, 1199-1217.
- Addison, P. F. E., Rumpff, L., Bau, S. S., Carey, J. M., Chee, Y. E., Jarrad, F. C., McBride, M. F. & Burgman, M. A. 2013. Practical solutions for making models indispensable in conservation decision-making. *Diversity and Distributions*, 19, 490-502.
- Agresti, A. 2003. Building and Applying Logistic Regression Models. *Categorical Data Analysis*. John Wiley & Sons, Inc.
- Agresti, A. & Kateri, M. 2011. Categorical Data Analysis. In: LOVRIC, M. (ed.) *International Encyclopedia of Statistical Science*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Albornoz, F. E., Lambers, H., Turner, B. L., Teste, F. P. & Laliberté, E. 2016. Shifts in symbiotic associations in plants capable of forming multiple root symbioses across a long-term soil chronosequence. *Ecology and Evolution*, 6, 2368-2377.
- Allen, C. D., Macalady, A. K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., Kitzberger, T., Rigling, A., Breshears, D. D., Hogg, E. H., Gonzalez, P., Fensham, R., Zhang, Z., Castro, J., Demidova, N., Lim, J.-H., Allard, G., Running, S. W., Semerci, A. & Cobb, N. 2010. A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management*, 259, 660-684.
- Altman, D. G. 1991. *Practical statistics for medical research / Douglas G. Altman*, London, London : Chapman and Hall.
- Anderson, J. E., Kriedemann, P. E., Austin, M. P. & Farquhar, G. D. 2000. Eucalypts forming a canopy functional type in dry sclerophyll forests respond differentially to environment. *Australian Journal of Botany*, 48, 759-775.
- Anderson, M. G., Comer, P. J., Beier, P., Lawler, J. J., Schloss, C. A., Buttrick, S., Albano, C. M. & Faith, D. P. 2015. Case studies of conservation plans that incorporate geodiversity. *Conservation Biology*, 29, 680-691.
- Anderson, M. G. & Ferree, C. E. 2010. Conserving the stage: climate change and the geophysical underpinnings of species diversity. *PLoS One*, 5, e11554.
- Araújo, M. B., Anderson, R. P., Márcia Barbosa, A., Beale, C. M., Dormann, C. F., Early, R., Garcia, R. A., Guisan, A., Maiorano, L., Naimi, B., O'hara, R. B., Zimmermann, N. E. & Rahbek, C. 2019. Standards for distribution models in biodiversity assessments. *Science Advances*, 5, eaat4858.
- Araújo, M. B. & Luoto, M. 2007. The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography*, 16, 743-753.
- Araújo, M. B. & Peterson, A. T. 2012. Uses and misuses of bioclimatic envelope modeling. *Ecology*, 93, 1527-1539.
- Armitage, D. W. & Jones, S. E. 2020. Coexistence barriers confine the poleward range of a globally distributed plant. *Ecology Letters*, 23, 1838-1848.
- Ashcroft, M. B. 2006. A method for improving landscape scale temperature predictions and the implications for vegetation modelling. *Ecological Modelling*, 197, 394-404.
- Ashcroft, M. B. 2010. Identifying refugia from climate change. *Journal of Biogeography*, 37, 1407-1413.

- Ashcroft, M. B., Chisholm, L. A. & French, K. O. 2009. Climate change at the landscape scale: predicting fine-grained spatial heterogeneity in warming and potential refugia for vegetation. *Global Change Biology*, 15, 656-667.
- Ashcroft, M. B. & Gollan, J. R. 2012. Fine-resolution (25 m) topoclimatic grids of near-surface (5 cm) extreme temperatures and humidities across various habitats in a large (200 × 300 km) and diverse region. *International journal of climatology*, 32, 2134-2148.
- Ashcroft, M. B. & Gollan, J. R. 2013. Moisture, thermal inertia, and the spatial distributions of near-surface soil and air temperatures: Understanding factors that promote microrefugia. *Agricultural and Forest Meteorology*, 176, 77-89.
- Ashcroft, M. B., Gollan, J. R., Warton, D. I. & Ramp, D. 2012. A novel approach to quantify and locate potential microrefugia using topoclimate, climate stability, and isolation from the matrix. *Global Change Biology*, 18, 1866-1879.
- Austin, M. P. 2013. Vegetation and Environment: Discontinuities and Continuities. *Vegetation Ecology*. John Wiley & Sons, Ltd.
- Austin, M. P., Smith, T. M., Van Niel, K. P. & Wellington, A. B. 2009. Physiological responses and statistical models of the environmental niche: a comparative study of two co-occurring Eucalyptus species. *Journal of Ecology*, 97, 496-507.
- Australian Government 2015. Collaborative Australian Protected Area Database (CAPAD) (WA). In: ENERGY, D. O. T. E. A. (ed.).
- Bardgett, R. D. & Wardle, D. A. 2010. *Aboveground-belowground linkages: biotic interactions, ecosystem processes, and global change*, Oxford University Press.
- Barrows, C. W., Ramirez, A. R., Sweet, L. C., Morelli, T. L., Millar, C. I., Frakes, N., Rodgers, J. & Mahalovich, M. F. 2020. Validating climate-change refugia: empirical bottom-up approaches to support management actions. *Frontiers in Ecology and the Environment*, 18, 298-306.
- Bates, B. C., Hope, P., Ryan, B., Smith, I. & Charles, S. 2008. Key findings from the Indian Ocean Climate Initiative and their impact on policy development in Australia. *Climatic Change*, 89, 339-354.
- 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.
- Beaumont, L. J., Hughes, L. & Pitman, A. J. 2008. Why is the choice of future climate scenarios for species distribution modelling important? *Ecology Letters*, 11, 1135-1146.
- Beaumont, L. J., Pitman, A., Poulsen, M. & Hughes, L. 2007. Where will species go? Incorporating new advances in climate modelling into projections of species distributions. *Global Change Biology*, 13, 1368-1385.
- Belbin, L. 1987. The use of non-hierarchical allocation methods for clustering large sets of data. *Australian Computer Journal*, 19, 32-41.
- Belbin, L. 1990. *PATN: pattern analysis package*, Division of Wildlife and Ecology. CSIRO.
- Belbin, L. & McDonald, C. 1993. Comparing three classification strategies for use in ecology. *Journal of Vegetation Science*, 4, 341-348.
- Bennie, J., Hill, M. O., Baxter, R. & Huntley, B. 2006. Influence of slope and aspect on long-term vegetation change in British chalk grasslands. *Journal of ecology*, 94, 355-368.
- Beven, K., Kirkby, M., Schofield, N. & Tagg, A. 1984. Testing a physically-based flood forecasting model (TOPMODEL) for three UK catchments. *Journal of hydrology*, 69, 119-143.
- Beven, K. J. & Kirkby, M. J. 1979. A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin*, 24, 43-69.

- Blonder, B., Nogués - Bravo, D., Borregaard, M. K., Donoghue Ii, J. C., Jørgensen, P. M., Kraft, N. J. B., Lessard, J. P., Morueta - Holme, N., Sandel, B., Svenning, J. C., Violle, C., Rahbek, C. & Enquist, B. J. 2015. Linking environmental filtering and disequilibrium to biogeography with a community climate framework. *Ecology*, 96, 972-985.
- Boehner, J., Koethe, R., Conrad, O., Gross, J., Ringeler, A. & Selige, T. 2002. *Soil regionalisation by means of terrain analysis and process parameterisation*, Luxembourg.
- Böhner, J. & Selige, T. 2006. Spatial prediction of soil attributes using terrain analysis and climate regionalisation.
- Bond, W. J. & Keeley, J. E. 2005. Fire as a global 'herbivore': the ecology and evolution of flammable ecosystems. *Trends in Ecology & Evolution*, 20, 387-394.
- Borcard, D., Gillet, F. & Legendre, P. 2011. *Numerical ecology with R*, Springer Science & Business Media.
- Bourel, M. & Segura, A. M. 2018. Multiclass classification methods in ecology. *Ecological Indicators*, 85, 1012-1021.
- Bradshaw, C. J. A. 2012. Little left to lose: deforestation and forest degradation in Australia since European colonization. *Journal of Plant Ecology*, 5, 109-120.
- Bradshaw, F., Collins, P. & Mcnamara, P. 1997. Forest Mapping in the South West of Western Australia. In: DEPARTMENT OF CONSERVATION AND LAND MANAGEMENT, W. A. (ed.). Department of Conservation and Land Management, Western Australia.
- Bray, J. R. & Curtis, J. T. 1957. An ordination of the upland forest communities of southern Wisconsin. *Ecological monographs*, 27, 325-349.
- Breiman, L. 2001. Random forests. *Machine learning*, 45, 5-32.
- Briscoe, N. J., Elith, J., Salguero-Gómez, R., Lahoz-Monfort, J. J., Camac, J. S., Giljohann, K. M., Holden, M. H., Hradsky, B. A., Kearney, M. R., McMahon, S. M., Phillips, B. L., Regan, T. J., Rhodes, J. R., Vesk, P. A., Wintle, B. A., Yen, J. D. L. & Guillera-Arroita, G. 2019. Forecasting species range dynamics with process-explicit models: matching methods to applications. *Ecol Lett*, 22, 1940-1956.
- Briscoe, N. J., Kearney, M. R., Taylor, C. A. & Wintle, B. A. 2016. Unpacking the mechanisms captured by a correlative species distribution model to improve predictions of climate refugia. *Glob Chang Biol*, 22, 2425-2439.
- Brouwers, N., Matusick, G., Ruthrof, K., Lyons, T. & Hardy, G. 2013. Landscape-scale assessment of tree crown dieback following extreme drought and heat in a Mediterranean eucalypt forest ecosystem. *Landscape Ecology*, 28, 69-80.
- Bruehlheide, H., Dengler, J., Jiménez-Alfaro, B., Purschke, O., Hennekens, S. M., Chytrý, M., Pillar, V. D., Jansen, F., Kattge, J., Sandel, B., Aubin, I., Biurrun, I., Field, R., Haider, S., Jandt, U., Lenoir, J., Peet, R. K., Peyre, G., Sabatini, F. M., Schmidt, M., Schrodte, F., Winter, M., Ačić, S., Agrillo, E., Alvarez, M., Ambarlı, D., Angelini, P., Apostolova, I., Arfin Khan, M. a. S., Arnst, E., Attorre, F., Baraloto, C., Beckmann, M., Berg, C., Bergeron, Y., Bergmeier, E., Bjorkman, A. D., Bondareva, V., Borchardt, P., Botta-Dukát, Z., Boyle, B., Breen, A., Brisse, H., Byun, C., Cabido, M. R., Casella, L., Cayuela, L., Černý, T., Chepinoga, V., Csiky, J., Curran, M., Čuštěrvská, R., Dajić Stevanović, Z., De Bie, E., De Ruffray, P., De Sanctis, M., Dimopoulos, P., Dressler, S., Ejrnæs, R., El-Sheikh, M. a. E.-R. M., Enquist, B., Ewald, J., Fagúndez, J., Finckh, M., Font, X., Forey, E., Fotiadis, G., García-Mijangos, I., De Gasper, A. L., Golub, V., Gutierrez, A. G., Hatim, M. Z., He, T., Higuchi, P., Holubová, D., Hölzel, N., Homeier, J., Indreica, A., Işık Gürsoy, D., Jansen, S., Janssen, J., Jedrzejek,

- B., Jiroušek, M., Jürgens, N., Kački, Z., Kavgacı, A., Kearsley, E., Kessler, M., Knollová, I., Kolomiychuk, V., Korolyuk, A., Kozhevnikova, M., Kozub, Ł., Krstonošić, D., Köhl, H., Kühn, I., Kuzemko, A., Küzmič, F., Landucci, F., Lee, M. T., et al. 2019. sPlot – A new tool for global vegetation analyses. *Journal of Vegetation Science*, 30, 161-186.
- Brundrett, M. C. 2017. Distribution and evolution of mycorrhizal types and other specialised roots in Australia. *Biogeography of mycorrhizal symbiosis*. Springer.
- Brundrett, M. C. & Abbott, L. 1991. Roots of jarrah forest plants. I. Mycorrhizal associations of shrubs and herbaceous plants. *Australian Journal of Botany*, 39, 445-457.
- Brundrett, M. C. & Abbott, L. 1994. Mycorrhizal fungus propagules in the jarrah forest. *New Phytologist*, 127, 539-546.
- Buckley, S. 2019. *NASADEM: Creating a New NASA Digital Elevation Model and Associated Products* [Online]. NASA, EARTHDATA. Available: <https://earthdata.nasa.gov/community/community-data-system-programs/measures-projects/nasadem> [Accessed 6/2/2019 2019].
- Bunn, R., Lekberg, Y. & Zabinski, C. 2009. Arbuscular mycorrhizal fungi ameliorate temperature stress in thermophilic plants. *Ecology*, 90, 1378-1388.
- Burrows, N. & Mccaw, L. 2013. Prescribed burning in southwestern Australian forests. *Frontiers in Ecology and the Environment*, 11, e25-e34.
- Burrows, N., Ward, B., Wills, A., Williams, M. & Cranfield, R. 2019. Fine-scale temporal turnover of jarrah forest understory vegetation assemblages is independent of fire regime. *Fire Ecology*, 15, 10.
- Burrows, N. & Wardell-Johnson, G. 2003. Fire and plant interactions in forested ecosystems of south-west Western Australia. *Fire in ecosystems of south-west Western Australia: impacts and management*, 2, 225.
- Butt, N., Possingham, H., De Los Rios, C., Maggini, R., Fuller, R., Maxwell, S. & Watson, J. 2016. Challenges in assessing the vulnerability of species to climate change to inform conservation actions. *Biological Conservation*, 199, 10-15.
- Byrne, M. 2008. Evidence for multiple refugia at different time scales during Pleistocene climatic oscillations in southern Australia inferred from phylogeography. *Quaternary Science Reviews*, 27, 2576-2585.
- Byrne, M., Steane, D. A., Joseph, L., Yeates, D. K., Jordan, G. J., Crayn, D., Aplin, K., Cantrill, D. J., Cook, L. G. & Crisp, M. D. 2011. Decline of a biome: evolution, contraction, fragmentation, extinction and invasion of the Australian mesic zone biota. *Journal of Biogeography*, 38, 1635-1656.
- Capon, S. J. & Dowe, J. L. 2007. Diversity and dynamics of riparian vegetation. In: LOVETT, S. & PRICE, P. (eds.) *Principles for riparian lands management*. Canberra ACT: Land & Water Australia.
- Cartwright, J. M., Dwire, K. A., Freed, Z., Hammer, S. J., Mclaughlin, B., Misztal, L. W., Schenk, E. R., Spence, J. R., Springer, A. E. & Stevens, L. E. 2020. Oases of the future? Springs as potential hydrologic refugia in drying climates. *Frontiers in Ecology and the Environment*, 18, 245-253.
- Carwardine, J., Klein, C. J., Wilson, K. A., Pressey, R. L. & Possingham, H. P. 2009. Hitting the target and missing the point: target-based conservation planning in context. *Conservation Letters*, 2, 4-11.
- Catford, J. A., Downes, B. J., Gippel, C. J. & Vesik, P. A. 2011. Flow regulation reduces native plant cover and facilitates exotic invasion in riparian wetlands. *Journal of Applied Ecology*, 48, 432-442.
- Chave, J. 2013. The problem of pattern and scale in ecology: what have we learned in 20 years? *Ecology Letters*, 16, 4-16.

- Chiew, F. H., Peel, M., McMahon, T. & Siriwardena, L. 2006. Precipitation elasticity of streamflow in catchments across the world. IAHS Publ. 308. *Climate Variability and Change (Proceedings of the Fifth FRIEND World Conference)*. Havana, Cuba.
- Chiew, F. H. S. 2006. Estimation of rainfall elasticity of streamflow in Australia. *Hydrological Sciences Journal*, 51, 613-625.
- Churchward, H. & Mearthar, W. 1980. Landforms and soils of the Darling System, Western Australia. *Atlas of Natural Resources, Darling System, Western Australia. Department of Conservation and Environment*. Perth: Western Australia Dept. of Conservation and Environment.
- Chytrý, M., Hennekens, S. M., Jiménez-Alfaro, B., Knollová, I., Dengler, J., Jansen, F., Landucci, F., Schaminée, J. H. J., Ačić, S., Agrillo, E., Ambarlı, D., Angelini, P., Apostolova, I., Attorre, F., Berg, C., Bergmeier, E., Biurrun, I., Botta-Dukát, Z., Brisse, H., Campos, J. A., Carlón, L., Čarni, A., Casella, L., Csiky, J., Čušterevska, R., Dajić Stevanović, Z., Danihelka, J., De Bie, E., De Ruffray, P., De Sanctis, M., Dickoré, W. B., Dimopoulos, P., Dubyna, D., Dziuba, T., Ejrnæs, R., Ermakov, N., Ewald, J., Fanelli, G., Fernández-González, F., Fitzpatrick, Ú., Font, X., García-Mijangos, I., Gavilán, R. G., Golub, V., Guarino, R., Haveman, R., Indreica, A., Işık Gürsoy, D., Jandt, U., Janssen, J. a. M., Jiroušek, M., Kački, Z., Kavğacı, A., Kleikamp, M., Kolomyichuk, V., Krstivojević Ćuk, M., Krstonošić, D., Kuzemko, A., Lenoir, J., Lysenko, T., Marcenò, C., Martynenko, V., Michalcová, D., Moeslund, J. E., Onyshchenko, V., Pedashenko, H., Pérez-Haase, A., Peterka, T., Prokhorov, V., Rašomavičius, V., Rodríguez-Rojo, M. P., Rodwell, J. S., Rogova, T., Ruprecht, E., Rūsiņa, S., Seidler, G., Šibík, J., Šilc, U., Škvorc, Ž., Sopotlieva, D., Stančić, Z., Svenning, J.-C., Swacha, G., Tsiripidis, I., Turtureanu, P. D., Uğurlu, E., Uogintas, D., Valachovič, M., Vashenyak, Y., Vassilev, K., Venanzoni, R., Virtanen, R., Weekes, L., Willner, W., Wohlgemuth, T. & Yamalov, S. 2015. European Vegetation Archive (EVA): an integrated database of European vegetation plots. *Applied Vegetation Science*, n/a-n/a.
- Chytrý, M., Hennekens, S. M., Jiménez-Alfaro, B., Knollová, I., Dengler, J., Jansen, F., Landucci, F., Schaminée, J. H., Ačić, S. & Agrillo, E. 2016. European Vegetation Archive (EVA): an integrated database of European vegetation plots. *Applied vegetation science*, 19, 173-180.
- Clarke, J. M., Whetton, P. H. & Hennessy, K. J. 2011. Providing application-specific climate projections datasets: CSIRO's Climate Futures Framework *19th International Congress on Modelling and Simulation*. Perth, Australia.
- Cohen, J. 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20, 37-46.
- Commonwealth of Australia 1998. *A Regional Forest Agreement for Western Australia : comprehensive regional assessment / prepared by officials to support the Western Australian South-West Forest Regional Forest Agreement process*, Joint Commonwealth and Western Australian Regional Forest Agreement Steering Committee, Perth.
- Commonwealth of Australia 2010. Australia's Strategy for the National Reserve System 2009–2030. Australian Government.
- Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V. & Böhner, J. 2015. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geosci. Model Dev.*, 8, 1991-2007.
- Conservation Commission of Western Australia 2013. Forest Management Plan 2014-2023. Conservation Commission of Western Australia, Perth.

- Cooley, S. W. 2015. *GIS4Geomorphology* [Online]. Available: <http://www.gis4geomorphology.com> [Accessed 3 March 2017].
- Corporation, W. W. 2015. Water Forever: South West, Final Report. WA Water Corporation.
- Couronné, R., Probst, P. & Boulesteix, A.-L. 2018. Random forest versus logistic regression: a large-scale benchmark experiment. *BMC Bioinformatics*, 19, 270.
- Cox, R. L. & Underwood, E. C. 2011. The importance of conserving biodiversity outside of protected areas in mediterranean ecosystems. *PLOS ONE*, 6, e14508.
- Crawley, M. J. 2012. Multivariate Statistics. *The R Book*. John Wiley & Sons, Ltd.
- Csiro 2012. Water yields and demands in south-west Western Australia. A report to the Australian Government from the CSIRO South-West Western Australia Sustainable Yields Project.: Western Australian Government.
- Csiro & Bureau of Meteorology 2015. Climate change in Australia information for Australia's natural resource management regions: Technical report. CSIRO and Bureau of Meteorology Australia.
- Davis, J., Pavlova, A., Thompson, R. & Sunnucks, P. 2013. Evolutionary refugia and ecological refuges: key concepts for conserving Australian arid zone freshwater biodiversity under climate change. *Global change biology*, 19, 1970-1984.
- Davis, L., Thoms, M. C., Fellows, C. & Bunn, S. 2002. Physical and ecological associations in dryland refugia: waterholes of the Cooper Creek, Australia. *International Association of Hydrological Sciences, Publication*, 276, 77-84.
- Dawson, T. E. & Pate, J. S. 1996. Seasonal water uptake and movement in root systems of Australian phraeatophytic plants of dimorphic root morphology: a stable isotope investigation. *Oecologia*, 107, 13-20.
- De Cáceres, M., Chytrý, M., Agrillo, E., Attorre, F., Botta-Dukát, Z., Capelo, J., Czúcz, B., Dengler, J., Ewald, J., Faber-Langendoen, D., Feoli, E., Franklin, S. B., Gavilán, R., Gillet, F., Jansen, F., Jiménez-Alfaro, B., Krestov, P., Landucci, F., Lengyel, A., Loidi, J., Mucina, L., Peet, R. K., Roberts, D. W., Roleček, J., Schaminée, J. H. J., Schmidtlein, S., Theurillat, J.-P., Tichý, L., Walker, D. A., Wildi, O., Willner, W. & Wisser, S. K. 2015. A comparative framework for broad-scale plot-based vegetation classification. *Applied Vegetation Science*, 18, 543-560.
- De Cáceres, M. & Legendre, P. 2009. Associations between species and groups of sites: indices and statistical inference. *Ecology*, 90, 3566-3574.
- De Cáceres, M., Legendre, P. & Moretti, M. 2010. Improving indicator species analysis by combining groups of sites. *Oikos*, 119, 1674-1684.
- De Deyn, G. B. & Van Der Putten, W. H. 2005. Linking aboveground and belowground diversity. *Trends in Ecology & Evolution*, 20, 625-633.
- Dekker, S. C., Rietkerk, M. & Bierkens, M. F. 2007. Coupling microscale vegetation–soil water and macroscale vegetation–precipitation feedbacks in semiarid ecosystems. *Global change biology*, 13, 671-678.
- Dell, B., Havel, J. & Malajczuk, N. 1989. *The jarrah forest: a complex Mediterranean ecosystem*, Springer Science & Business Media.
- Devia, G. K., Ganasri, B. P. & Dwarakish, G. S. 2015. A Review on Hydrological Models. *Aquatic Procedia*, 4, 1001-1007.
- Di Marco, M., Harwood, T. D., Hoskins, A. J., Ware, C., Hill, S. L. & Ferrier, S. 2019. Projecting impacts of global climate and land-use scenarios on plant biodiversity using compositional-turnover modelling. *Global change biology*, 25, 2763-2778.
- Di Virgilio, G., Laffan, S. W. & Ebach, M. C. 2012. Fine-scale quantification of floral and faunal breaks and their geographic correlates, with an example from south-eastern Australia. *Journal of Biogeography*, 39, 1862-1876.

- Dillon, S., Mcevoy, R., Baldwin, D. S., Southerton, S., Campbell, C., Parsons, Y. & Rees, G. N. 2015. Genetic diversity of *Eucalyptus camaldulensis* Dehnh. following population decline in response to drought and altered hydrological regime. *Austral Ecology*, 40, 558-572.
- Doerr, V. A., Barrett, T. & Doerr, E. D. 2011. Connectivity, dispersal behaviour and conservation under climate change: a response to Hodgson et al. *Journal of Applied Ecology*, 48, 143-147.
- Dufrêne, M. & Legendre, P. 1997. Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecological monographs*, 67, 345-366.
- Elith, J., Graham, C., Valavi, R., Abegg, M., Bruce, C., Ford, A., Guisan, A., Hijmans, R. J., Huettmann, F., Lohmann, L., Loiselle, B., Moritz, C., Overton, J., Peterson, A. T., Phillips, S., Richardson, K., Williams, S., Wiser, S. K., Wohlgemuth, T. & Zimmermann, N. E. 2020. Presence-only and Presence-absence Data for Comparing Species Distribution Modeling Methods. *Biodiversity informatics*, 15, 69-80.
- Elith, J. & Graham, C. H. 2009. Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography*, 32, 66-77.
- Elith, J., H. Graham, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann, F., R. Leathwick, J., Lehmann, A., Li, J., G. Lohmann, L., A. Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Mcc. M. Overton, J., Townsend Peterson, A., J. Phillips, S., Richardson, K., Scachetti-Pereira, R., E. Schapire, R., Soberón, J., Williams, S., S. Wisz, M. & E. Zimmermann, N. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29, 129-151.
- Elith, J. & Leathwick, J. R. 2009. Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, 40, 677.
- Elith, J., Leathwick, J. R. & Hastie, T. 2008. A working guide to boosted regression trees. *Journal of Animal Ecology*, 77, 802-813.
- Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E. & Yates, C. J. 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, 17, 43-57.
- Engemann, K., Sandel, B., Enquist, B. J., Jørgensen, P. M., Kraft, N., Marcuse-Kubitza, A., McGill, B., Morueta-Holme, N., Peet, R. K., Violle, C., Wiser, S. & Svenning, J. C. 2016. Patterns and drivers of plant functional group dominance across the Western Hemisphere: a macroecological re-assessment based on a massive botanical dataset. *Botanical Journal of the Linnean Society*, 180, 141-160.
- Enquist, B., Condit, R., Peet, R., Schildhauer, M. & Thiers, B. 2009. The Botanical Information and Ecology Network (BIEN): Cyberinfrastructure for an integrated botanical information network to investigate the ecological impacts of global climate change on plant biodiversity. *Salt Lake City, UT: iPlant Collaborative*.
- Enquist, B. J., Condit, R., Peet, R. K., Schildhauer, M. & Thiers, B. M. 2016. Cyberinfrastructure for an integrated botanical information network to investigate the ecological impacts of global climate change on plant biodiversity. PeerJ Preprints.
- Enright, N., Fontaine, J., Westcott, V., Lade, J. & Miller, B. 2011. Fire interval effects on persistence of resprouter species in Mediterranean-type shrublands. *Plant Ecology*, 1-13.
- Enright, N. J., Fontaine, J. B., Bowman, D. M., Bradstock, R. A. & Williams, R. J. 2015. Interval squeeze: altered fire regimes and demographic responses interact

- to threaten woody species persistence as climate changes. *Frontiers in Ecology and the Environment*, 13, 265-272.
- Enright, N. J., Fontaine, J. B., Lamont, B. B., Miller, B. P. & Westcott, V. C. 2014. Resistance and resilience to changing climate and fire regime depend on plant functional traits. *Journal of Ecology*, 102, 1572-1581.
- Faber-Langendoen, D., Baldwin, K., Peet, R. K., Meidinger, D., Muldavin, E., Keeler-Wolf, T. & Josse, C. 2018. The EcoVeg approach in the Americas: US, Canadian and international vegetation classifications. *Phytocoenologia*, 48, 215-237.
- Faber-Langendoen, D., Keeler-Wolf, T., Meidinger, D., Tart, D., Hoagland, B., Josse, C., Navarro, G., Ponomarenko, S., Saucier, J.-P., Weakley, A. & Comer, P. 2014. EcoVeg: a new approach to vegetation description and classification. *Ecological Monographs*, 84, 533-561.
- Feng, D. & Beighley, E. 2020. Identifying uncertainties in hydrologic fluxes and seasonality from hydrologic model components for climate change impact assessments. *Hydrol. Earth Syst. Sci.*, 24, 2253-2267.
- Ferrier, S. & Guisan, A. 2006. Spatial modelling of biodiversity at the community level. *Journal of applied ecology*, 43, 393-404.
- Ferrier, S., Manion, G., Elith, J. & Richardson, K. 2007. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and distributions*, 13, 252-264.
- Fitzpatrick, M. C., Gove, A. D., Sanders, N. J. & Dunn, R. R. 2008. Climate change, plant migration, and range collapse in a global biodiversity hotspot: the Banksia (Proteaceae) of Western Australia. *Global Change Biology*, 14, 1337-1352.
- Foody, G. M. 1999. Applications of the self-organising feature map neural network in community data analysis. *Ecological Modelling*, 120, 97-107.
- Forester, B. R., Dechaine, E. G. & Bunn, A. G. 2013. Integrating ensemble species distribution modelling and statistical phylogeography to inform projections of climate change impacts on species distributions. *Diversity and Distributions*, 19, 1480-1495.
- Forrister, D. L., Endara, M.-J., Younkin, G. C., Coley, P. D. & Kursar, T. A. 2019. Herbivores as drivers of negative density dependence in tropical forest saplings. *Science*, 363, 1213-1216.
- Frampton, W. J., Dash, J., Watmough, G. & Milton, E. J. 2013. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS journal of photogrammetry and remote sensing*, 82, 83-92.
- Franklin, J., Serra-Diaz, J. M., Syphard, A. D. & Regan, H. M. 2016. Big data for forecasting the impacts of global change on plant communities. *Global Ecology and Biogeography*.
- Fu, P. & Rich, P. M. 2002. A geometric solar radiation model with applications in agriculture and forestry. *Computers and Electronics in Agriculture*, 37, 25-35.
- Gavin, D. G., Fitzpatrick, M. C., Gugger, P. F., Heath, K. D., Rodríguez-Sánchez, F., Dobrowski, S. Z., Hampe, A., Hu, F. S., Ashcroft, M. B. & Bartlein, P. J. 2014. Climate refugia: joint inference from fossil records, species distribution models and phylogeography. *New Phytologist*, 204, 37-54.
- Gellie, N. J. H., Hunter, J. T., Benson, J. S., Kirkpatrick, J. B., Cheal, D. C., McCreery, K. & Brocklehurst, P. 2018. Overview of plot-based vegetation classification approaches within Australia. *Phytocoenologia*, 48, 251 - 272.
- Gemitzi, A., Ajami, H. & Richnow, H.-H. 2017. Developing empirical monthly groundwater recharge equations based on modeling and remote sensing data – Modeling future groundwater recharge to predict potential climate change impacts. *Journal of Hydrology*, 546, 1-13.

- Gentili, J. 1989. Climate of the jarrah forest. *In*: DELL, B., HAVEL, J. J. & MALAJCZUK, N. (eds.) *The Jarrah Forest, A complex mediterranean ecosystem*. Dordrecht: Kluwer Academic Publishers.
- Gioia, P. & Hopper, S. D. 2017. A new phytogeographic map for the Southwest Australian Floristic Region after an exceptional decade of collection and discovery. *Botanical Journal of the Linnean Society*, 184, 1-15.
- Goldsmith, G. R., Morueta-Holme, N., Sandel, B., Fitz, E. D., Fitz, S. D., Boyle, B., Casler, N., Engemann, K., Jørgensen, P. M. & Kraft, N. J. 2016. Plant-O-Matic: a dynamic and mobile guide to all plants of the Americas. *Methods in Ecology and Evolution*.
- Gollan, J. R., Ashcroft, M. B. & Ramp, D. 2013. Fine-grained climate data alters the interpretation of a trait-based cline. *Ecosphere*, 4.
- Gómez, A. & Lunt, D. H. 2007. Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. *Phylogeography of southern European refugia*. Springer.
- Gray, M. 2004. *Geodiversity: valuing and conserving abiotic nature*, John Wiley & Sons.
- Groom, P. K. 2002. Seedling water stress response of two sandplain Banksia species differing in ability to tolerate drought. *Journal of Mediterranean Ecology*, 3, 3-10.
- Groom, P. K. 2004. Seedling growth and physiological responses of two sandplain Banksia species differing in flood tolerance. *Journal of the Royal Society of Western Australia*, 87, 115-121.
- Groves, C. & Game, E. 2016. *Conservation Planning: Informed Decisions for a Healthier Planet*. Greenwood Village, CO: Roberts & Company.
- Groves, C. R., Game, E. T., Anderson, M. G., Cross, M., Enquist, C., Ferdana, Z., Girvetz, E., Gondor, A., Hall, K. R. & Higgins, J. 2012. Incorporating climate change into systematic conservation planning. *Biodiversity and Conservation*, 21, 1651-1671.
- Grundy, M., Rossel, R. V., Searle, R., Wilson, P., Chen, C. & Gregory, L. 2015. Soil and landscape grid of Australia. *Soil Research*, 53, 835-844.
- Guillera-Arroita, G., Lahoz-Monfort, J. J. & Elith, J. 2014. Maxent is not a presence-absence method: a comment on Thibaud et al. *Methods in Ecology and Evolution*, 5, 1192-1197.
- Guillera-Arroita, G., Lahoz-Monfort, J. J., Elith, J., Gordon, A., Kujala, H., Lentini, P. E., Mccarthy, M. A., Tingley, R. & Wintle, B. A. 2015. Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography*, 24, 276-292.
- Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I., Regan, T. J., Brotons, L., Mcdonald-Madden, E. & Mantyka-Pringle, C. 2013. Predicting species distributions for conservation decisions. *Ecology letters*, 16, 1424-1435.
- Guisan, A. & Zimmermann, N. E. 2000. Predictive habitat distribution models in ecology. *Ecological modelling*, 135, 147-186.
- Gülagiz, F. K. & Sahin, S. 2017. Comparison of hierarchical and non-hierarchical clustering algorithms. *International Journal of Computer Engineering and Information Technology*, 9, 6.
- Hallgren, W., Beaumont, L., Bowness, A., Chambers, L., Graham, E., Holewa, H., Laffan, S., Mackey, B., Nix, H., Price, J., Vanderwal, J., Warren, R. & Weis, G. 2016. The Biodiversity and Climate Change Virtual Laboratory: Where ecology meets big data. *Environmental Modelling & Software*, 76, 182-186.

- Hampton, S. E., Strasser, C. A., Tewksbury, J. J., Gram, W. K., Budden, A. E., Batcheller, A. L., Duke, C. S. & Porter, J. H. 2013. Big data and the future of ecology. *Frontiers in Ecology and the Environment*, 11, 156-162.
- Hancock, C., Ladd, P. & Froend, R. 1996. Biodiversity and management of riparian vegetation in Western Australia. *Forest Ecology and Management*, 85, 239-250.
- Hannah, L., Flint, L., Syphard, A. D., Moritz, M. A., Buckley, L. B. & McCullough, I. M. 2014. Fine-grain modeling of species' response to climate change: holdouts, stepping-stones, and microrefugia. *Trends Ecol Evol*, 29, 390-397.
- Harpole, W. S. & Tilman, D. 2007. Grassland species loss resulting from reduced niche dimension. *Nature*, 446, 791-3.
- Harrell, F. E. 2019. Package 'Hmisc'. CRAN2018.
- Harris, R. M., Beaumont, L. J., Vance, T. R., Tozer, C. R., Remenyi, T. A., Perkins-Kirkpatrick, S. E., Mitchell, P. J., Nicotra, A., Mcgregor, S. & Andrew, N. 2018. Biological responses to the press and pulse of climate trends and extreme events. *Nature Climate Change*, 8, 579.
- Harris, T. 2018. *Crown Reserve dataset (LGATE-227)* [Online]. Available: <https://catalogue.data.wa.gov.au/dataset/reserves> [Accessed 5 November 2018].
- Harvey, E., Gounand, I., Ward, C. L. & Altermatt, F. 2017. Bridging ecology and conservation: from ecological networks to ecosystem function. *Journal of Applied Ecology*, 54, 371-379.
- Harwood, T. D., Donohue, R. J., Williams, K. J., Ferrier, S., Mcvicar, T. R., Newell, G. & White, M. 2016. Habitat Condition Assessment System: a new way to assess the condition of natural habitats for terrestrial biodiversity across whole regions using remote sensing data. *Methods in Ecology and Evolution*, 7, 1050-1059.
- Havel, J. 1975a. Site vegetation mapping in the northern jarrah forest (Darling Range): I. Definition of site-vegetation types.: Forest Department Western Australia.
- Havel, J. 1975b. Site vegetation mapping in the northern jarrah forest (Darling Range): II. Location and mapping of site-vegetation types. . Forest Department Western Australia.
- Havel, J. & Matiske, E. 2000. Vegetation mapping of south west forest regions of Western Australia. CALM Science Supplement.
- Havel, J. J. 2000. *Ecology of the forests of south western Australia in relation to climate and landforms*. Murdoch University.
- Hedde, E., Loneragan, O. & Havel, J. 1980. Vegetation complexes of the Darling system, Western Australia. *Department of Conservation and Environment, Atlas of Natural Resources, Darling System, Western Australia*.
- Hijmans, R. J. & Elith, J. 2015. Species distribution modeling with R.
- Hijmans, R. J. & Elith, J. 2019. Spatial Distribution Models.
- Hijmans, R. J., Phillips, S., Leathwick, J., Elith, J. & Hijmans, M. R. J. 2017. Package 'dismo'. *Circles*, 9, 1-68.
- Hijmans, R. J. & Van Etten, J. 2012. Geographic analysis and modeling with raster data.
- Hijmans, R. J., Van Etten, J., Mattiuzzi, M., Sumner, M., Greenberg, J., Lamigueiro, O., Bevan, A., Racine, E. & Shortridge, A. 2013. Raster package in R.
- Hjort, J., Gordon, J. E., Gray, M. & Hunter, M. L. 2015. Why geodiversity matters in valuing nature's stage. *Conservation Biology*, 29, 630-639.
- Hlavac, M. 2015. stargazer: Well-Formatted Regression and Summary Statistics Tables. R package version 5.2. .
- Hobbs, R. J., Higgs, E. & Harris, J. A. 2009. Novel ecosystems: implications for conservation and restoration. *Trends in ecology & evolution*, 24, 599-605.
- Hodgson, J. A., Thomas, C. D., Wintle, B. A. & Moilanen, A. 2009. Climate change, connectivity and conservation decision making: back to basics. *Journal of Applied Ecology*, 46, 964-969.

- Hoekstra, J. M., Boucher, T. M., Ricketts, T. H. & Roberts, C. 2005. Confronting a biome crisis: global disparities of habitat loss and protection. *Ecology Letters*, 8, 23-29.
- Hopper, S., Brown, A. & Marchant, N. 1997. Plants of Western Australian granite outcrops. *Journal of the Royal Society of Western Australia*, 80, 141-158.
- Hopper, S. D. 1979. Biogeographical Aspects of Speciation in the Southwest Australian Flora. *Annual Review of Ecology and Systematics*, 10, 399-422.
- Hopper, S. D. 2009. OCBIL theory: towards an integrated understanding of the evolution, ecology and conservation of biodiversity on old, climatically buffered, infertile landscapes. *Plant and Soil*, 322, 49-86.
- Hopper, S. D. & Gioia, P. 2004. The Southwest Australian Floristic Region: Evolution and Conservation of a Global Hot Spot of Biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, 35, 623-650.
- Hopper, S. D., Silveira, F. A. & Fiedler, P. L. 2015. Biodiversity hotspots and Ocbil theory. *Plant and Soil*, 1-50.
- Hothorn, T., Zeileis, A., Farebrother, R. W., Cummins, C., Millo, G., Mitchell, D. & Zeileis, M. A. 2019. Package 'lmtest'.
- House, A., Hilbert, D., Ferrier, S., Martin, T., Dunlop, M., Harwood, T., Williams, K., Fletcher, C., Murphy, H. & Gobbett, D. 2012. The implications of climate change for biodiversity conservation and the National Reserve System: sclerophyll forests of southeastern Australia. CSIRO Climate Adaptation Flagship Working Paper.
- Howard, T., Burrows, N., Smith, T., Daniel, G. & Mccaw, L. 2020. A framework for prioritising prescribed burning on public land in Western Australia. *International journal of wildland fire*, 29, 314-325.
- Hughes, J., Petrone, K. & Silberstein, R. 2012. Drought, groundwater storage and stream flow decline in southwestern Australia. *Geophysical Research Letters*, 39.
- Iwahashi, J. & Pike, R. J. 2007. Automated classifications of topography from DEMs by an unsupervised nested-means algorithm and a three-part geometric signature. *Geomorphology*, 86, 409-440.
- Jablonski, D. 2008. Biotic interactions and macroevolution: extensions and mismatches across scales and levels. *Evolution*, 62, 715-739.
- Janis 1997. Nationally agreed criteria for the establishment of a comprehensive, adequate and representative reserve system for forests in Australia. A Joint ANZECC/MCFFA National Forest Policy Statement Implementation Subcommittee (JANIS) report.
- Kassambara, A. 2017. *Practical guide to cluster analysis in R: Unsupervised machine learning*, Sthda.
- Kaufman, L. & Rousseeuw, P. J. 2008. Fuzzy Analysis (Program FANNY). *Finding Groups in Data*. John Wiley & Sons, Inc.
- Kaufman, L. & Rousseeuw, P. J. 2009. *Finding groups in data: an introduction to cluster analysis*, John Wiley & Sons.
- Keppel, G. & Kavousi, J. 2015. Effective climate change refugia for coral reefs. *Global Change Biology*, n/a-n/a.
- Keppel, G., Mokany, K., Wardell-Johnson, G. W., Phillips, B. L., Welbergen, J. A. & Reside, A. E. 2015. The capacity of refugia for conservation planning under climate change. *Frontiers in Ecology and the Environment*, 13, 106-112.
- Keppel, G., Robinson, T. P., Wardell-Johnson, G. W., Yates, C. J., Van Niel, K. P., Byrne, M. & Schut, A. G. 2017. A low-altitude mountain range as an important refugium for two narrow endemics in the Southwest Australian Floristic Region biodiversity hotspot. *Annals of Botany*, 119, 289-300.

- Keppel, G., Van Niel, K. P., Wardell-Johnson, G. W., Yates, C. J., Byrne, M., Mucina, L., Schut, A. G., Hopper, S. D. & Franklin, S. E. 2012. Refugia: identifying and understanding safe havens for biodiversity under climate change. *Global Ecology and Biogeography*, 21, 393-404.
- Keppel, G. & Wardell-Johnson, G. W. 2012. Refugia: keys to climate change management. *Global Change Biology*, 18, 2389-2391.
- Keppel, G. & Wardell-Johnson, G. W. 2015. Refugial capacity defines holdouts, microrefugia and stepping-stones: a response to Hannah et al. *Trends in ecology & evolution*, 20, 1-2.
- Kikvidze, Z., Pugnaire, F. I., Brooker, R. W., Choler, P., Lortie, C. J., Michalet, R. & Callaway, R. M. 2005. LINKING PATTERNS AND PROCESSES IN ALPINE PLANT COMMUNITIES: A GLOBAL STUDY. *Ecology*, 86, 1395-1400.
- Klausmeyer, K. R. & Shaw, M. R. 2009. Climate Change, Habitat Loss, Protected Areas and the Climate Adaptation Potential of Species in Mediterranean Ecosystems Worldwide (Mediterranean Climate Change). *PLoS ONE*, 4, e6392.
- Klein, C., Wilson, K., Watts, M., Stein, J., Berry, S., Carwardine, J., Smith, M. S., Mackey, B. & Possingham, H. 2009. Incorporating ecological and evolutionary processes into continental-scale conservation planning. *Ecological applications*, 19, 206-217.
- Kopecký, M. & Čížková, Š. 2010. Using topographic wetness index in vegetation ecology: does the algorithm matter? *Applied Vegetation Science*, 13, 450-459.
- Kopecký, M., Macek, M. & Wild, J. 2021. Topographic Wetness Index calculation guidelines based on measured soil moisture and plant species composition. *Science of The Total Environment*, 757, 143785.
- Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, O., Mayer, Z., Kenkel, B., Benesty, M., Lescarbeau, R., Ziem, A., Scrucca, L., Tang, Y., Candan, C. & Hunt, T. 2017. caret: Classification and Regression Training. R package. version 6.0-78. ed.
- Kujala, H., Lahoz-Monfort, J. J., Elith, J. & Moilanen, A. 2018. Not all data are equal: Influence of data type and amount in spatial conservation prioritisation. *Methods in Ecology and Evolution*, 9, 2249-2261.
- Kukkala, A. S. & Moilanen, A. 2013. Core concepts of spatial prioritisation in systematic conservation planning. *Biological Reviews*, 88, 443-464.
- Laing, I. & Hauck, E. 1997. Water harvesting from granite outcrops in Western Australia. *Journal of the Royal Society of Western Australia*, 80, 181-184.
- Laliberté, E., Zemunik, G. & Turner, B. L. 2014. Environmental filtering explains variation in plant diversity along resource gradients. *Science*, 345, 1602-1605.
- Lamont, B. 1982. Mechanisms for enhancing nutrient uptake in plants, with particular reference to mediterranean South Africa and Western Australia. *The Botanical Review*, 48, 597-689.
- Langford, W. T., Gordon, A., Bastin, L., Bekessy, S. A., White, M. D. & Newell, G. 2011. Raising the bar for systematic conservation planning. *Trends in Ecology & Evolution*, 26, 634-640.
- Lavorel, S. 1999. Ecological diversity and resilience of Mediterranean vegetation to disturbance. *Diversity and Distributions*, 5, 3-13.
- Lawley, V., Lewis, M., Clarke, K. & Ostendorf, B. 2016. Site-based and remote sensing methods for monitoring indicators of vegetation condition: An Australian review. *Ecological Indicators*, 60, 1273-1283.
- Le Roux, P. C., Lenoir, J., Pellissier, L., Wisz, M. S. & Luoto, M. 2013. Horizontal, but not vertical, biotic interactions affect fine-scale plant distribution patterns in a low-energy system. *Ecology*, 94, 671-682.

- Le Roux, P. C., Pellissier, L., Wisz, M. S. & Luoto, M. 2014. Incorporating dominant species as proxies for biotic interactions strengthens plant community models. *Journal of Ecology*, 102, 767-775.
- Lehtomäki, J. & Moilanen, A. 2013. Methods and workflow for spatial conservation prioritization using Zonation. *Environmental Modelling & Software*, 47, 128-137.
- Leipold, M., Tausch, S., Poschlod, P. & Reisch, C. 2017. Species distribution modeling and molecular markers suggest longitudinal range shifts and cryptic northern refugia of the typical calcareous grassland species *Hippocrepis comosa* (horseshoe vetch). *Ecology and Evolution*, 7, 1919-1935.
- Levin, S. A. 1992. The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. *Ecology*, 73, 1943-1967.
- Lindgren, M., Dakos, V., Gröger, J. P., Gårdmark, A., Kornilovs, G., Otto, S. A. & Möllmann, C. 2012. Early detection of ecosystem regime shifts: a multiple method evaluation for management application. *PloS one*, 7, e38410-e38410.
- Lippitt, C. D., Rogan, J., Toledano, J., Sangermano, F., Eastman, J. R., Mastro, V. & Sawyer, A. 2008. Incorporating anthropogenic variables into a species distribution model to map gypsy moth risk. *ecological modelling*, 210, 339-350.
- Liu, B., Chen, H. Y. H. & Yang, J. 2018. Understory Community Assembly Following Wildfire in Boreal Forests: Shift From Stochasticity to Competitive Exclusion and Environmental Filtering. *Frontiers in Plant Science*, 9.
- Lötter, M., Mucina, L. & Witkowski, E. 2013. The classification conundrum: species fidelity as leading criterion in search of a rigorous method to classify a complex forest data set. *Community Ecology*, 14, 121-132.
- Lubomír, T., Milan, C. & Petr, Š. 2011. Evaluating the stability of the classification of community data. *Ecography (Copenhagen)*, 34, 807-813.
- Ma, J., Lin, G., Chen, J. & Yang, L. An improved topographic wetness index considering topographic position. 2010 18th International Conference on Geoinformatics, 2010. IEEE, 1-4.
- Mackey, B., Berry, S., Hugh, S., Ferrier, S., Harwood, T. D. & Williams, K. J. 2012. Ecosystem greenspots: identifying potential drought, fire, and climate-change micro-refuges. *Ecological Applications*, 22, 1852-1864.
- Macqueen, J. 1967. Some methods for classification and analysis of multivariate observations. *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, 1, 281-297.
- Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. & Hornik, K. 2012. Cluster: cluster analysis basics and extensions. *R package version*, 1, 56.
- Margules, C. R. & Pressey, R. L. 2000. Systematic conservation planning. *Nature*, 405, 243-253.
- Mattiske, E. 2012. Flora and Vegetation Assessment of Keats.
- Mattiske, E. & Havel, J. 1998. Vegetation Complexes of the South-west Forest Region of Western Australia. *Maps and report prepared as part of the Regional Forest Agreement, Western Australia for the Department of Conservation and Land Management and Environment Australia.*
- Matusick, G., Ruthrof, K., Brouwers, N., Dell, B. & Hardy, G. 2013. Sudden forest canopy collapse corresponding with extreme drought and heat in a mediterranean-type eucalypt forest in southwestern Australia. *European Journal of Forest Research*, 132, 497-510.
- Mazzoleni, S., Di Pasquale, G., Mulligan, M., Di Martino, P. & Rego, F. 2004. *Recent dynamics of the Mediterranean vegetation and landscape*, John Wiley & Sons.
- Mcdonald, J. H. 2008. *Handbook of Biological Statistics*, Baltimore, Sparky House.

- Mcdowell, N., Pockman, W. T., Allen, C. D., Breshears, D. D., Cobb, N., Kolb, T., Plaut, J., Sperry, J., West, A., Williams, D. G. & Yezzer, E. A. 2008. Mechanisms of plant survival and mortality during drought: why do some plants survive while others succumb to drought? Oxford, UK.
- Mcguire, J. L., Lawler, J. J., Mcrae, B. H., Nuñez, T. A. & Theobald, D. M. 2016. Achieving climate connectivity in a fragmented landscape. *Proceedings of the National Academy of Sciences*, 113, 7195-7200.
- Mcintosh, E. J., Pressey, R. L., Lloyd, S., Smith, R. J. & Grenyer, R. 2017. The impact of systematic conservation planning. *Annual Review of Environment and Resources*, 42, 677-697.
- Mckenzie, N. 2008. *Guidelines for surveying soil and land resources / N.J. McKenzie ... [et al.]*, Collingwood, Vic., Collingwood, Vic. : CSIRO Pub.
- Mckenzie, N., Hopper, S., Wardell-Johnson, G. & Gibson, N. 1996. Assessing the conservation reserve system in the Jarrah Forest Bioregion. *Journal of the Royal Society of Western Australia*, 79, 241-248.
- Mclaughlin, B. C., Ackerly, D. D., Klos, P. Z., Natali, J., Dawson, T. E. & Thompson, S. E. 2017. Hydrologic refugia, plants, and climate change. *Global Change Biology*, 23, 2941-2961.
- Médail, F. & Diadema, K. 2009. Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of Biogeography*, 36, 1333-1345.
- Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A. & Leisch, F. 2017. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. Version 1.6-8.
- Michalak, J. L., Stralberg, D., Cartwright, J. M. & Lawler, J. J. 2020. Combining physical and species-based approaches improves refugia identification. *Frontiers in Ecology and the Environment*, 18, 254-260.
- Miller, J. A. & Franklin, J. 2010. Incorporating Spatial Autocorrelation in Species Distribution Models. In: FISCHER, M. M. & GETIS, A. (eds.) *Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Mod, H. K., Heikkinen, R. K., Roux, P. C., Wisz, M. S. & Luoto, M. 2016. Impact of biotic interactions on biodiversity varies across a landscape. *Journal of Biogeography*, 43, 2412-2423.
- Mohamedou, C., Korhonen, L., Eerikäinen, K. & Tokola, T. 2019. Using LiDAR-modified topographic wetness index, terrain attributes with leaf area index to improve a single-tree growth model in south-eastern Finland. *Forestry: An International Journal of Forest Research*, 92, 253-263.
- Moilanen, A., Wilson, K. A. & Possingham, H. 2009. *Spatial conservation prioritization: quantitative methods and computational tools*, Oxford University Press.
- Mokany, K., Bush, A. & Ferrier, S. 2019. Community assembly processes restrict the capacity for genetic adaptation under climate change. *Ecography*, 42, 1164-1174.
- Mokany, K., Jordan, G. J., Harwood, T. D., Harrison, P. A., Keppel, G., Gilfedder, L., Carter, O. & Ferrier, S. 2016. Past, present and future refugia for Tasmania's palaeoendemic flora. *Journal of Biogeography*.
- Molloy, S. W., Davis, R. A. & Van Etten, E. J. 2016. An evaluation and comparison of spatial modelling applications for the management of biodiversity: a case study on the fragmented landscapes of south-western Australia. *Pacific Conservation Biology*, 22, 338-349.

- Moore, I. D., Grayson, R. & Ladson, A. 1991. Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. *Hydrological processes*, 5, 3-30.
- Morales-Castilla, I., Matias, M. G., Gravel, D. & Araújo, M. B. 2015. Inferring biotic interactions from proxies. *Trends in Ecology & Evolution*.
- Moreira, F., Allsopp, N., Esler, K. J., Wardell-Johnson, G., Ancillotto, L., Arianoutsou, M., Clary, J., Brotons, L., Clavero, M. & Dimitrakopoulos, P. G. 2019. Priority questions for biodiversity conservation in the Mediterranean biome: Heterogeneous perspectives across continents and stakeholders. *Conservation Science and Practice*, 1.
- Morelli, T. L., Barrows, C. W., Ramirez, A. R., Cartwright, J. M., Ackerly, D. D., Eaves, T. D., Ebersole, J. L., Krawchuk, M. A., Letcher, B. H., Mahalovich, M. F., Meigs, G. W., Michalak, J. L., Millar, C. I., Quiñones, R. M., Stralberg, D. & Thorne, J. H. 2020. Climate-change refugia: biodiversity in the slow lane. *Frontiers in Ecology and the Environment*, 18, 228-234.
- Morelli, T. L., Daly, C., Dobrowski, S. Z., Dulen, D. M., Ebersole, J. L., Jackson, S. T., Lundquist, J. D., Millar, C. I., Maher, S. P., Monahan, W. B., Nydick, K. R., Redmond, K. T., Sawyer, S. C., Stock, S. & Beissinger, S. R. 2016. Managing Climate Change Refugia for Climate Adaptation. *PLOS ONE*, 11, e0159909.
- Morton, S. R., Short, J. & Barker, R. D. 1995. Refugia for biological diversity in arid and semi-arid Australia.
- Morueta-Holme, N., Blonder, B., Sandel, B., McGill, B. J., Peet, R. K., Ott, J. E., Violle, C., Enquist, B. J., Jørgensen, P. M. & Svenning, J. C. 2016. A network approach for inferring species associations from co-occurrence data. *Ecography*, 39, 1139-1150.
- Mucina, L. 2019. Biome: evolution of a crucial ecological and biogeographical concept. *New Phytologist*, 222, 97-114.
- Mucina, L. & Wardell-Johnson, G. 2011. Landscape age and soil fertility, climate stability, and fire regime predictability: beyond the OBCIL framework. *Plant and Soil*, 341, 1-23.
- Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A. & Kent, J. 2000. Biodiversity hotspots for conservation priorities. *Nature*, 403, 853-858.
- Naiman, R. J. & Decamps, H. 1997. The Ecology of Interfaces: Riparian Zones. *Annual Review of Ecology and Systematics*, 28, 621-658.
- Nakazawa, M. 2014. fmsb: Functions for medical statistics book with some demographic data. *R package version 0.5*, 2.
- Nasa 2011. One Second Shuttle Radar Topography Mission (SRTM) Hydrologically Enforced Digital Elevation Model (DEM-H). 1.0 ed. Elevation Information System (ELVIS): GeoSciences Australia.
- Nruthya, K. & Srinivas, V. V. 2015. Evaluating Methods to Predict Streamflow at Ungauged Sites Using Regional Flow Duration Curves: A Case Study. *Aquatic Procedia*, 4, 641-648.
- Nvis Technical Working Group 2017. Australian Vegetation Attribute Manual: National Vegetation Information System, Version 7.0. *In: ENERGY, D. O. T. E. A. (ed.)*. Canberra: Department of the Environment and Energy.
- Oksanen, J., Blanchet, F., Kindt, R., Legendre, P., Minchin, P., O'hara, R., Simpson, G., Solymos, P., Henry, M. & Stevens, M. 2015. Vegan: community ecology package. Ordination methods, diversity analysis and other functions for community and vegetation ecologists. *Version 2.3-1. R package version 2.3*, 1.
- Orians, G. H. & Milewski, A. V. 2007. Ecology of Australia: the effects of nutrient-poor soils and intense fires. *Biological Reviews*, 82, 393-423.

- Palmquist, K. A., Peet, R. K. & Mitchell, S. R. 2015. Scale-dependent responses of longleaf pine vegetation to fire frequency and environmental context across two decades. *Journal of Ecology*, 103, 998-1008.
- Pate, J. S., Froend, R. H., Bowen, B. J., Hansen, A. & Kuo, J. 1990. Seedling Growth and Storage Characteristics of Seeder and Resprouter Species of Mediterranean-type Ecosystems of S. W. Australia. *Annals of Botany*, 65, 585-601.
- Paula, S. & Pausas, J. G. 2006. Leaf traits and resprouting ability in the Mediterranean basin. *Functional Ecology*, 20, 941-947.
- Pausas, J. G. & Bond, W. J. 2019. Humboldt and the reinvention of nature. *Journal of Ecology*, 107, 1031-1037.
- Peet, R. K. & Roberts, D. W. 2013. Classification of Natural and Semi-natural Vegetation. *Vegetation Ecology*. John Wiley & Sons, Ltd.
- Pei, T., Qin, C.-Z., Zhu, A.-X., Yang, L., Luo, M., Li, B. & Zhou, C. 2010. Mapping soil organic matter using the topographic wetness index: A comparative study based on different flow-direction algorithms and kriging methods. *Ecological Indicators*, 10, 610-619.
- Petrone, K. C., Hughes, J. D., Van Niel, T. G. & Silberstein, R. P. 2010. Streamflow decline in southwestern Australia, 1950–2008. *Geophysical Research Letters*, 37.
- Phillips, S. J., Anderson, R. P. & Schapire, R. E. 2006. Maximum entropy modeling of species geographic distributions. *Ecological modelling*, 190, 231-259.
- Poot, P. & Veneklaas, E. 2013. Species distribution and crown decline are associated with contrasting water relations in four common sympatric eucalypt species in southwestern Australia. *An International Journal on Plant-Soil Relationships*, 364, 409-423.
- Possingham, H. P., Franklin, J., Wilson, K. & Regan, T. J. 2005. The roles of spatial heterogeneity and ecological processes in conservation planning. *Ecosystem function in heterogeneous landscapes*. Springer New York.
- Pressey, R. L. 2004. Conservation Planning and Biodiversity: Assembling the Best Data for the Job. *Conservation Biology*, 18, 1677-1681.
- Pressey, R. L. & Bottrill, M. C. 2009. Approaches to landscape-and seascape-scale conservation planning: convergence, contrasts and challenges. *Oryx*, 43, 464-475.
- Prober, S. M., Thiele, K. R., Rundel, P. W., Yates, C. J., Berry, S. L., Byrne, M., Christidis, L., Gosper, C. R., Grierson, P. F. & Lemson, K. 2012. Facilitating adaptation of biodiversity to climate change: a conceptual framework applied to the world's largest Mediterranean-climate woodland. *Climatic Change*, 110, 227-248.
- R Core Team 2017. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
- Randin, C. F., Ashcroft, M. B., Bolliger, J., Cavender-Bares, J., Coops, N. C., Dullinger, S., Dirnböck, T., Eckert, S., Ellis, E., Fernández, N., Giuliani, G., Guisan, A., Jetz, W., Joost, S., Karger, D., Lembrechts, J., Lenoir, J., Luoto, M., Morin, X., Price, B., Rocchini, D., Schaepman, M., Schmid, B., Verburg, P., Wilson, A., Woodcock, P., Yoccoz, N. & Payne, D. 2020. Monitoring biodiversity in the Anthropocene using remote sensing in species distribution models. *Remote Sensing of Environment*, 239, 111626.
- Regional Forest Agreement Steering Committee 1998. National Estate Identification and Assessment in the South West Forest Region of Western Australia. In: GOVERNMENT, C. O. A. A. W. A. (ed.).
- Reschke, C., Reid, R., Jones, J., Feeney, T. & Potter, H. 1999. Conserving Great Lakes Alvars.

- Reside, A., Vanderwal, J., Phillips, B., Shoo, L., Rosauer, D., Anderson, B., Welbergen, J., Moritz, C., Ferrier, S., Harwood, T., Williams, K., Mackey, B., Hugh, S. & Williams, S. 2013. Climate change refugia for terrestrial biodiversity: Defining areas that promote species persistence and ecosystem resilience in the face of global climate change. National Climate Change Adaptation Research Facility, Gold Coast.
- Reside, A. E., Butt, N. & Adams, V. M. 2018. Adapting systematic conservation planning for climate change. *Biodiversity and Conservation*, 27, 1-29.
- Reside, A. E., Welbergen, J. A., Phillips, B. L., Wardell-Johnson, G. W., Keppel, G., Ferrier, S., Williams, S. E. & Vanderwal, J. 2014. Characteristics of climate change refugia for Australian biodiversity. *Austral Ecology*, 39, 887-897.
- Richardson, D. M., Holmes, P. M., Esler, K. J., Galatowitsch, S. M., Stromberg, J. C., Kirkman, S. P., Pyšek, P. & Hobbs, R. J. 2007. Riparian vegetation: degradation, alien plant invasions, and restoration prospects. *Diversity and Distributions*, 13, 126-139.
- Ricotta, C., Carboni, M. & Acosta, A. T. 2015. Let the concept of indicator species be functional! *Journal of vegetation science*, 26, 839-847.
- Ripley, B., Venables, W. & Ripley, M. B. 2016. Package ‘nnet’. *R package version*, 7, 3-12.
- Riviera, F. 2019. *Patterns and drivers of kwongan vegetation restored after mining: A multifaceted approach*. PhD Thesis, University of Western Australia.
- Roberts, D. W. 1986. Ordination on the basis of fuzzy set theory. *Vegetatio*, 66, 123-131.
- Roberts, D. W. 2015. Vegetation classification by two new iterative reallocation optimization algorithms. *Plant Ecology*, 216, 741-758.
- Roberts, D. W. & Peres-Neto, P. 2017. Distance, dissimilarity, and mean–variance ratios in ordination. *Methods in ecology and evolution*, 8, 1398-1407.
- Robinson, T., Di Virgilio, G., Temple-Smith, D., Hesford, J. & Wardell-Johnson, G. 2019. Characterisation of range restriction amongst the rare flora of Banded Ironstone Formation ranges in semiarid south-western Australia. *Australian Journal of Botany*, 67, 234-247.
- Rondinini, C., Wilson, K. A., Boitani, L., Grantham, H. & Possingham, H. P. 2006. Tradeoffs of different types of species occurrence data for use in systematic conservation planning. *Ecology letters*, 9, 1136-1145.
- Rosenheim, J. A. & Gratton, C. 2017. Ecoinformatics (Big Data) for Agricultural Entomology: Pitfalls, Progress, and Promise. *Annual Review of Entomology*, 62, 399-417.
- Roux, P. C. L., Virtanen, R., Heikkinen, R. K. & Luoto, M. 2012. Biotic interactions affect the elevational ranges of high-latitude plant species. *Ecography*, 35, 1048-1056.
- Ruane, S. 2018. Using a worldview lens to examine complex policy issues: a historical review of bushfire management in the south west of Australia. *Local Environment*, 23, 777-795.
- Rull, V. 2009. Microrefugia. *Journal of Biogeography*, 36, 481-484.
- Rundle, G. E. 1996. History of conservation reserves in the south-west of Western Australia (Vol 79, Part 4, 1996). *Journal of the Royal Society of Western Australia*, 79, 225-240.
- Runge, C. A., Tulloch, A. I. T., Possingham, H. P., Tulloch, V. J. D. & Fuller, R. A. 2016. Incorporating dynamic distributions into spatial prioritization. *Diversity and Distributions*, 22, 332-343.
- Russell, B. 1948. *Human Knowledge: Its Scope and Limits.*, London, George Allen and Unwin.

- Rykiel Jr, E. J. 1996. Testing ecological models: the meaning of validation. *Ecological Modelling*, 90, 229-244.
- Sala, O. E., Stuart Chapin, F., Iii, Armesto, J. J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L. F., Jackson, R. B., Kinzig, A., Leemans, R., Lodge, D. M., Mooney, H. A., Oesterheld, M. N., Poff, N. L., Sykes, M. T., Walker, B. H., Walker, M. & Wall, D. H. 2000. Global Biodiversity Scenarios for the Year 2100. *Science*, 287, 1770-1774.
- Sander, J. & Wardell-Johnson, G. 2011. Fine-scale patterns of species and phylogenetic turnover in a global biodiversity hotspot: Implications for climate change vulnerability. *Journal of Vegetation Science*, 22, 766-780.
- Sankarasubramanian, A., Vogel, R. M. & Limbrunner, J. F. 2001. Climate elasticity of streamflow in the United States. *Water Resources Research*, 37, 1771-1781.
- Schut, A. G., Wardell-Johnson, G. W., Yates, C. J., Keppel, G., Baran, I., Franklin, S. E., Hopper, S. D., Van Niel, K. P., Mucina, L. & Byrne, M. 2014. Rapid characterisation of vegetation structure to predict refugia and climate change impacts across a global biodiversity hotspot. *PLoS One*, 9, e82778.
- Singh, V. P. 2018. Hydrologic modeling: progress and future directions. *Geoscience Letters*, 5, 15.
- Slavich, E., Warton, D. I., Ashcroft, M. B., Gollan, J. R. & Ramp, D. 2014. Topoclimate versus macroclimate: how does climate mapping methodology affect species distribution models and climate change projections? *Diversity and Distributions*, 20, 952-963.
- Sparrow, A. D. 1991. *A Geobotanical Study of the Remnant Natural Vegetation of Temperate South Australia*. Doctor of Philosophy, University of Adelaide.
- Starkweather, J. & Moske, A. K. 2011. Multinomial logistic regression. *Consulted page at September 10th: http://www.unt.edu/rss/class/Jon/Benchmarks/MLR_JDS_Aug2011.pdf*, 29, 2825-2830.
- Stein, A., Gerstner, K. & Kreft, H. 2014. Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecology Letters*, 17, 866-880.
- Stein, A. & Kreft, H. 2015. Terminology and quantification of environmental heterogeneity in species-richness research. *Biological Reviews*, 90, 815-836.
- Stephens, S. L., Collins, B. M., Fettig, C. J., Finney, M. A., Hoffman, C. M., Knapp, E. E., North, M. P., Safford, H. & Wayman, R. B. 2018. Drought, Tree Mortality, and Wildfire in Forests Adapted to Frequent Fire. *BioScience*, 68, 77-88.
- Stewart, J. R. & Lister, A. M. 2001. Cryptic northern refugia and the origins of the modern biota. *Trends in Ecology & Evolution*, 16, 608-613.
- Stewart, J. R., Lister, A. M., Barnes, I. & Dalén, L. 2009. Refugia revisited: individualistic responses of species in space and time. *Proc Biol Sci*, 277, 661-671.
- Stewart, J. R., Lister, A. M., Barnes, I. & Dalén, L. 2010. Refugia revisited: individualistic responses of species in space and time. *Proceedings of the Royal Society B: Biological Sciences*, 277, 661-671.
- Stralberg, D., Arseneault, D., Baltzer, J. L., Barber, Q. E., Bayne, E. M., Boulanger, Y., Brown, C. D., Cooke, H. A., Devito, K., Edwards, J., Estevo, C. A., Flynn, N., Frelich, L. E., Hogg, E. H., Johnston, M., Logan, T., Matsuoka, S. M., Moore, P., Morelli, T. L., Morissette, J. L., Nelson, E. A., Nenzén, H., Nielsen, S. E., Parisien, M.-A., Pedlar, J. H., Price, D. T., Schmiegelow, F. K., Slattery, S. M., Sonnentag, O., Thompson, D. K. & Whitman, E. 2020a. Climate-change refugia in boreal North America: what, where, and for how long? *Frontiers in Ecology and the Environment*, 18, 261-270.

- Stralberg, D., Carroll, C. & Nielsen, S. E. 2020b. Toward a climate-informed North American protected areas network: Incorporating climate-change refugia and corridors in conservation planning. *Conservation Letters*, 13, e12712.
- Stralberg, D., Wang, X., Parisien, M. A., Robinne, F. N., Sóllymos, P., Mahon, C. L., Nielsen, S. E. & Bayne, E. M. 2018. Wildfire-mediated vegetation change in boreal forests of Alberta, Canada. *Ecosphere*, 9, n/a-n/a.
- Su, Y., Guo, Q., Fry, D. L., Collins, B. M., Kelly, M., Flanagan, J. P. & Battles, J. J. 2016. A vegetation mapping strategy for conifer forests by combining airborne LiDAR data and aerial imagery. *Canadian Journal of Remote Sensing*, 42, 1-15.
- Sweet, L. C., Green, T., Heintz, J. G., Frakes, N., Graver, N., Rangitsch, J. S., Rodgers, J. E., Heacox, S. & Barrows, C. W. 2019. Congruence between future distribution models and empirical data for an iconic species at Joshua Tree National Park. *Ecosphere*, 10, e02763.
- Tapper, S. L., Byrne, M., Yates, C., Keppel, G., Hopper, S., Van Niel, K., Schut, A., Mucina, L. & Wardell-Johnson, G. 2014b. Isolated with persistence or dynamically connected? Genetic patterns in a common granite outcrop endemic. *Diversity and Distributions*, 20, 987-1001.
- Tapper, S. L., Byrne, M., Yates, C. J., Keppel, G., Hopper, S. D., Van Niel, K., Schut, A. G., Mucina, L. & Wardell-Johnson, G. W. 2014a. Prolonged isolation and persistence of a common endemic on granite outcrops in both mesic and semi-arid environments in south-western Australia. *Journal of Biogeography*, 41, 2032-2044.
- Thackway, R. & Cresswell, I. 1995. An Interim Biogeographic Regionalisation for Australia: A framework for setting priorities in the national reserve system cooperative program. Australian Nature Conservation Agency, Canberra.
- The State of Western Australia 1999. The Regional Forest Agreement for the South-West Forest Region of Western Australia.
- Thibaud, E., Petitpierre, B., Broennimann, O., Davison, A. C. & Guisan, A. 2014. Measuring the relative effect of factors affecting species distribution model predictions. *Methods in ecology and evolution*, 5, 947-955.
- Thorne, J. H., Gogol-Prokurat, M., Hill, S., Walsh, D., Boynton, R. M. & Choe, H. 2020. Vegetation refugia can inform climate-adaptive land management under global warming. *Frontiers in Ecology and the Environment*, 18, 281-287.
- Thuiller, W., Guéguen, M., Renaud, J., Karger, D. N. & Zimmermann, N. E. 2019. Uncertainty in ensembles of global biodiversity scenarios. *Nature Communications*, 10, 1-9.
- Tichý, L., Chytrý, M. & Botta-Dukát, Z. 2014. Semi-supervised classification of vegetation: preserving the good old units and searching for new ones. *Journal of Vegetation Science*, 25, 1504-1512.
- Tichý, L., Chytrý, M., Hájek, M., Talbot, S. S. & Botta-Dukát, Z. 2010. OptimClass: Using species-to-cluster fidelity to determine the optimal partition in classification of ecological communities. *Journal of Vegetation Science*, 21, 287-299.
- Tozer, M. G., Simpson, C. C., Jansens, I. B. & Keith, D. A. 2017. Biogeography of Australia's dry sclerophyll forests: Drought, nutrients and fire. In: KEITH, D. A. (ed.) *Australian vegetation*. Cambridge, UK: Cambridge University Press.
- Trotter, L., Robinson, T., Wardell-Johnson, G., Grigg, A. & Luxton, S. 2018. FIMS: a free and open-source spatial database system for plant observation and mobile data collection.
- Tzedakis, P. C., Emerson, B. C. & Hewitt, G. M. 2013. Cryptic or mystic? Glacial tree refugia in northern Europe. *Trends in Ecology & Evolution*, 28, 696-704.

- Van Der Putten, W. H., Macel, M. & Visser, M. E. 2010. Predicting species distribution and abundance responses to climate change: why it is essential to include biotic interactions across trophic levels. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365, 2025-2034.
- Visser, A. G., Beevers, L. & Patidar, S. 2019. A coupled modelling framework to assess the hydroecological impact of climate change. *Environmental Modelling & Software*, 114, 12-28.
- Walker, D. A., Breen, A. L., Druckenmiller, L. A., Wirth, L. W., Fisher, W., Reynolds, M. K., Šibík, J., Walker, M. D., Hennekens, S. & Boggs, K. 2016. The Alaska Arctic Vegetation Archive (AVA-AK). *Phytocoenologia*, 46, 221-229.
- Walther, G.-R. 2010. Community and ecosystem responses to recent climate change. *Philosophical Transactions of the Royal Society B*, 365, 2019-2024.
- Wan, H. Y., Cushman, S. A. & Ganey, J. L. 2020. The effect of scale in quantifying fire impacts on species habitats. *Fire Ecology*, 16, 1-15.
- Wardell-Johnson, G., Calver, M., Burrows, N. & Di Virgilio, G. 2015. Integrating rehabilitation, restoration and conservation for a sustainable jarrah forest future during climate disruption. *Pacific Conservation Biology*, 21, 175-185.
- Wardell-Johnson, G. & Horwitz, P. 1996. Conserving biodiversity and the recognition of heterogeneity in ancient landscapes: a case study from south-western Australia. *Forest Ecology and Management*, 85, 219-238.
- Wardell-Johnson, G. & Horwitz, P. 2000. The recognition of heterogeneity and restricted endemism in the management of forested ecosystems in south-western Australia. *Australian Forestry*, 63, 218-225.
- Wardell-Johnson, G., Luxton, S., Craig, K., Brown, V., Evans, N. & Kennedy, S. 2018. Implications of floristic patterns, and changes in stand structure following a large-scale, intense fire across forested ecosystems in south-western Australia's high-rainfall zone. *Pacific Conservation Biology*, 23, 399-412.
- Wardell-Johnson, G. & Roberts, J. D. 1993. Biogeographic Barriers in a Subdued Landscape: The Distribution of the *Geocrinia rosea* (Anura: Myobatrachidae) Complex in South-Western Australia. *Journal of Biogeography*, 20, 95-108.
- Wardell-Johnson, G., Wardell-Johnson, A., Schultz, B., Dortch, J., Robinson, T., Collard, L. & Calver, M. 2019. The contest for the tall forests of south-western Australia and the discourses of advocates. *Pacific Conservation Biology*, 25, 50-71.
- Wardell-Johnson, G. W., Williams, M. R., Mellican, A. E. & Annells, A. 2004. Floristic patterns and disturbance history in karri forest, south-western Australia: 1. Environment and species richness. *Forest Ecology and Management*, 199, 449-460.
- Wardell-Johnson, G. W., Williams, M. R., Mellican, A. E. & Annells, A. 2007. Floristic patterns and disturbance history in karri (*Eucalyptus diversicolor*: Myrtaceae) forest, south-western Australia: 2. Origin, growth form and fire response. *Acta Oecologica*, 31, 137-150.
- Ware, C., Williams, K. J., Harding, J., Hawkins, B., Harwood, T., Manion, G., Perkins, G. C. & Ferrier, S. 2018. Improving biodiversity surrogates for conservation assessment: A test of methods and the value of targeted biological surveys. *Diversity and Distributions*, 24, 1333-1346.
- Weber, L. C., Vanderwal, J., Schmidt, S., McDonald, W. J. & Shoo, L. P. 2014. Patterns of rain forest plant endemism in subtropical Australia relate to stable mesic refugia and species dispersal limitations. *Journal of Biogeography*, 41, 222-238.
- Wei, T., Simko, V., Levy, M., Xie, Y., Jin, Y. & Zemla, J. 2017. Package 'corrplot'. *Statistician*, 56, 316-324.
- Weiss, S. & Ferrand, N. 2007. *Phylogeography of southern European refugia*, Springer.

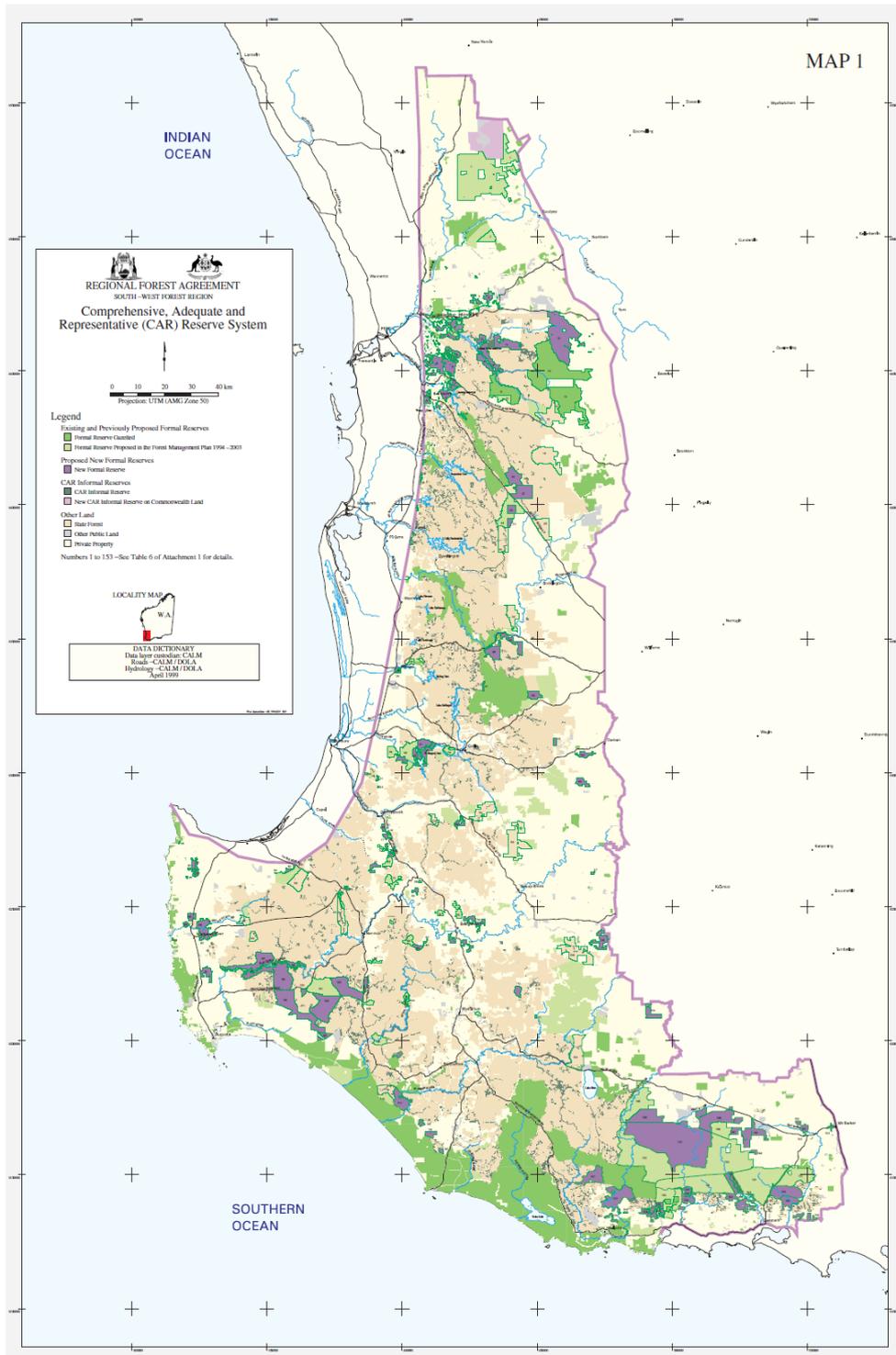
- Western Australian Herbarium 1998-. FloraBase—the Western Australian Flora. *In*: DEPARTMENT OF BIODIVERSITY, C. A. A. (ed.).
- Whetton, P., Hennessy, K., Clarke, J., Mcinnes, K. & Kent, D. 2012. Use of Representative Climate Futures in impact and adaptation assessment. *Climatic Change*, 115, 433-442.
- White, H. A. 2017. *Resistance, resilience and adaptation to climate change in riparian ecosystems*. The University of Western Australia.
- Whittaker, R. H. 1967. Gradient analysis of vegetation. *Biological reviews*, 42, 207-264.
- Wickham, H., Francois, R., Henry, L. & Müller, K. 2017. dplyr: A Grammar of Data Manipulation. R package. version 0.7.4. ed.
- Williams, K., De Vries, R., Stein, J., Rosauer, D. & Storey, R. 2018. 0.01 degree stack of edaphic and physiographic layers for continental analysis of biodiversity pattern, version 1.0. v2 ed.: Commonwealth Scientific and Industrial Research Organisation (CSIRO).
- Williams, K. & Mitchell, D. 2001. Jarrah forest 1 (JF1—northern jarrah forest subregion). *A Biodiversity Audit of Western Australia's*, 53, 369-381.
- Williams, R. J., Bradstock, R. A., Cary, G. J., Enright, N. J., Gill, A. M., Leidloff, A. C., Lucas, C., Whelan, R. J., Andersen, A. N., Bowman, D. J. M. S., Clarke, P. J., Cook, G. D., Hennessy, K. J. & York, A. 2009. Interactions between climate change, fire regimes and biodiversity in Australia: A preliminary assessment. . *In*: DEPARTMENT OF CLIMATE CHANGE AND DEPARTMENT OF THE ENVIRONMENT, W., HERITAGE AND THE ARTS (ed.). Canberra, Australia.
- Wilson, K. A., Cabeza, M. & Klein, C. J. 2009. Fundamental concepts of spatial conservation prioritization. *Spatial Conservation Prioritization: quantitative methods and computational tools*: Oxford University Press, New York.
- Wiser, S. K., Bellingham, P. J. & Burrows, L. E. 2001. Managing biodiversity information: development of New Zealand's National Vegetation Survey databank. *New Zealand Journal of Ecology*, 1-17.
- Wiser, S. K. & De Cáceres, M. 2013. Updating vegetation classifications: an example with New Zealand's woody vegetation. *Journal of Vegetation Science*, 24, 80-93.
- Wisz, M. S., Pottier, J., Kissling, W. D., Pellissier, L., Lenoir, J., Damgaard, C. F., Dormann, C. F., Forchhammer, M. C., Grytnes, J. A. & Guisan, A. 2013. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biological Reviews*, 88, 15-30.
- Wolock, D. M. & McCabe Jr, G. J. 1995. Comparison of single and multiple flow direction algorithms for computing topographic parameters in TOPMODEL. *Water Resources Research*, 31, 1315-1324.
- Wong, H. J. & Hartigan, J. 1979. Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28, 100-108.
- Xie, Y., Sha, Z. & Yu, M. 2008. Remote sensing imagery in vegetation mapping: a review. *Journal of plant ecology*, 1, 9-23.
- Yates, C. J., McNeill, A., Elith, J. & Midgley, G. F. 2010. Assessing the impacts of climate change and land transformation on Banksia in the South West Australian Floristic Region. *Diversity and Distributions*, 16, 187-201.
- Yates, C. J., Robinson, T., Wardell-Johnson, G. W., Keppel, G., Hopper, S. D., Schut, A. G. T. & Byrne, M. 2019. High species diversity and turnover in granite inselberg floras highlight the need for a conservation strategy protecting many outcrops. *Ecology and Evolution*, 9, 7660-7675.

- Yates, K. L., Bouchet, P. J., Caley, M. J., Mengersen, K., Randin, C. F., Parnell, S., Fielding, A. H., Bamford, A. J., Ban, S., Barbosa, A. M., Dormann, C. F., Elith, J., Embling, C. B., Ervin, G. N., Fisher, R., Gould, S., Graf, R. F., Gregr, E. J., Halpin, P. N., Heikkinen, R. K., Heinänen, S., Jones, A. R., Krishnakumar, P. K., Lauria, V., Lozano-Montes, H., Mannocci, L., Mellin, C., Mesgaran, M. B., Moreno-Amat, E., Mormede, S., Novaczek, E., Oppel, S., Ortuño Crespo, G., Peterson, A. T., Rapacciuolo, G., Roberts, J. J., Ross, R. E., Scales, K. L., Schoeman, D., Snelgrove, P., Sundblad, G., Thuiller, W., Torres, L. G., Verbruggen, H., Wang, L., Wenger, S., Whittingham, M. J., Zharikov, Y., Zurell, D. & Sequeira, A. M. M. 2018. Outstanding Challenges in the Transferability of Ecological Models. *Trends in Ecology & Evolution*, 33, 790-802.
- Zurell, D., Franklin, J., König, C., Bouchet, P. J., Dormann, C. F., Elith, J., Fandos, G., Feng, X., Guillera-Arroita, G., Guisan, A., Lahoz-Monfort, J. J., Leitão, P. J., Park, D. S., Peterson, A. T., Rapacciuolo, G., Schmatz, D. R., Schröder, B., Serra-Diaz, J. M., Thuiller, W., Yates, K. L., Zimmermann, N. E. & Merow, C. 2020. A standard protocol for reporting species distribution models. *Ecography*, 43, 1261-1277.
- Zuur, A. F., Ieno, E. N. & Elphick, C. S. 2010. A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, 1, 3-14.

Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged.

Appendix A

A1 The current reserve system for forested region of SWA meets minimum CAR (Comprehensive, Adequate and Representative) requirements. Major creek lines are recognised in the ‘CAR Informal Reserve’ category. Source: Department of Biodiversity, Conservation and Attractions: Map 1, CAR Reserve Map, Regional Forest Agreement for the South-West Forest Region of Western Australia, May 1999.

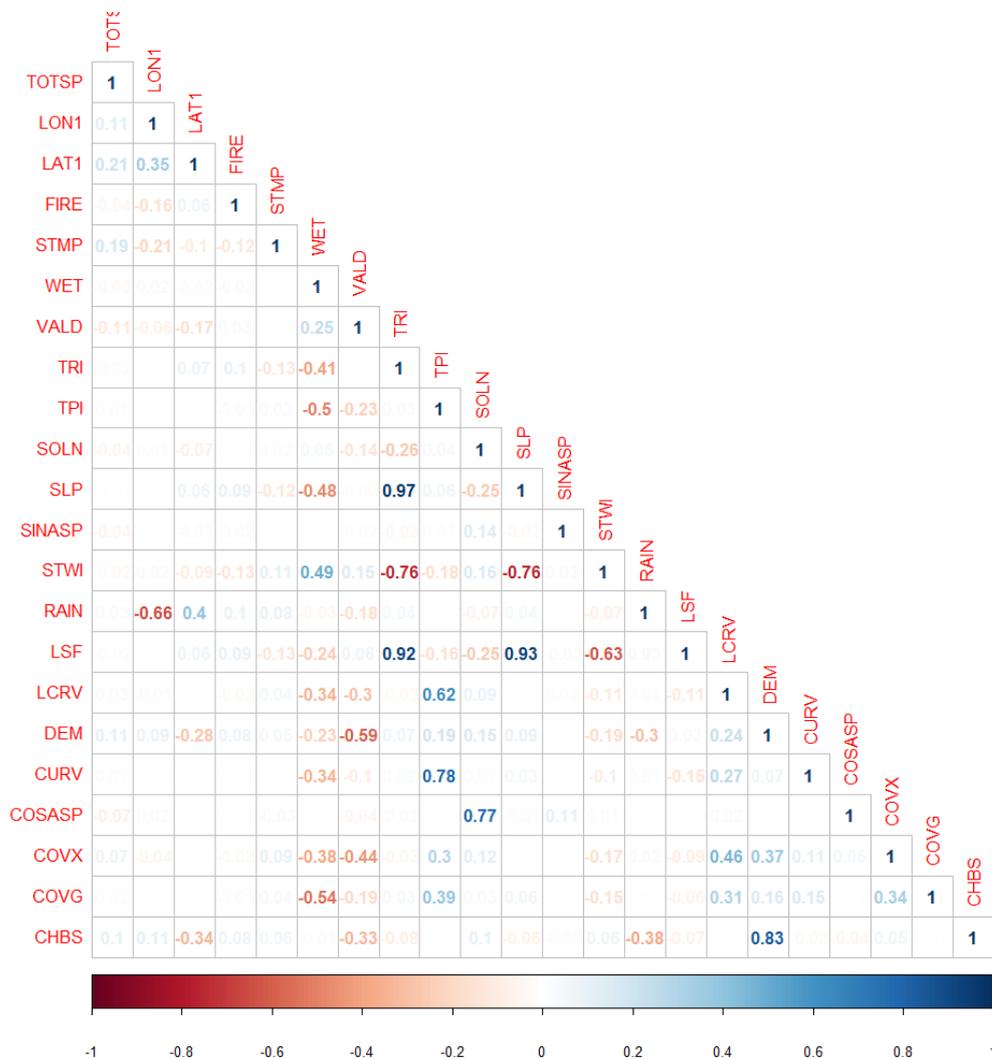


Appendix B

- B1** Alcoa raw data, access database.
- B2** Alcoa traits data, access database.
- B3** Northern jarrah forest traits data, access database.
- B4** Family data, all.
- B5** Vegetation complexes, raw data.
- B6** Conservation indices, raw data.

Appendix C

- C1** Correlation matrix for potential environmental descriptors for community classification analyses.

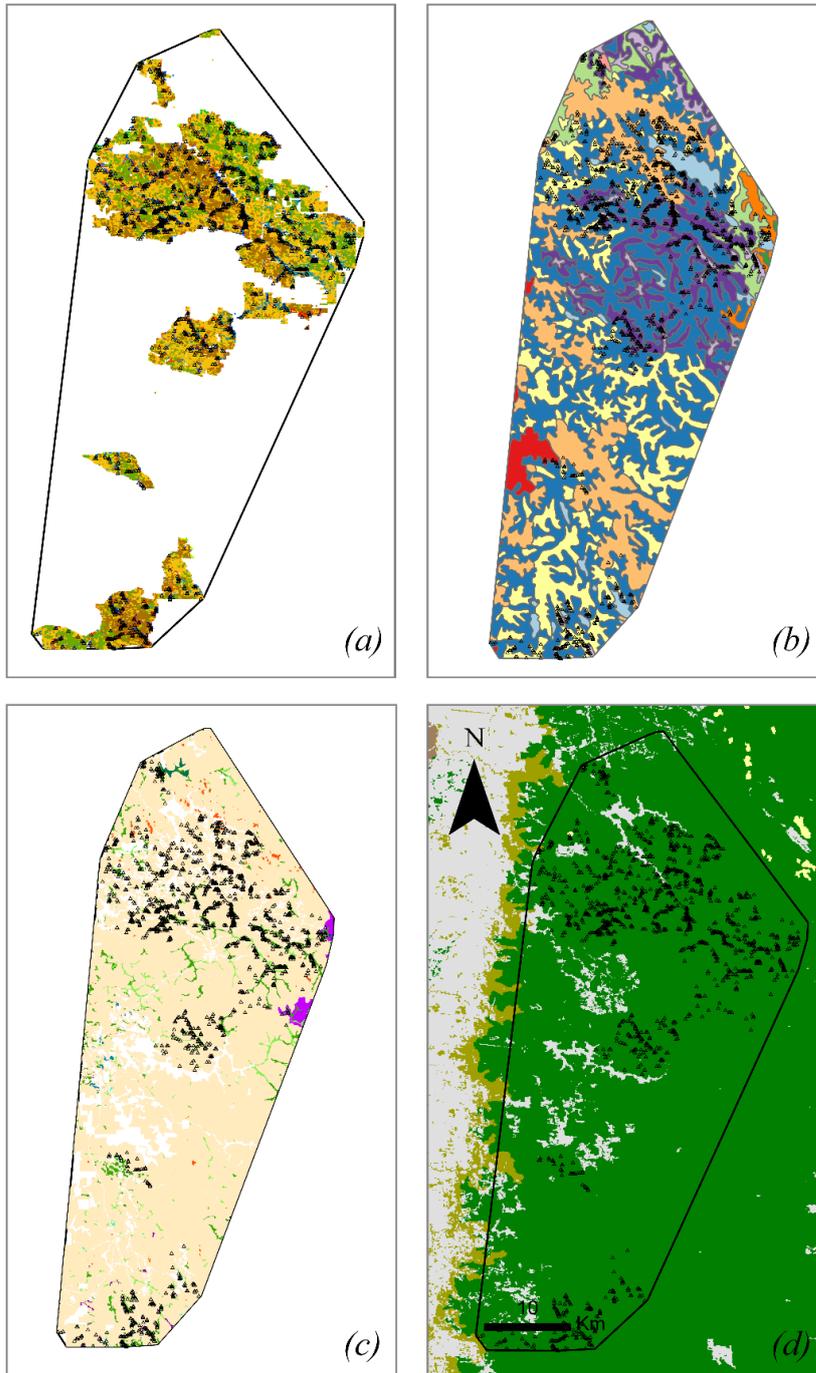


C2 Choice of clustering algorithm – additional information.

Cluster analysis was done using the non-hierarchical (NH) allocation algorithm ‘ALOC’ in PATN v 3.12 (Belbin 1987; Belbin and McDonald 1993). By calculating object-group associations (not all object-object associations) it is effective in clustering large datasets. ALOC allocates objects to a set of seeds – starting with the first object in the dataset and iterating from there (although prior groups can also be used). It has been found to be reasonably insensitive to seeds (Belbin 1987) however, if results are compared to other classification methods it is likely that the classifications will differ (Belbin and McDonald 1993). The number of groups produced by ALOC is controlled by an allocation radius (an association value between zero and one that acts as a threshold for the generation of additional seeds). The number of iterations is \sqrt{N} – with more iterations required to reach optimality as the size of the dataset increases. Belbin (1987) first argued for the use of NH approaches, particularly in larger datasets. He states “While the aim of NH clustering methods is to produce some optimal partitioning of objects into clusters, SAHN methods are designed to optimise the fusions and therefore define clusters only in an indirect fashion. Most tend to use SAHN methods to provide a partitioning, making little or no use of the inter-group structure.” (pg. 32, Belbin 1987). ALOC has been tested against hierarchical methods and shown to have good recovery of clusters (Belbin 1987; Belbin and McDonald 1993).

Alternative non-hierarchical clustering algorithms that were investigated, but not selected for analysis include: K-means, Partitioning Around Medoids (PAM), Clustering Large Applications (CLARA), OPTPART and OPTSIL (MacQueen 1967; Kaufman and Rousseeuw 2009; Roberts 2015; Kassambara 2017). K-means is a commonly used unsupervised machine learning algorithm (Kassambara 2017). It classifies objects so that within cluster similarity is maximised and between-cluster minimised. There are several k-means algorithms however, the standard (Wong and Hartigan 1979), uses Euclidean distance, which is sensitive to outliers and not appropriate in data with many zeros. The number of clusters (k) must also be pre-set, which may force groups that are not optimal. PAM (or K-medoids clustering) is less sensitive than k-means to outliers, and more likely to converge at a similar solution from different starting medoids (Kaufman and Rousseeuw 2009). However, PAM is considered suitable for at most two to three thousand plots (R help file for the cluster package). CLARA is the recommended non-hierarchical alternative in R, but also uses Euclidean (or Manhattan) distance. Distances measures are not ideal for analysing ecological data, as while they quantify the differences between objects, they do not consider what the two objects have in common. This is in contrast to dissimilarity measurements like Bray Curtis, which generally compare this difference to what the objects have in common and scale the difference accordingly (Roberts and Peres-Neto 2017). Roberts (2015) states an updated case for the use of non-hierarchical clustering with his algorithms OPTPART and OPTSIL. However, OPTSIL took 50 hours of CPU time for 209 species and 424 sites (Intel I7 processor). As this dataset has over 400 species and 30,000 sites, the computing time would have been large. While high-performance computing provides solutions to this, fast desktop alternatives like PATN are a practical and effective alternative.

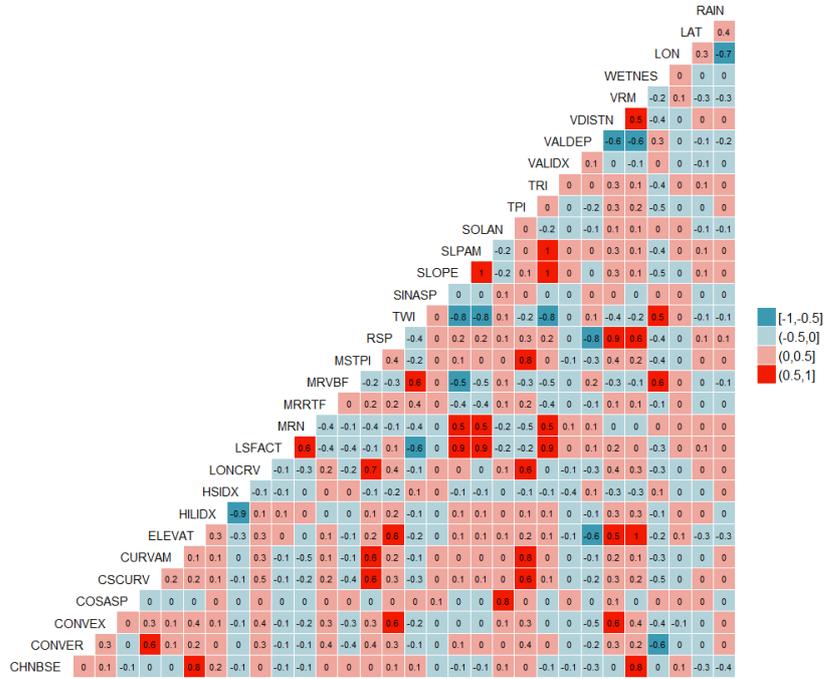
- C3** R script, classification.
C4 Species frequency per group, 30-group classification (Excel file).
C5 Results from two-group ALOC vegetation classification, overlain over current vegetation mapping systems available for the study area: (a) site vegetation types, (b) vegetation complexes, (c) forest ecosystem units; and (d) the National Vegetation information System (NVIS).



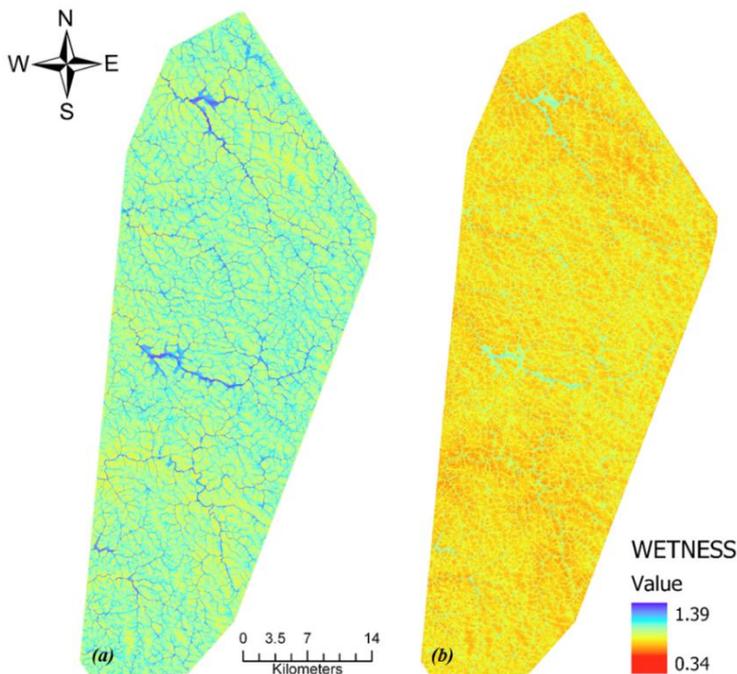
- C6** Two-step, two way table.

Appendix D

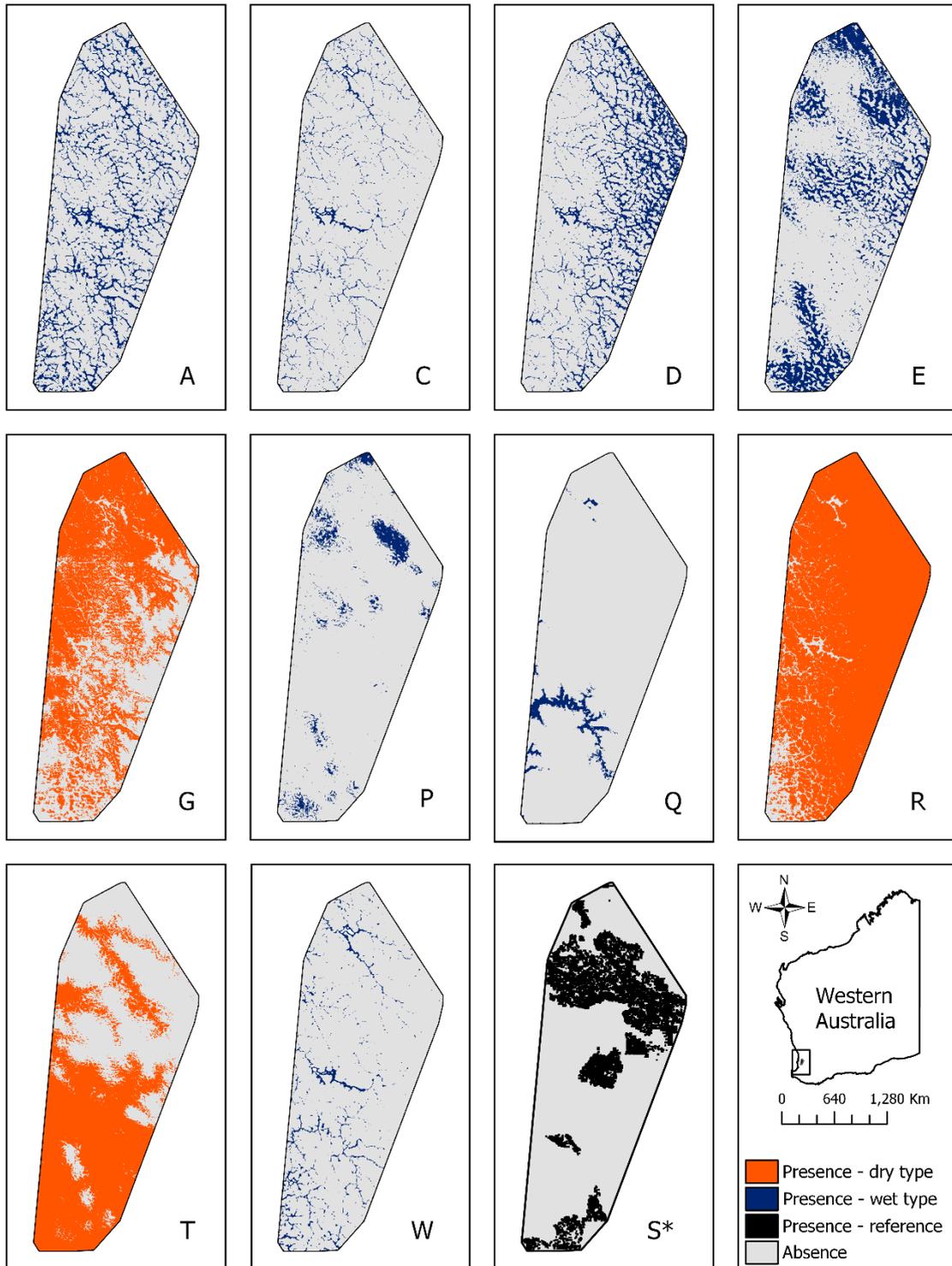
- D1** Site vegetation types metadata.
- D2** Environmental metadata.
- D3** R script, multinomial logistic regression modelling.
- D4** Correlation matrix for potential topographic predictors.



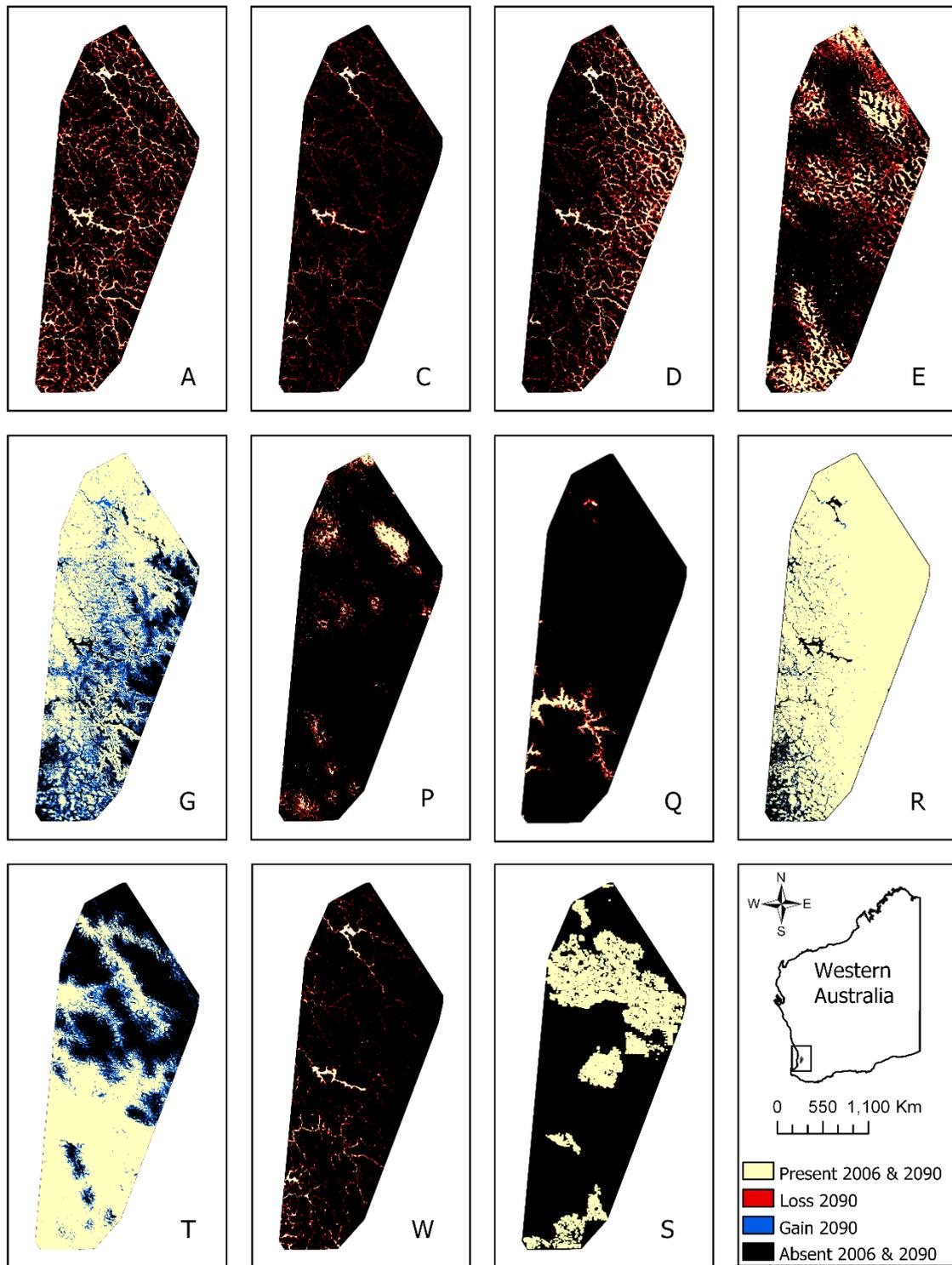
- D5** SAGA wetness index (SWI) surfaces for (a) 2006 and (b) 2090, based on RCP 4.5 rainfall projections and streamflow-rainfall elasticity calculations.



D6 Binomial logistic regression maps for current (2006) conditions for each site vegetation type.



D7 Binomial logistic regression maps of change from 2006 to 2090, for each site vegetation type.



D8 Summary of additional binomial and multinomial logistic regression accuracy assessment results. Binomial models show higher individual accuracy, but were not used due to overlapping distributions and non-independence.

| SVT | Accuracy Statistic | Modelling method | | | |
|-----|-----------------------|------------------|------------|---------------|---------------|
| | | Binomial 1 | Binomial 2 | Multinomial 1 | Multinomial 2 |
| A | Balanced accuracy (%) | 0.8 | 0.78 | 0.72 | 0.73 |
| | Kappa | 0.56 | 0.56 | 0.4 | 0.45 |
| C | Balanced accuracy (%) | 0.77 | 0.73 | 0.73 | 0.69 |
| | Kappa | 0.48 | 0.43 | 0.45 | 0.22 |
| D | Balanced accuracy (%) | 0.76 | 0.73 | 0.64 | 0.62 |
| | Kappa | 0.45 | 0.41 | 0.18 | 0.22 |
| E | Balanced accuracy (%) | 0.81 | 0.8 | 0.71 | 0.77 |
| | Kappa | 0.62 | 0.53 | 0.35 | 0.5 |
| G | Balanced accuracy (%) | 0.79 | 0.81 | 0.75 | 0.71 |
| | Kappa | 0.54 | 0.71 | 0.32 | 0.42 |
| P | Balanced accuracy (%) | 0.74 | 0.7 | 0.54 | 0.62 |
| | Kappa | 0.48 | 0.42 | 0.03 | 0.21 |
| Q | Balanced accuracy (%) | 0.88 | 0.8 | 0.8 | 0.82 |
| | Kappa | 0.72 | 0.54 | 0.12 | 0.3 |
| R | Balanced accuracy (%) | 0.85 | 0.75 | 0.71 | 0.55 |
| | Kappa | 0.67 | 0.44 | 0.05 | 0.09 |
| S | Balanced accuracy (%) | 0.67 | 0.63 | NA | NA |
| | Kappa | 0.3 | 0.27 | NA | NA |
| T | Balanced accuracy (%) | 0.72 | 0.69 | 0.68 | 0.6 |
| | Kappa | 0.4 | 0.35 | 0.2 | 0.2 |
| W | Balanced accuracy (%) | 0.75 | 0.68 | 0.7 | 0.65 |
| | Kappa | 0.4 | 0.4 | 0.4 | 0.07 |

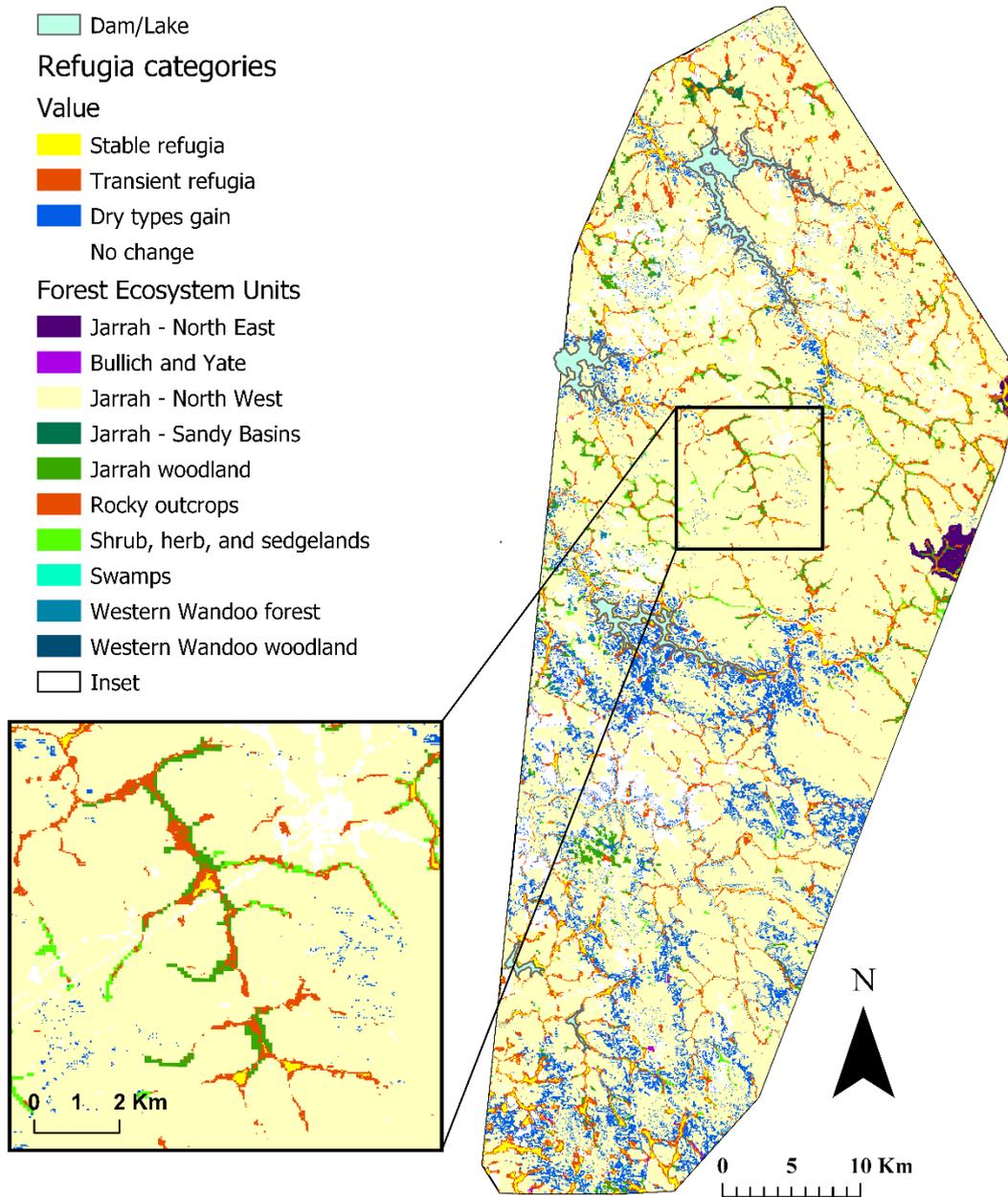
Binomial 1 Accuracy test - Arcmap method, raw data

Binomial 2 K-folds cross validation, range-standardised data. Mean accuracy = 0.81; mean kappa = 0.5

Multinomial1 Binomial eqn's from multinomial logistic regression + INTERACTIONS - arcmap method, Accuracy 0.5.

Multinomial2 Binomial eqn's from multinomial logistic regression - arcmap method - Accuracy 0.45, Kappa 0.13, p val < 0.001.

D9 Forest ecosystem units reflect changes in major vegetation units between upland and stream zone sites. Refugia track the most restricted units, but are also modelled across the general forest type ('Jarrah – North West'), particularly in the north east and south west of the study area.



Appendix E

E1 (A) Key habitat characteristics for the indicator species used in modelling and; (B) a summary of key species responses (percent contribution to refugia area).

(A)

| SVT | SVT* | Indicator Species | Habit | Key Habitat Characteristics |
|-----|---------|-----------------------------------|-------|--|
| A | C | <i>Astartea scoparia</i> | S | Loam, sand. |
| A | C | <i>Banksia littoralis</i> * | T | Grey or black peaty sand. Low-lying, seasonally damp areas, along watercourses. |
| A | W | <i>Eucalyptus patens</i> | T | Gravelly soils, sandy clay, loam. Depressions, stream banks, valleys. |
| A | C | <i>Eucalyptus rudis</i> | T | Sandy or loam soils. Wetter parts of south-western WA, flats, hillsides. |
| A | | <i>Hakea varia</i> * | S | White, grey or red loamy sand, clay loam, laterite. Seasonally-wet flats. |
| A | D, E, W | <i>Hypocalymma angustifolium</i> | S | Grey to white sand, peaty soils, sandy clay, sandstone. Flats, swamps, watercourses, near permanent fresh-water springs, outcrops, hillsides. |
| A | | <i>Melaleuca preissiana</i> * | S/T | Sandy soils. Swamps. |
| A | D | <i>Pericalymma ellipticum</i> | S | Leached sand with some clayey sands, lateritic soils. In elevated areas on seasonally swampy platforms. |
| C | | <i>Acacia alata</i> | S | Variety of soils. Near water, rocky hills, breakaways, salt pans, clay flats. |
| C | | <i>Acacia divergens</i> | S | Variety of habitats, often along watercourses or near swamps. |
| C | | <i>Boronia molloyae</i> | S | Sandy soils. Along creeks & streams, swamps. |
| C | | <i>Gahnia decomposita</i> | G | White sand, wet black sandy loam. Swamps, streams, seasonally wet flats. |
| C | | <i>Gahnia trifida</i> | G | Grey or white sand, clay, sometimes saline. Swamps, creeks. |
| C | D, E, W | <i>Lepidosperma squamatum</i> | G | Calcareous, peaty or lateritic sand, sandy clay, gravel. Dunes, swamps. |
| C | | <i>Lepidosperma tetragetrum</i> | G | Black peaty sand. Gullies, swamps, streams. |
| C | | <i>Taxandria linearifolia</i> | S/T | Loam, clay or sand, gravel, quartzite, laterite. Bordering swamps & watercourses. |
| D | E | <i>Babingtonia camphorosmae</i> * | S | Occurs mainly in heavy soils on the Darling Scarp |
| D | W | <i>Corymbia calophylla</i> | T | Red-brown clay loam, orange-brown sandy clay, gravel, grey sand over limestone, granite, laterite. Flats, hills, slopes, breakaways, wetlands, fringing salt marches, beside drainage lines. |
| D | | <i>Hakea prostrata</i> | S | Sandy soils, often over laterite, loam, gravel. Hillslopes, granite outcrops, coastal dunes. |
| D | | <i>Trymalium ledifolium</i> | S | Calcareous grey sand, clay, loam, gravel, granite, limestone, laterite. Ridges, outcrops, dunes. |
| E | | <i>Eucalyptus marginata</i> | T | Grey sand, clay or sandy loam, laterite. Hills, rises. |
| E | | <i>Kingia australis</i> * | M | Sand, sandy loam, clayey loam. |
| W | | <i>Acacia extensa</i> * | S | Often on sandy & sandy lateritic soils. Damp areas, along watercourses, near swamps. |
| W | | <i>Eucalyptus megacarpa</i> | T | Sand, sandy loam, limestone. Hills, near swamps & streams. |
| W | | <i>Hakea lissocarpha</i> | S | White, grey or yellow sand, sandy loam, granitic soils, laterite. |

(B)

| Species | SVT | Habitat Description | Stable (%) | Transient (%) | Non-refugia (%) | Total (%) |
|----------------------------------|------------|--|-------------|---------------|-----------------|--------------|
| <i>Acacia divergens</i> | C | Along watercourses and near swamps | 0.65 | 0.53 | 0.29 | 1.5 |
| <i>Acacia extensa</i> | W | Along watercourses and near swamps | 1 | 2.31 | 9.25 | 12.6 |
| <i>Astartea scoparia</i> | A & C | Loam, sand | 1.28 | 1.85 | 1.7 | 4.8 |
| <i>Babingtonia camphorosmae</i> | D & E | Heavy soils on the Darling scarp | 1.36 | 3.08 | 5.46 | 9.9 |
| <i>Banksia littoralis</i> | A & C | Along watercourses. Low-lying, damp areas | 1.36 | 3.33 | 3.83 | 8.5 |
| <i>Boronia molloyae</i> | C | Along watercourses and near swamps | 0.89 | 1.65 | 3.63 | 6.2 |
| <i>Eucalyptus patens</i> | A & W | Depressions, streambeds, valleys | 1.35 | 1.87 | 0.84 | 4.1 |
| <i>Eucalyptus rudis</i> | A & C | Wetter areas but not restricted (flats and hillsides). | 1.38 | 3.42 | 4.23 | 9.0 |
| <i>Gahnia decomposita</i> | C | Along watercourses and near swamps | 1.12 | 1.29 | 0.62 | 3.0 |
| <i>Gahnia trifida</i> | C | Along watercourses and near swamps | 1.09 | 0.86 | 0.31 | 2.3 |
| <i>Hakea prostrata</i> | D | Sandy soils, hillslopes, granite outcrops | 1.25 | 2.29 | 1.35 | 4.9 |
| <i>Hakea varia</i> | A | Sandy, loam. Seasonally wet flats. | 1.21 | 2.28 | 1.71 | 5.2 |
| <i>Hypocalymma angustifolium</i> | A, D, E, W | Along watercourses and near swamps. Flats | 1.15 | 2.22 | 2.07 | 5.4 |
| <i>Lepidosperma tetragetrum</i> | C | Along watercourses and near swamps | 0.66 | 0.86 | 1.94 | 3.5 |
| <i>Melaleuca preissiana</i> | A | Sandy soils, swamps | 1.37 | 3.24 | 1.69 | 6.3 |
| <i>Pericalymma ellipticum</i> | A & D | Leached sands. Elevated, seasonally swampy platforms. | 1.28 | 2.42 | 4.94 | 8.6 |
| <i>Taxandria linearifolia</i> | C | Along watercourses and near swamps | 1.37 | 2.02 | 0.84 | 4.2 |
| Total | | | 19.8 | 35.5 | 44.7 | 100.0 |

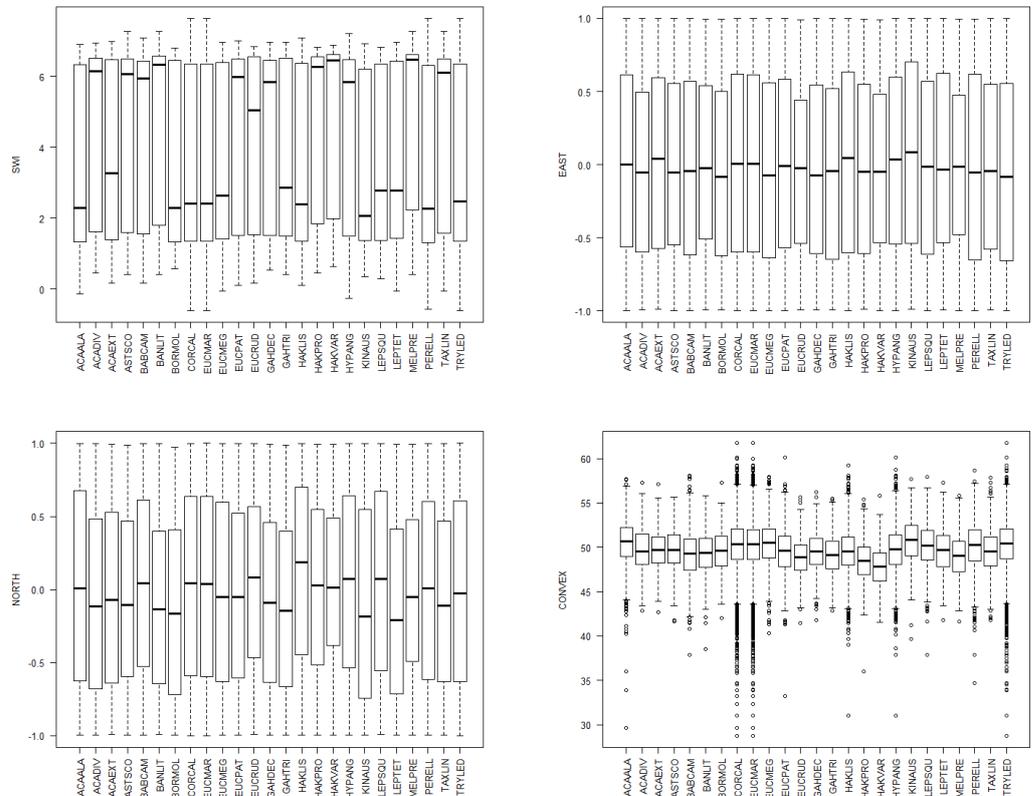
E2 Rainfall calculations for historic, current and future climate for southwestern Australia, 1970-2090. CSIRO sustainable yield reporting (CSYP), Climate Change in Australia website (CCA).

| Year | Period | Min AR (mm) | Max AR (mm) | Mean AR (mm) | Recorded/Projected Change (%), Source |
|----------|----------|-------------|-------------|--------------|---|
| pre-1975 | Historic | 600 | 1300 | 950 | NA, Havel (1975) |
| 2006 | Current | 525 | 1138 | 831 | -12.5% cf historic, CSYP, CCA |
| 2008 | Current | 510 | 1105 | 808 | -15% cf historic, Petrone & Hughes (2010) |
| 2030 | Future | 488 | 1058 | 773 | -7% cf current, CSYP |
| 2030 | Future | 494 | 1069 | 781 | -6% cf current, CCA |
| 2050 | Future | 488 | 1058 | 773 | -7% cf current, CCA |
| 2070 | Future | 483 | 1047 | 765 | -8% cf current, CCA |
| 2090 | Future | 478 | 1035 | 756 | -9% cf current, CCA |

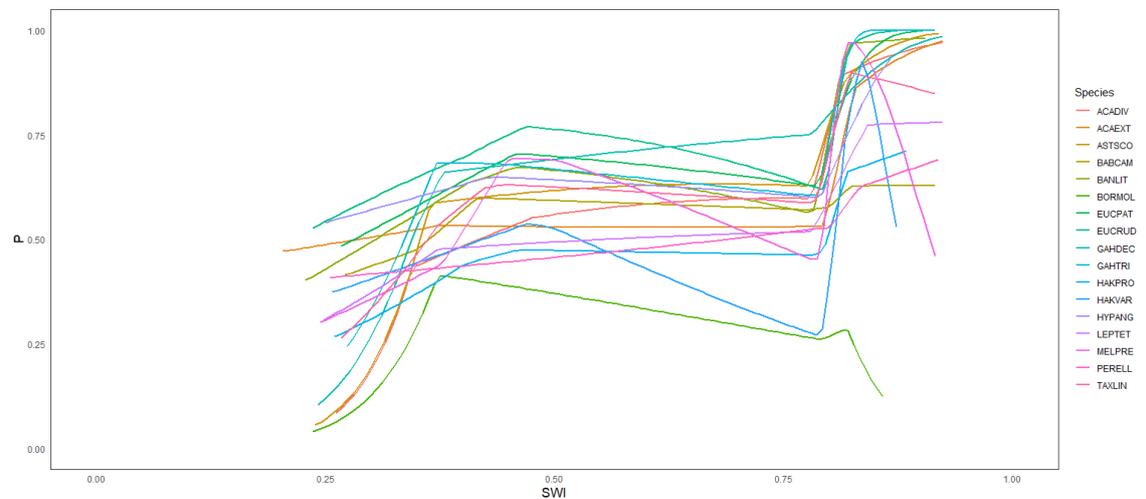
E3 Permutation importance (percent) for all variables used in modelling

| Species | CONVEX | EAST | SWI | LONGC | NORTH | TPI | TRI | VALDEP |
|--------------------------------------|--------|-------|-------|-------|-------|-------|-------|--------|
| <i>A - Astartea scoparia</i> | 0.766 | 25.88 | 27.07 | 14.94 | 17.11 | 2.888 | 11.35 | 0 |
| <i>A - Banksia littoralis</i> | 10.51 | 20 | 37.91 | 2.761 | 20.42 | 0 | 8.406 | 0 |
| <i>A - Eucalyptus patens</i> | 8.115 | 16.25 | 38.25 | 10.15 | 19 | 4.727 | 3.512 | 0 |
| <i>A - Eucalyptus rudis</i> | 21.23 | 19 | 27.1 | 4.327 | 11.5 | 1.711 | 15.13 | 0 |
| <i>A - Hakea varia</i> | 41.52 | 13.33 | 16.77 | 0.468 | 15.25 | 0.068 | 12.59 | 0 |
| <i>A - Hypocalymma angustifolium</i> | 4.89 | 14.46 | 58.27 | 0.372 | 9.187 | 0 | 12.82 | 0 |
| <i>A - Melaleuca preissiana</i> | 17.18 | 17.11 | 44.42 | 0.218 | 16.58 | 2.431 | 2.064 | 0 |
| <i>A - Pericalymma ellipticum</i> | 49.83 | 10.81 | 10.73 | 1.756 | 21.03 | 0.26 | 5.584 | 0 |
| <i>C - Acacia alata</i> | 35.6 | 3.341 | 2.183 | 0.726 | 2.573 | 0 | 8.761 | 46.82 |
| <i>C - Acacia divergens</i> | 1 | 10.88 | 39.22 | 19.98 | 7.411 | 6.716 | 14.79 | 0 |
| <i>C - Boronia molloyae</i> | 1.488 | 2.126 | 3.757 | 0.246 | 6.3 | 0 | 3.239 | 82.84 |
| <i>C - Gahnia decomposita</i> | 3.733 | 16.75 | 18.57 | 19.37 | 28.86 | 0.372 | 12.33 | 0 |
| <i>C - Gahnia trifida</i> | 12.53 | 20.16 | 16.41 | 7.263 | 36.39 | 0.246 | 6.998 | 0 |
| <i>C - Lepidosperma squamatum</i> | 18.3 | 6.557 | 1.716 | 0.002 | 23.83 | 7.804 | 24.82 | 16.96 |
| <i>C - Lepidosperma tetraquetrum</i> | 2.856 | 18.57 | 11.22 | 35.51 | 19.66 | 0.872 | 11.31 | 0 |
| <i>C - Taxandria linearifolia</i> | 9.184 | 10.88 | 35.77 | 20.11 | 19.14 | 3.534 | 1.385 | 0 |
| <i>D - Babingtonia camphorosmae</i> | 54.04 | 6.867 | 8.397 | 0 | 3.019 | 2.292 | 25.38 | 0 |
| <i>D - Corymbia calophylla</i> | 71.67 | 0.48 | 0.39 | 0 | 3.534 | 0 | 7.67 | 16.26 |
| <i>D - Hakea prostrata</i> | 50.03 | 3.008 | 14.22 | 0.315 | 5.365 | 0.817 | 26.25 | 0 |
| <i>D - Trymalium ledifolium</i> | 59.84 | 9.973 | 13.72 | 0.989 | 0.737 | 0.263 | 14.48 | 0 |
| <i>E - Eucalyptus marginata</i> | 76.6 | 0.285 | 0.031 | 0 | 0.5 | 0 | 10.31 | 12.27 |
| <i>E - Kingia australis</i> | 18.98 | 16.18 | 12.6 | 2.831 | 8.18 | 0 | 6.735 | 34.5 |
| <i>W - Acacia extensa</i> | 11.66 | 13.42 | 34.44 | 0 | 22.31 | 0 | 18.16 | 0 |
| <i>W - Eucalyptus megacarpa</i> | 16.91 | 21.08 | 8.521 | 19.6 | 15.87 | 0 | 18.02 | 0 |
| <i>W - Hakea lissocarpa</i> | 32.78 | 6.197 | 7.046 | 5.911 | 45 | 1.292 | 1.778 | 0 |

E4 Boxplots of the key predictors used in species distribution modelling show distinct difference in SWI and aspect (northness), but more subtle variation in eastness and convexity. SAGA wetness index (SWI), eastness (EAST), northness (NORTH) and convexity (CONVEX).



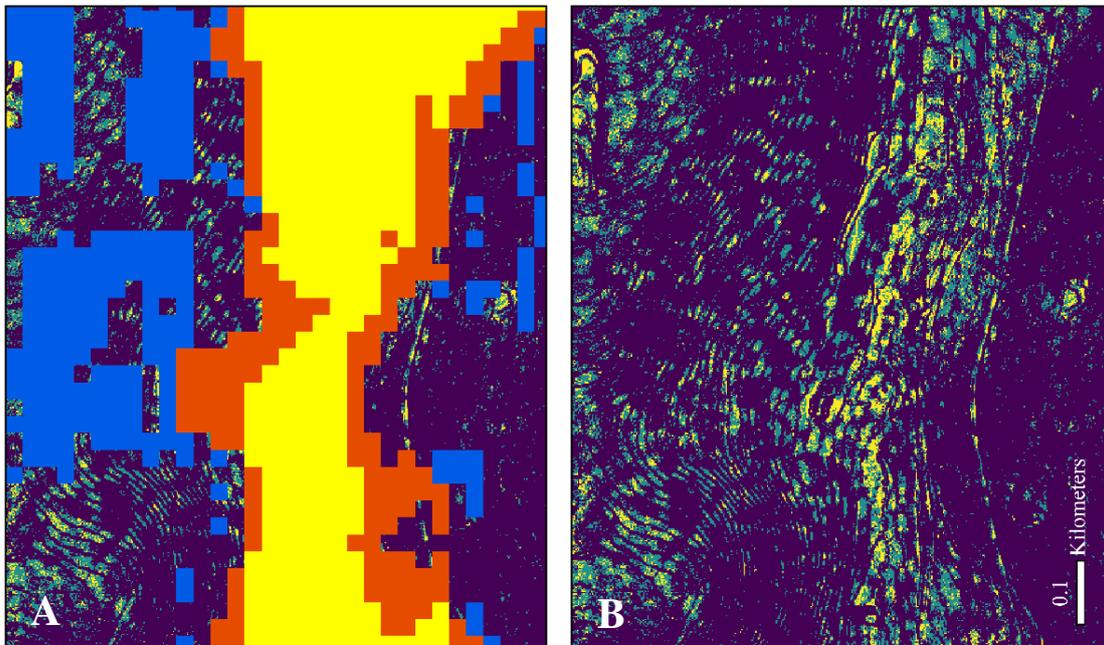
E5 Habitat suitability response curves (P) for the SAGA Wetness Index (SWI), for the species predicted to lose range from 2006 to 2090.



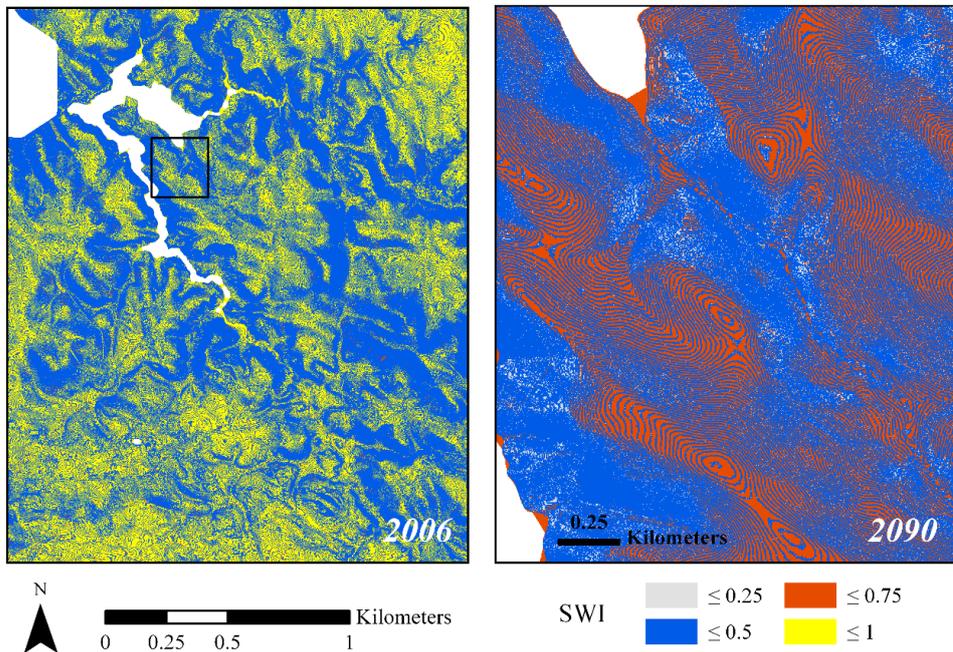
E6 The indicator species used in modelling each represent a particular site vegetation type, but occur across a larger number of types than expected.

| Modelled Species | Site Vegetation Type | | | | | | | | | | | Total |
|----------------------------------|----------------------|-----|------|----|-----|------|-----|-----|-------|------|------|-------|
| | A | C | D | E | G | P | Q | R | S | T | W | |
| <i>Acacia alata</i> | 101 | 224 | 53 | 8 | 5 | 420 | 1 | 17 | 453 | 105 | 129 | 1516 |
| <i>Acacia divergens</i> | 155 | 202 | 4 | | | 4 | | | 7 | | 31 | 405 |
| <i>Acacia extensa</i> | 35 | 42 | 25 | 2 | 7 | 43 | 14 | 2 | 101 | 92 | 56 | 419 |
| <i>Astartea scoparia</i> | 208 | 194 | 18 | | 4 | 15 | | 2 | 11 | 2 | 15 | 471 |
| <i>Babingtonia camphorosmae</i> | 78 | 84 | 383 | 8 | 56 | 132 | | 19 | 88 | 5 | 63 | 916 |
| <i>Banksia littoralis</i> | 392 | 273 | 103 | 3 | | 33 | | 3 | 70 | 19 | 106 | 1003 |
| <i>Boronia mollayae</i> | 86 | 89 | | | | 3 | 1 | | 5 | 11 | 8 | 203 |
| <i>Corymbia calophylla</i> | 337 | 537 | 939 | 74 | 174 | 4773 | 97 | 196 | 11099 | 5699 | 1098 | 25068 |
| <i>Eucalyptus marginata</i> | 442 | 615 | 1132 | 75 | 138 | 6300 | 100 | 218 | 12798 | 6287 | 1263 | 29415 |
| <i>Eucalyptus megacarpa</i> | 169 | 300 | 14 | 2 | 2 | 98 | | | 86 | 37 | 624 | 1335 |
| <i>Eucalyptus patens</i> | 579 | 576 | 227 | 3 | 4 | 56 | 29 | 5 | 132 | 128 | 727 | 2470 |
| <i>Eucalyptus rudis</i> | 115 | 42 | 32 | 1 | 1 | 4 | | 1 | 12 | 12 | 13 | 234 |
| <i>Gahnia decomposita</i> | 121 | 182 | 6 | | 3 | 3 | | 1 | 4 | | 8 | 329 |
| <i>Gahnia trifida</i> | 151 | 170 | 10 | | | 6 | 1 | | 6 | 5 | 22 | 371 |
| <i>Hakea lissocarpa</i> | 105 | 49 | 413 | 10 | 48 | 467 | 2 | 67 | 1903 | 332 | 202 | 3611 |
| <i>Hakea prostrata</i> | 57 | 6 | 250 | 1 | 1 | 17 | | 1 | 19 | 11 | 15 | 378 |
| <i>Hakea varia</i> | 149 | 8 | 48 | | 3 | 10 | | | 9 | 4 | 7 | 239 |
| <i>Hypocalymma angustifolium</i> | 289 | 185 | 711 | 35 | 126 | 822 | 8 | 115 | 1237 | 272 | 355 | 4163 |
| <i>Kingia australis</i> | 3 | 16 | 11 | 79 | 2 | 61 | 1 | 1 | 35 | 7 | 8 | 224 |
| <i>Lepidosperma squamatum</i> | 20 | 20 | 65 | 2 | 8 | 362 | | 6 | 516 | 144 | 45 | 1197 |
| <i>Lepidosperma tetraquetrum</i> | 129 | 262 | 2 | | 1 | 2 | 1 | | 9 | 7 | 24 | 437 |
| <i>Melaleuca preissiana</i> | 303 | 82 | 87 | | 1 | 19 | | | 23 | 4 | 40 | 560 |
| <i>Pericalymma ellipticum</i> | 30 | 6 | 147 | 4 | 29 | 72 | | 13 | 47 | 11 | 12 | 371 |
| <i>Taxandria linearifolia</i> | 515 | 767 | 57 | 8 | 2 | 36 | | 5 | 67 | 25 | 140 | 1622 |
| <i>Trymalium ledifolium</i> | 109 | 143 | 467 | 26 | 64 | 3755 | 6 | 74 | 5490 | 2140 | 512 | 12796 |

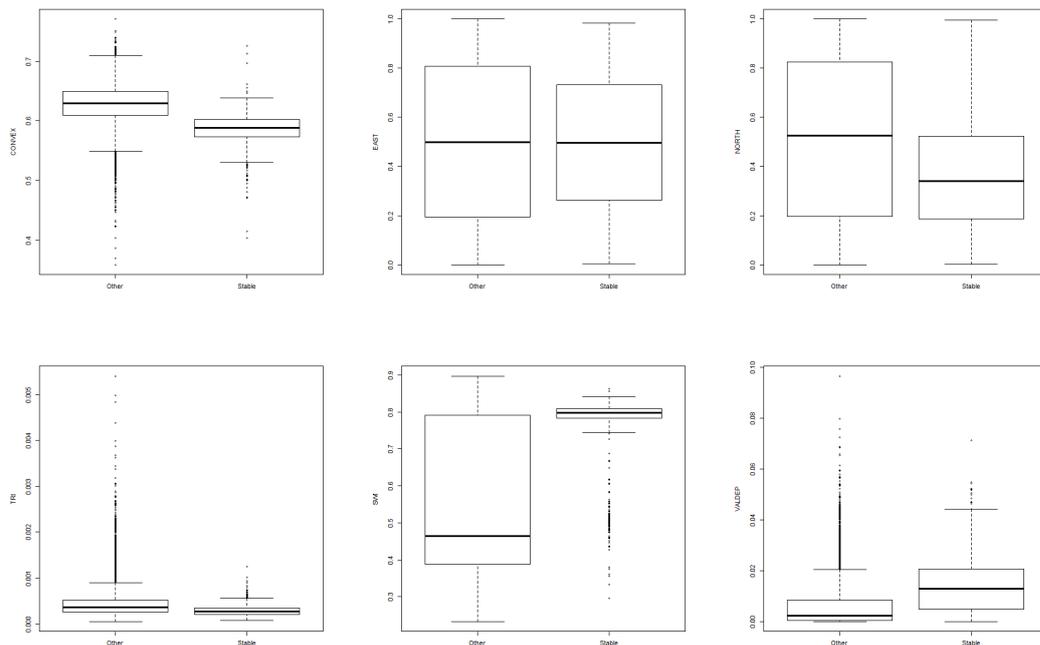
E7 The identification of refugia at two scales shows both broad and fine scale patterns and highlights the importance of micro-topography in the landscape. (A) community, 30 m scale; (B) species, 2 m scale.



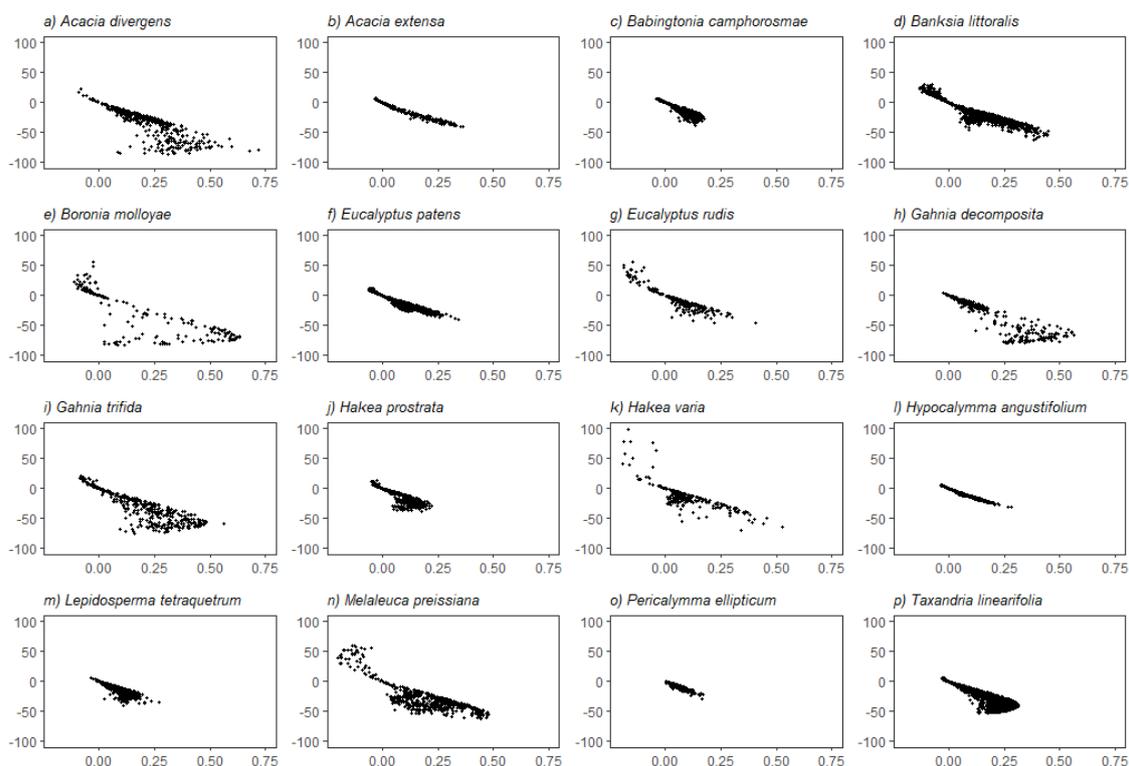
E8 SAGA Wetness Index (SWI) values for 2006 and 2090. Max SWI is projected to be 0.72 in 2090, indicating the possibility that no areas will provide stable hydrologic refugia under the RCP 4.5. future climate change scenario.



E9 Comparison of environmental predictors for stable vs other areas of refugia. Convexity = CONVEX, eastness = EAST, northness = NORTH, topographic roughness index = TRI, SAGA wetness index = SWI, distance to valley depth (VALDEP).



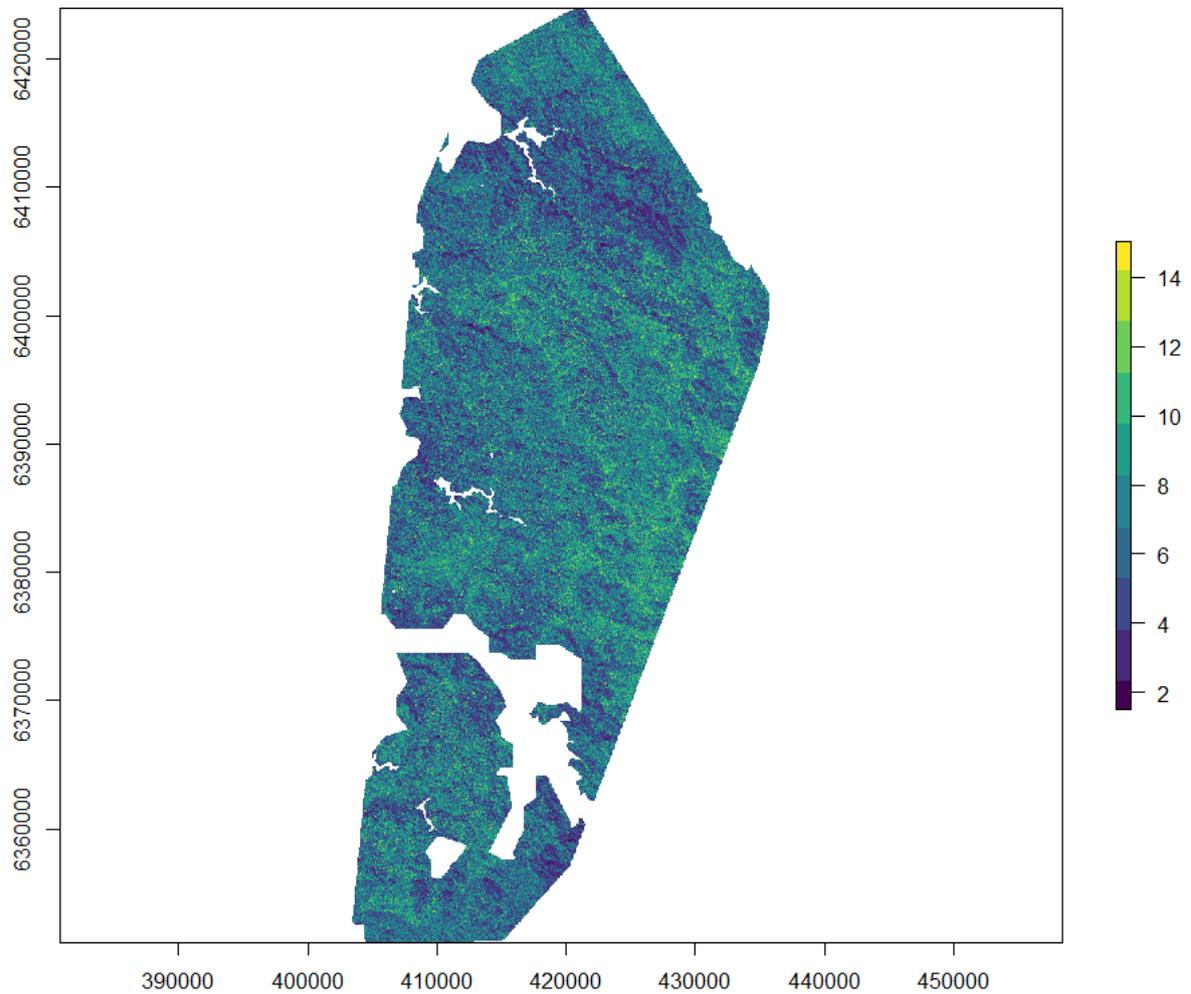
E10 Variation in raw (x-axis) vs percent (y-axis) change in habitat suitability for species predicted to contract in response to drying.



E11 Table with technical information on how the percent change in future rainfall was calculated. Weighted mean of average annual rainfall projections for the NRM super-region ‘Southern and SW Flatlands’ under RCP 4.5 for 2030, 2050, 2070 and 2090. Values were obtained from CSIRO’s Climate Futures website. Models (%): the number of GCM’s (global climate models) that agree on a particular trajectory; Weighted Value: model % multiplied by the rainfall category (e.g. 7% x -15%); Weighted Total: sum of weighted values; and Total Expected (Exp) Change in rainfall from baseline (%): Weighted Total divided by 100.

| Rainfall category (%) | 2030 | | 2050 | | 2070 | | 2090 | |
|-----------------------------|------------|----------------|------------|----------------|------------|----------------|------------|----------------|
| | Models (%) | Weighted Value |
| Much Drier (-15%) | 7 | -105 | 15 | -225 | 21 | -315 | 29 | -435 |
| Drier (- 10%) | 50 | -500 | 54 | -540 | 56 | -560 | 49 | -490 |
| Little Change (+/- 0%) | 40 | 0 | 25 | 0 | 19 | 0 | 16 | 0 |
| Wetter (+ 10%) | 3 | 30 | 4 | 40 | 4 | 40 | 4 | 40 |
| Much Wetter (+15%) | 0 | 0 | 1 | 15 | 0 | 0 | 1 | 15 |
| <i>Weighted Total</i> | | -575 | 99 | -710 | 100 | -835 | 99 | -870 |
| Total Exp Change (%) | | -6 | | -7 | | -8 | | -9 |

E12 Cumulative probability layer, showing the sum of species predicted occurrence in 2090, for taxa that are indicators of refugia (a positive correlation with SWI). See also Figure 5.5, which shows cumulative results based on binarised outputs.



E13 R scripts, maxent species distribution modelling.

Appendix F

F1 Conceptual, methodological and computing challenges – additional information.

The size and nature of the vegetation database that underpins this thesis brought conceptual, methodological and computing challenges. Conceptual challenges included the re-purposing of data collected for one purpose (systematic floristic survey and vegetation mapping) to another (an investigation of hydrological refugia). Generally, research is deductive, and the research question guides subsequent choices about data collection and analysis. The reversal of this process prompted the work for Chapter 2, to understand the specific characteristics of the data and judge its fitness for research purpose. From a general perspective this includes whether data are presence-only, presence-absence or abundance (Rondinini *et al.* 2006), the scale of collection and processes relevant to this scale (Elith and Leathwick 2009). For this project, as high-quality data was available for the larger region, the decision was made to set the taxonomic, geographic and conservation limits of the dataset. This context provided an operating framework for both interpreting results and ensuring they weren't applied too broadly, a key tenet of good modelling practise.

Methodological and computing issues included data-processing (Chapter 3 & 5); the selection of an appropriate algorithm for classification (Chapter 3), community (Chapter 4) and species modelling (Chapter 5); and visualisation of results (Chapters 3-5).

Data-processing issues (Chapter 3) occurred when creating high-order matrixes for further analysis – i.e. a Bray Curtis dissimilarity matrix for vegetation classification. For 30,000 plots, this exceeded Excel's column capacity (~16,000) and available ram on standard desktops (8-16 GB). Cloud computing and batch processing in R provided a solution before a non-hierarchical approach was taken. For remotely sensed data – one-meter resolution topographic layers for the study area in Chapter 5 were 15 GB each. Available computing resources (48 GB RAM on a virtual machine) were exceeded when using the 1 m DEM to calculate the SAGA topographic wetness index (SWI), as it holds several intermediate layers in memory (catchment area and slope, and modified catchment area). A process of re-sampling and conversion between formats (from 1 m to 2 m and ascii to tiff files) made a large difference to file size and enabled SWI to be calculated on the virtual machine. Two vs one-meter resolution layers also reduced the calculation time for each species distribution model (SDM) by ~75% (24 to 6-8 hours).

➤ Algorithm selection – Chapter 3 and 4

Algorithm selection formed a significant component of the work for Chapters 3 and 4. The vegetation classification literature is comprehensive, with many options available to cluster and reduce the dimensionality of data (Borcard *et al.* 2011). After a thorough investigation of the options available in R, including Roberts (2015) OPTPART and OPTSIL algorithms, ALOC in PATN was selected (Appendix C2). Reasons include short processing time for all plots (10-20 minutes); the ability to visualise results using group-based metrics (a dendrogram of groups, rather than sites); and that comparative work had been done to validate the ALOC algorithm (Belbin and McDonald 1993). While it performed well, the

processing time for OPTPART was significant in much smaller datasets (50 hours of CPU time, 424 samples and 209 species) (Roberts 2015). For the Alcoa dataset (30,000 samples, 436 species) this would have also been a barrier to using a cross-validation through multiple simulation runs (Appendix C3).

Algorithm selection for Chapter 4 was not as complex as in Chapter 3 but had significant implications for modelling. Multinomial (rather than binomial) logistic regression was used, as it simultaneously solves the predictor-response relationship for multiple (non-independent) classes, reducing the error across calculations. It also enabled the prediction of a single site vegetation type (SVT) per grid cell, eliminating the need for complicated post-modelling analyses to select a site (which was the case with the overlapping binomial results, Appendix D6, Appendix D7). However, the accuracy of the model was poor (accuracy = 0.44, kappa = 0.13) compared to individual binomial logistic regression models (mean accuracy = 0.81, mean kappa = 0.5; Appendix D8). Alternative analysis algorithms (e.g. Random Forest, Neural Networks, Generalised Dissimilarity Modelling) may improve accuracy and predictive power.

Additionally, low accuracy may be due to the dominance of type S in the dataset (42% of sites) ‘swamping’ results, however, equal sub-sampling did not alleviate the issue. More likely, is that the methods used to define the original SVT classification (Havel 1975a; 1975b) and the general overlapping nature of vegetation in the forest (Figure 4.4) make classification difficult. The indicator species for S, particularly *Eucalyptus marginata* and *Corymbia calophylla* are amongst the most wide-spread taxa and are frequently found with other indicator species complexes (Appendix E6) (Havel 1975a). Overall, misclassification is a function of the mixed vegetation patterns; difficulty in delineating widespread but specific indicator species and groups; and the dominance of several common species across elevational and wetness gradients in the forest (Figure 3.2, Table 5.2, Appendix E1, Appendix E6). The size of the dataset and variation in the subjective interpretation of SVTs by a large number of botanists during data collection (86) likely also contributed.

➤ Visualisation of results

The visualisation of results posed issues until variations on standard methods were found. This included plotting how groups of groups were related in cluster dendrograms, rather than groups of sites (Figure 3.4) and condensing multiple categories into interpretable classes for refugia mapping (Figure 4.5, Figure 5.5). Community modelling of refugia resulted in an array of potential changes in SVTs, with many permutations and combinations possible in the final results (Table 4.5, Appendix D3). Reducing classes based on each SVTs relationship to the SAGA topographic wetness index (SWI) (Table 4.3) enabled lumping into ‘wet’ and ‘dry’ groups (Chapter 4, 4.3.6) and subsequent classification as stable or transient refugia (Figure 4.5). A similar approach was used for species, with stacked SDMs and the proportion that overlapped used to define categories. These decisions influence how final refugia patterns appear in the landscape, but provide a practical way to visualise and relay results.

➤ Species distribution modelling and algorithm selection

The debate around the appropriate use of SDMs (and correlative modelling generally) prompted a cautionary approach (Elith *et al.* 2011; Guisan *et al.* 2013). Thus, the framework provided in Chapter 2 (parametrising the dataset within the context of widely available data for the region), enabled models to

be developed with data-limitations in mind. This included a focus on shrubs and trees and the exclusion of annual species from analysis (Figure 2.4, Table 2.1). Model extrapolation was also limited to the study area, based on the dominance of selected landforms and their influence on vegetation pattern (Table 2.2) (Elith and Leathwick 2009). This pre-processing led to a familiarity with the data that informed the selection of predictors at a relevant scale (meters vs kilometers) and therefore - the exclusion of available climate data from further analysis. This was a significant barrier to meeting project objectives until the streamflow elasticity equation was introduced. While additional criticisms linked to the projection of correlative models through time are not fully addressed by these precautions, it placed the work on a firm foundation for understanding the topographic drivers of plant-hydrological relationships in the forest.

After further reading of the SDM literature (Guillera-Arroita *et al.* 2014; Thibaud *et al.* 2014) other methods are likely better suited to a dataset of this size and structure than MaxEnt. For example, Thibaud *et al.* 2014 compared MaxEnt to three presence-absence methods and found a generalized linear model (GLM) with the probit link function, quadratic terms for each predictor and no interactions to have the best accuracy with a large sample size (n=500). MaxEnt outperformed other models in smaller sample sizes (n=100), but Guillera-Arroita *et al.* (2014) produced results that disagreed with this, and they argue for the use of GLM's in small datasets with presence-absence data. In hindsight, a GLM, or other model that takes advantage of presence-absence data may have led to better predictive performance. Additionally, final model results should have been aggregated up to 20 m resolution (e.g. mean value for a 20 m² grid cell), to mirror the scale that species data was collected at (plot size) – rather than producing maps at the scale of the LiDAR data (2 m).

- Hydrology: limitations of the topographic wetness index and streamflow elasticity estimator.

Streamflow elasticity was treated as a constant over the whole study area, which has implications for results. The original work that defined the rainfall-streamflow runoff relationship for southwestern Australia ('Elasticity' $E_p = 3$, where a 10% reduction in rainfall equates to a 30% decline in streamflow/runoff) was done at the catchment scale (Chiew 2006). However, raw data from the WA Water Corporation shows considerable variation in streamflow (discharge rate: m³/sec) from 1970 to 2010 (20-80% drop) (Figure 1.2, Chapter 1). As predictions did not incorporate variation in the rainfall-streamflow relationship, results will be incorrect at some locations (e.g. stream with especially high or low flow rates).

Ideally, a hydrological model based on the raw streamflow data would be used to calculate discharge rates for ungaged streams, and percent decline interpolated across the study area at the stream scale.

Hydrological modelling is a complex area of research within itself and outside of the scope of this study (Beven and Kirkby 1979; Devia *et al.* 2015; Gemitzi *et al.* 2017; Singh 2018; Visser *et al.* 2019; Feng and Beighley 2020).

Note: Figure 1.2, Chapter 1 shows percent decline (1970 to 2010) in minimum discharge rate (m³/sec) for gauged streams across southwestern Australia, based on daily flow data from the WA Water Corporation (<http://wir.water.wa.gov.au/Pages/Water-Information-Reporting.aspx>). Generalised linear models were fitted in R using averaged daily flow measurements.

F2 Attribution statement

A summary table stating the specific contribution by each author to each chapter is included below (SL - Sarah Luxton, GWJ - Grant Wardell-Johnson, TR - Todd Robinson, AS - Ashley Sparrow, LT - Lewis Trotter).

| Chapter | Contributor (%) |
|----------------------------------|---|
| 2 - Fit for purpose | |
| <i>Developed question/method</i> | SL (70%), GWJ (10%), LT (10%), TR (10%) |
| <i>Data analysis</i> | SL (70%), LT (20%), TR (10%) |
| <i>Wrote chapter</i> | SL (100%) |
| <i>Editing</i> | SL (80%), GWJ (10%), TR (10%) |
| 3 - Classification | |
| <i>Developed question/method</i> | SL (70%), AS (10%), GWJ (10%), TR (10%) |
| <i>Data analysis</i> | SL (80%), AS (15%), TR (5%) |
| <i>Wrote chapter</i> | SL (100%) |
| <i>Editing</i> | SL (85%), AS (5%), GWJ (5%), TR (5%) |
| 4 - Refugia community | |
| <i>Developed question/method</i> | SL (75%), AS (10%), GWJ (5%), TR (10%) |
| <i>Data analysis</i> | SL (80%), AS (10%), TR (10%) |
| <i>Wrote chapter</i> | SL (100%) |
| <i>Editing</i> | SL (85%), AS (5%), GWJ (5%), TR (5%) |
| 5 - Refugia species | |
| <i>Developed question/method</i> | SL (90%), AS (5%), TR (5%) |
| <i>Data analysis</i> | SL (90%), AS (5%), TR (5%) |
| <i>Wrote chapter</i> | SL (100%) |
| <i>Editing</i> | SL (90%), GWJ (10%) |

F3 Copyright permissions

I confirm that I have written copyright permission to reproduce the reserve system map presented in Appendix A1 and the Forest Ecosystem Units presented in Appendix C5(c).

Appendix A

From: Rayner, Martin <Martin.Rayner@DPaW.wa.gov.au>
Sent: Friday, 25 November 2016 3:19 PM
To: Banks, Geoffrey <Geoffrey.Banks@DPaW.wa.gov.au>
Cc: Sarah Luxton <s.luxton@postgrad.curtin.edu.au>; Tarrant, David <David.Tarrant@DPaW.wa.gov.au>; Miller, Jodie <Jodie.Miller@dpaw.wa.gov.au>
Subject: RE: DPWA reserve data (Map 2, Forest Management Plan)

Hi Geoff,

Many thanks for checking with us. I have no issues with Sarah including Map 2 from the FMP in her research documentation, and wish her well in her studies.

Regards
Martin

From: Banks, Geoffrey
Sent: Wednesday, 23 November 2016 3:16 PM
To: Rayner, Martin <Martin.Rayner@DPaW.wa.gov.au>
Cc: s.luxton@postgrad.curtin.edu.au
Subject: FW: DPWA reserve data (Map 2, Forest Management Plan)

Hi Martin.

Are you happy for the map attached to be used as part of Sarah's research paper (please see below) – it is from the current FMP report.

I have no issues with this but as the custodian of the FMP report process, I believe this should be approved (or not) by FMB.

As she has mentioned, DPaW will be acknowledged in the report.

Thanks
Geoff

Geoffrey Banks
Spatial Database Administrator (GIS)
Department of Parks and Wildlife (DPaW)
17 Dick Perry Ave Kensington WA 6151
Phone: 9219 9562
geoffrey.banks@DPaW.wa.gov.au

From: Sarah Luxton [<mailto:s.luxton@postgrad.curtin.edu.au>]
Sent: Wednesday, 23 November 2016 2:44 PM
To: Banks, Geoffrey
Subject: DPWA reserve data (Map 2, Forest Management Plan)

Hello Geoffery,

I'd like to use the attached (Map 2, from the southwest forest management plan 2014-2023) as part of a figure for my research, but am not sure whether there are any restrictions regarding re-use for academic purposes. Can you tell me if there would be any limitations to my re-using part of the map in a figure? It would be acknowledged and referenced properly.

Thank you, I appreciate your response.

Kind regards, Sarah

PhD Candidate | Department of Environment & Agriculture

Email | s.luxton@postgrad.curtin.edu.au

Web | <http://curtin.edu.au>



Appendix C5(c)

From: Martin Rayner <martin.rayner@dbca.wa.gov.au>
Sent: Thursday, 15 April 2021 4:35 PM
To: Sarah Luxton <s.luxton@postgrad.curtin.edu.au>
Cc: Carmel Easton <carmel.easton@dbca.wa.gov.au>
Subject: RE: Permission to Publish - Forest Ecosystem Units

Hi Sarah,

I hope your work is progressing well.

I confirm DBCA also grants permission to publish the map in your thesis.

Best wishes

Martin

Dr Martin Rayner
Manager Forest Management Branch
Conservation and Ecosystem Management Division
Parks and Wildlife Service
Department of Biodiversity, Conservation and Attractions
Ph: 9725 5927 | email: Martin.Rayner@dbca.wa.gov.au



Department of Biodiversity,
Conservation and Attractions



*We're working for
Western Australia.*

From: Sarah Luxton <s.luxton@postgrad.curtin.edu.au>
Sent: Thursday, 15 April 2021 8:41 AM
To: Martin Rayner <martin.rayner@dbca.wa.gov.au>
Subject: Re: Permission to Publish - Forest Ecosystem Units

[External Email] This email was sent from outside the department – be cautious, particularly with links and attachments.

Hi Martin,

Further to the below email - can I please also confirm that I have DBCA's permission to publish the map in the final version of my PhD thesis? It will be submitted electronically to the Curtin University Library and available through the library's website. The thesis chapter that refers to the map is effectively an earlier draft of the AJB manuscript that you reviewed.

Kind regards,
Sarah

PhD Candidate | Department of Environment & Agriculture

Email | s.luxton@postgrad.curtin.edu.au

Web | <http://curtin.edu.au>



Curtin University