Evaluating the existing Private Native Forestry Koala Prescription Map

Jane Elith, Honorary Professor, University of Melbourne

November 2024

Jane Elith was engaged by the NSW Natural Resources Commission (NRC) to conduct an independent expert evaluation of the Private Native Forestry Koala Prescription Map using readily available data. This technical paper, prepared by Jane Elith, should be read in conjunction with the summary document prepared by the NRC, which provides background to this work. The evaluation was conducted with input from a cross-agency technical review team and support from scientists in the then Department of Planning and Environment (DPE) (now the NSW Department of Climate Change, Energy, the Environment and Water). In addition, the NRC asked for feedback from Dr Scott Foster, a statistician with CSIRO, on the draft paper.

Summary

This paper presents analyses of the Private Native Forestry (PNF) Koala Prescription Map (the PNF Map). The PNF Map was released in May 2022. It maps areas of High Koala Habitat Suitability on private land; these areas are then subject to additional koala prescriptions under the PNF Codes of Practice. The analyses presented in this paper use climate suitability and presence-absence and foliage cover data for selected browse tree species from over 3,000 vegetation plots surveyed after the Map was completed. This paper presents the evaluation methods and results, and shows that the Map inputs generally have sensible, positive relationships with relevant independent data. However, it also shows there is a lot of variation in the Map inputs that cannot be explained by those data. The evaluation raises questions about why the relationship with independent data is not stronger.

This report evaluated PNF Map inputs that, whilst being cross-tenure, are very similar to the final PNF Map. The assigned task was to evaluate the PNF Map inputs with available data. This meant that the most direct measure of habitat suitability, the presence and absence (or abundance) of koalas could not be used, because the good quality data currently being collected is not yet available for use. Instead measures of browse tree species and climate suitability for koalas were used. The evaluation found that the Map inputs are not strongly related to observed counts of browse tree species, nor to the total cover of browse tree species, regardless of which of the three selected lists of koala browse tree species are used. The analysis suggests that the models and methods leading to the final PNF Map need to be revisited. Koala habitat is challenging to model well, but it is reasonable to expect that the relationship with browse trees and climate suitability would be stronger. By PNF code region, the evaluation of the Map showed better performance in the Southern NSW region than Northern NSW, with weakest performance in the Cypress and Western Hardwoods Forests region. There was insufficient data for meaningful analysis of the River Red Gum Forests region.

It is, therefore, worth revisiting the steps in making the PNF koala prescription map, including the creation of tree distribution models, the summary of those models in an index, the modelling of koala distribution, and the conversion of those distribution maps to a final binary product that intends to represent high suitability koala habitat. Even when each step is scrutinised and re-run in line with accepted good practice it is possible that the relationship with independent relevant data will still be noisy. However, re-visiting each step and using new data including recently improved spatial products may lead to meaningful improvements.

Introduction

The purpose of these analyses is to evaluate the existing PNF koala prescription map (PNF Map; see Box 1 for summary of how this was made). The evaluation aims to test whether the PNF Map (specifically, the cross-tenure input to it) is accurate in the sense that it tends to be "1" (high koala habitat suitability) in places that – according to independent data – are suitable for koalas. Any map based on modelling and generalisation will not be perfect, so errors are expected. Both types of error are important: failing to map "1" where koala habitat suitability is high, and failing to map "0" where the suitability is in fact low. The actual PNF Map is restricted to regulated private lands. Since relatively few data are available on these lands, the evaluation in this report focusses on the input to the PNF Map (see Glossary and data sections) and evaluates that input across all tenures. The logic behind this is that the inputs to the PNF Map will be driving its accuracy, and it is best to evaluate that input with as much data as can be collated.

For this evaluation independent data are required: data not used to make the map. The ideal data would be accurate records of the presence and absence (or the abundance) of koalas. Whilst good quality data of koala occurrence are now being collected, these are not yet available. Existing presence-only records of koala presence are not suitable for evaluation because they have no reliable information on where koalas are absent. My task here was to use available data to evaluate the PNF Map, so measures of the count of browse tree species or total foliage cover of those trees, and of climate suitability as modelled in a koala-focused eco-physiological model, will be used instead. As discussed later, this introduces some uncertainty in how to interpret the results, but its advantage is that it is achievable with current data.

Koalas may use multiple tree species for browsing, but are also known to occur in high densities in some habitats dominated by a single canopy species. Hence, both count of number of species and foliage cover of those species are useful measures. Presence-absence (PA) records for tree species at sites across the state are readily available, and these can be used to count the number of browse tree species at a site. Several thousand locations had been surveyed for full floristic records since the PNF Map was created. These recent data are used because the models used to create the PNF Map (Box 1) used all PA tree records available at that time to create the main predictor in the models, the koala tree species index. Re-using such data in evaluation would bias the results towards the modelled product. The more recent data are independent, so are the best available for evaluation.

In this analysis, these recent PA data are used to test whether the PNF Map tends to predict higher suitability where trees known as important browse species have been observed. At the same PA sites, foliage cover was also recorded, though with variation in methods across sites. This variation means that the raw recorded cover scores cannot be used for evaluation. Because we were keen to use some measure related to abundance of relevant trees, DPE staff transformed the cover values to a consistent measure across surveys (McNellie and Etminaniesfarhani 2023). Thus, foliage cover of browse trees important to koalas is used in this analysis in a second step, to test whether higher suitability for koalas is predicted at sites with higher foliage cover.

Note that the assertion here is that it is reasonable to expect that the general trend should be that the PNF Map values are higher where one or more browse tree species have been observed. Clearly there are a number of caveats here. The relationship between koala habitat and browse tree species is impacted by the vagaries of koala preferences for tree species, which vary across landscapes. Our knowledge of koala preferences is incomplete, and our ability to map the tree species is imperfect. Hence whilst the general trend should hold, we expect deviations from that trend regardless of whether the PNF Map is correct. This is discussed as results are interpreted.

Glossary

C1 / C2 species – a subset of a statewide list of tree species important to koalas. 135 tree species were allocated to categories C1, C2 … C5, listed in a spreadsheet called 'FavouredKoalaTrees' supplied by DPE staff in February 2023 and modified by the technical review team in May and November 2023. Note that this list is different to the list used to make the koala tree species (KTS) index (see below), which was used to create the PNF Map. In this report only those species considered most important browse trees (C1) and less important browse trees (C2) were used. Other lists (KTS species and PNF species) are also used in this evaluation to reflect the various efforts to define important koala tree species. These lists of species are all supplied (see main text).

KMR – koala modelling region – 9 regions in NSW

KTSIndex – Koala tree species (KTS) index – an index reporting the probability that at least one koala tree species (mostly browse, sometimes shelter) is present at a location. Based on models of tree distribution, and lists of tree species per KMR. Made by DPE staff in 2018/19.

KTS species – the species used in the KTSIndex. These vary by KMR region and are documented in DPIE (2019).

PNF Map (or the Map) – final regulatory Private Native Forestry (PNF) koala prescription map released in May 2022.

PNF1 / PNF2 species – primary (1) and secondary (2) koala tree species listed under the PNF codes. In this report these are simplified to one list per PNF code region, but in the code more detail (species per KMR within PNF code regions) is provided. The lists here were created by including all species mentioned for the relevant PNF region.

Map_input_raw – predictions from koala habitat suitability models, combined and transformed into binary (1/0) data to indicate whether predictions show high habitat suitability (1) or not (0). As summarised in Box 1, these predictions were made across the state, and across tenures. Steps to create the underlying models are detailed in DPIE (2019) and in Law et al. (2017), and methods for combining the modelled output documented in section 4.4 of NRC (2022). This Map_input_raw was available from DPE in mapped (raster) format at 30m grid cell resolution.

Map_input – the Map_input_raw, which has been customised so that sites outside of vegetation classified as "tree cover" or "tree cover matrix" always had a value of 0, as did sites in small patches (<2ha). This customisation reflects steps taken to create the PNF Map as documented in NRC (2022). However, this is a cross-tenure product.

Validation sites_initial – data available for evaluating the Map_input (3488 sites) These were sourced from vegetation surveys in the NSW BioNet Vegetation Information System (DPE 2023) by DPE staff. The sites were selected based on:

- date: surveyed after September 2018
- sites that were surveyed systematically – that is fixed area plot (between 0.04 and 0.1 hectares) where all observed species were recorded.
- excluded restricted and semi-restricted sites prior to using the data in building or validating models.
- excluded sites that were quarantined by BioNet data officers as 'suspect'.

Validation sites_final – data used to evaluate the Map_input (3184 sites)

Box 1: Description of how the PNF Map was developed

The summary document prepared by the NRC provides background on the reasoning and broad approach for creating the existing PNF Map. As discussed there, the Map was prepared by the Department of Planning and Environment (DPE), using existing modelled outputs. The purpose of this box is to briefly summarise the main inputs and steps in creating the PNF Map. All steps except those noted were developed and implemented by DPE, with consultation as documented in their reports.

The main inputs were modelled predictions of koala habitat suitability. These already existed as:

a. Predictions made by DPE and detailed in the Koala Habitat Information Base Technical Guide (DPIE 2019) - see section on Koala Habitat Suitability model (KHSM)
b. Predictions made by DPI detailed in Law et al. (2017). Only available for north-east NSW.

- Both these inputs were made using presence-only koala records and fitted in the Maxent software (Elith et al. 2011); only model b (the DPI model) was field validated.
- The models made by DPE were developed both regionally and statewide: six regional models focussed on the six most easterly koala management regions (KMRs, Fig A; north coast, central coast; south coast; northern tablelands; north west slopes; central and southern tablelands). Each of the six was developed independently, but using the same set of five predictor variables (koala tree species index, depth to bedrock; land soil capability index; cold air drainage; remotely sensed projected foliage cover). The 7th model was used for the 3 most western KMRs and only included climate variables; this was fitted across the full range of the koala but clipped to the 3 most western KMRs
- Model b (the DPI model) was only for NE NSW and used a set of predictors covering climate, vegetation, disturbance, topography and soil

The PNF Map was developed for each of the four PNF regions (Fig A) then joined into a statewide product. Since the KMRs can span PNF boundaries (Fig A) different combinations of the inputs detailed above were used.
The 4 PNF regions are:

1. Northern NSW
2. Southern NSW
3. River red gum forests (mostly the region north of Moama/Corowa/Albury, Figure A, though these are classified as River red gum forest regardless of where they are)
4. Cypress and western hardwood forests (western and northern parts of the state, Figure A)

For regions 1 and 2 (coastal PNFs)

- Inputs: regional DPE models and DPI model for NE NSW
- Each input used at 30m grid cell size and rescaled so predictions span 0 to 1.
- Each input divided into 6 classes with highest classes representing highest koala habitat suitability (details in NRC 2022).
- For pixels in the northeast where the DPI model overlapped with DPE models, highest class across the two products was allocated.
- To align the final map with modelled plant community type (PCT) boundaries, the PCT polygons from the state vegetation type map (SVTM; https://datasets.seed.nsw.gov.au/dataset/nsw-state-vegetation-type-map) were used as follows: in each polygon, all pixels of the koala habitat models were extracted and a median value estimated from the class values 1 to 6. That median value was assigned to the whole polygon.
- Classes 5 and 6 were considered high koala habitat suitability and were retained (i.e. given a value of “1”). Non-retained land can be considered “0”.

(cont’d over page)

For regions 3 and 4:

- Inputs: regional and statewide models from DPE. Statewide models only used where regional models did not not exist (i.e. western-most regions)
- Rescale and make 6 classes as above
- Only retain class 6 (because class 5 on inspection included too much marginal habitat) – i.e. class 6 becomes “1” and the rest, “0”

Insufficient PCT mapping to allow alignment with SVTM

The outputs from above were combined then masked:

- All non-native vegetation (5m cell size) removed (i.e. considered “0”).
- Remove various regulated land so map is only focused on areas where PNF Map is relevant
- Remove all non-connected patches (direct connection at 5m cell size) of koala habitat <2ha in size – removes scattered trees and unconnected small patches – i.e. these are also considered “0”.

The final product can be considered a binary map where 1 indicates predicted high quality koala habitat.

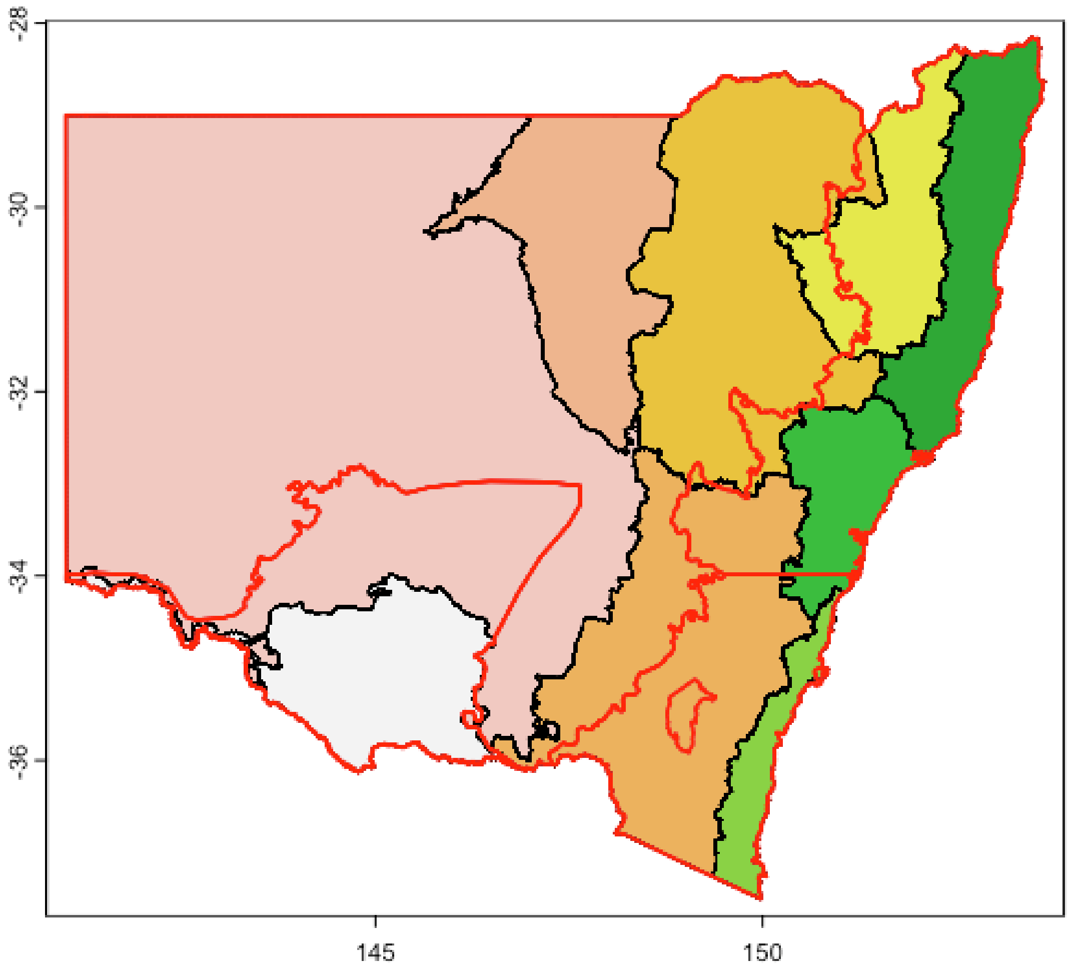


Figure A: KMR regions (filled, with black outlines) compared with PNF regions (red outlines).

Data used for the analysis

1. Map_input_raw – See Glossary. This is a binary map (Fig 1) that shows high koala habitat suitability (1) or not high suitability (0). It is based on koala habitat suitability models (details in Glossary, and Box 1).

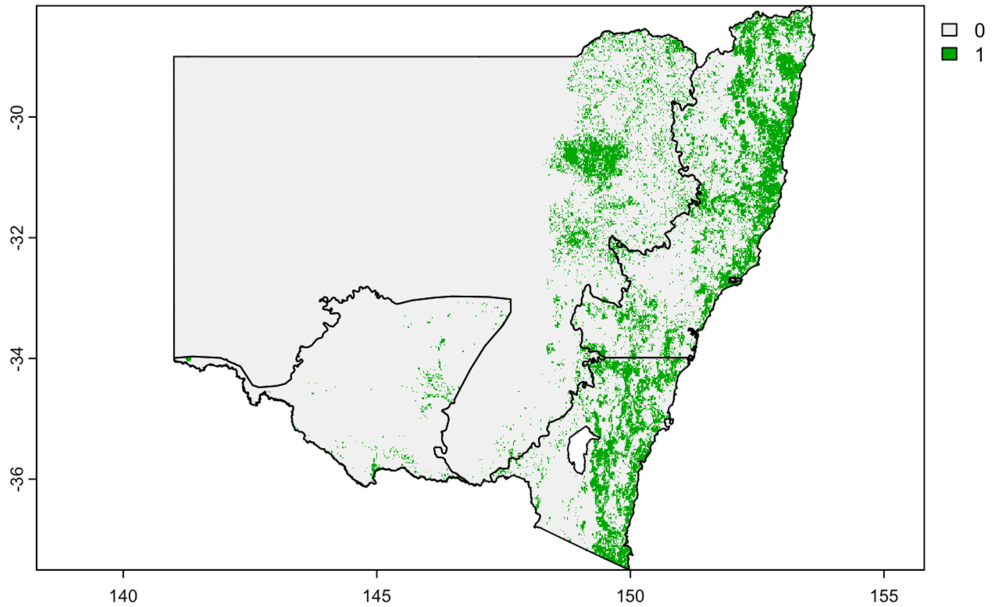


Fig 1. Map_input_raw. '1' (green shading) indicates the area has been mapped as high koala habitat suitability; '0' (grey) indicates not high suitability. Black lines show PNF regions.

2. For this analysis independent data are required – i.e. data not used in any modelling for the PNF Map. Of existing datasets, the most appropriate data identified were 3,488 sites where full floristics were recorded, surveyed after September 2018, which is after the creation of the Map_input_raw. These will be called Validation sites_initial (see Glossary). These sites are spread across the state (Fig 2). As detailed in the Methods, Results and Discussion section, the full set of 3,488 sites were checked for (a) whether they are in locations with climates that might support koalas; (b) whether they are likely to have been impacted by fires prior to survey; and (c) their inter-site distance. Any required adjustments were made as described in the Methods section.

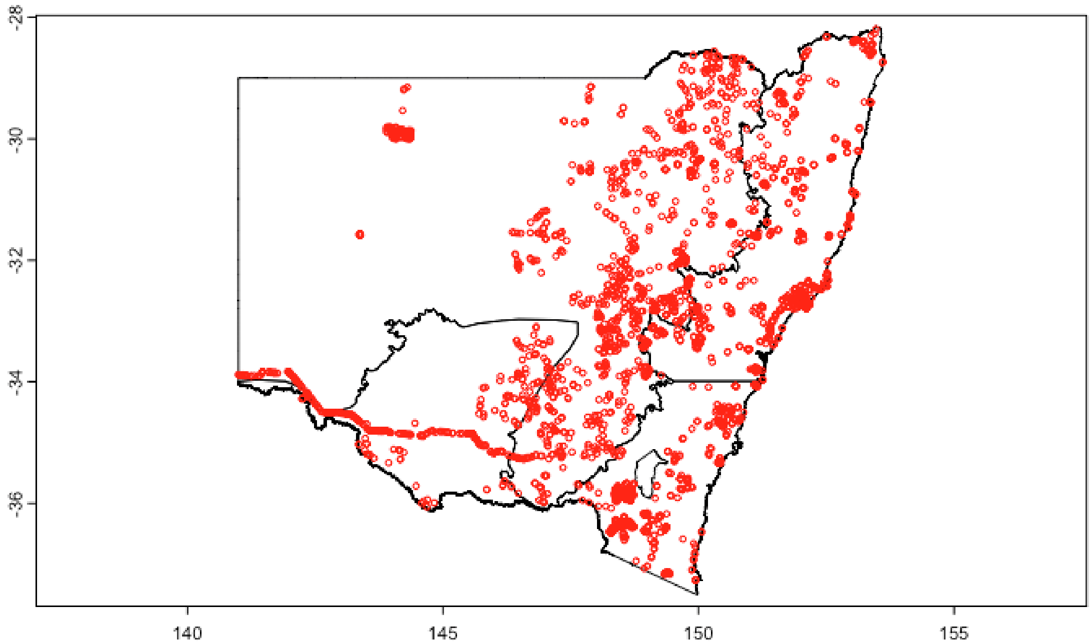


Fig 2. Distribution of 3488 validation sites across NSW. Black lines show PNF regions.

3. The full floristic data at the validation sites was used to quantify what browse tree species are present at each site. Identifying which tree species are important browse species for koalas is a complex task, particularly since the focus here is state-wide. In this analysis 3 lists of tree species were used:

   a. Based on a list provided by DPE staff and described in the glossary, the most important browse tree species (C1) and less important browse tree species (C2). These are referred to as C1 and C2 species in this report.

   b. Browse trees (primary and secondary) named in the PNF codes. In this analysis species identified in each PNF code region (Fig. 1) are used (4 regions: Northern, Southern, River Red Gum, Cypress & Western Hardwoods). These are referred to as PNF1 (primary browse trees) and PNF2 (secondary browse trees) species in this paper.

   c. Tree species contributing to the koala tree species index (KTS index) that was used by DPE in their koala habitat suitability modelling (details of modelling in DPIE (2019)). These are referred to as KTS species in this report. KTS species vary by koala modelling region (KMR; Fig.3). The KTS index was made prior to September 2018, so records of the KTS species at the validation sites are independent of the modelling behind the PNF Map.

   Note that the DPE koala habitat suitability modelling is not the only modelling input to the Map_input_raw: the DPI map of Law et al (2017) for northeast NSW was combined with that of DPE, for that area (details in Box 1). Due to the complexity of dealing with that area separately to the rest of the state, and noting that the DPI model used forest type mapping rather than defined species lists, the same approach to validation was used in the Northern NSW PNF region as for the rest of NSW.

   Species lists for a to c are available in Attachment 1.zip (see the files with names beginning "species" in the input folder). Since full floristic information is available from the validation sites, the number and foliage cover of tree species falling into lists a to c above can be calculated at each site. These are independent observations, not used in any modelling, and so constitute an independent validation dataset.

4. Also included in the analyses below is KTS index, the koala tree species index developed for the DPE koala habitat suitability modelling that was the major input to the Map_input_raw (Fig 1). This KTS index is based on models of tree distribution, which model the probability of presence of tree species at sites. KTS index values at the validation sites used in this analysis are based on predictions from those models, and express the probability that at least one tree species relevant to koalas is present (see Glossary). Details of the KTS index are in DPIE (2019). The KTS index was one of five predictors used in most of DPE's koala habitat suitability models, and was the most important predictor of koala habitat suitability. Because it was used in the models underpinning the PNF Map this will not be a fully independent evaluation of the PNF Map and its Map inputs (raw and customized; see glossary), but analyses using it may shed light on validation results.

5. Since the suitability of the climate for koalas might mediate the relationship between Map inputs (raw or customized) and number of browse trees, we included a koala climate suitability layer created from a mechanistic model[1] (Briscoe et al 2016) (i.e. based on the physiology of the animals,

[1] Koala climate suitability index mapping from Briscoe et al (2016) was used for this purpose; using the 'high' scenario - presuming koalas always have access to eucalypt leaves with high (66%) leaf water content.

not on their known locations) – Fig 3. Climate suitability index values of less than 0.25 were considered unsuitable for koalas, unless access to permanent water was available.

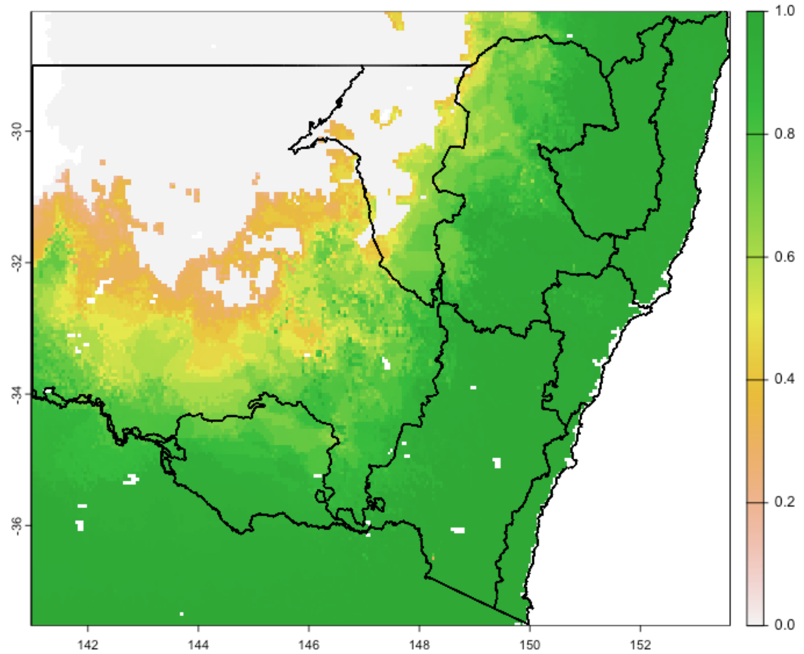


Fig 3. Koala climate suitability from Briscoe et al. (2016), using the "high" scenario (uses a high foliage water content to delineate non-suitable areas - this is most conservative). Black lines show KMR regions. '1' indicates most suitable and '0' least suitable.

Methods, Results and Discussion

Note: all data manipulation and modelling was done in R version 4.2.2 (R Core Team, 2022). Code is available along with inputs and outputs for that code (see Attachment 1.zip).

a. <u>Assessing the 3,488 sites in the initial validation dataset.</u>

The climate suitability filter (i.e. climate suitability < 0.25, which are also more than 500m from a permanent river) applied to the 3,488 Validation sites_initial resulted in the loss of 104 sites, therefore leaving 3,384 sites.

There were 17 sites that had a climate suitability < 0.25 but were within 500m from a river/permanent water source. These sites were retained, but the climate suitability value was changed to a value above the threshold of 0.25, because the presence of water will help koalas to physiologically adapt to high temperatures. On consultation with Dr Natalie Briscoe, the author of the climate suitability map, values at these sites had 0.35 added to their current value.

A fire filter was applied, on the premise that some of these sites might have been so badly burnt pre-survey that the target trees would not be recognisable. For this we used Fire Extent and Severity Mapping (FESM) for the year 2019/2020 (DPIE, 2020), covering the worst and most extensive fires. Any site surveyed after July 2019 with a FESM severity value of 4 (high) or 5 (extreme) was removed from the dataset on the assumption that canopy scorch or consumption would make it difficult to identify tree species. This led to the loss of 40 sites, with 3,344 sites remaining.

Three sites had a "no data" value for the Map_input_raw, and these sites were removed, leaving 3,341 sites. Finally, to avoid sites unlikely to be independent, these 3,341 sites were thinned to only include sites at least 100m apart. For sites closer than 100m, the retained site was randomly selected. This left 3,184 sites, forming the final validation dataset. At these 3,184 sites, the Map_input_raw value showed 877 sites with a value of 1 (high koala habitat suitability), and 2,307 with a value of 0 (not high habitat suitability) (Fig 4).

b. <u>Changing the Map input_raw value based on additional filters used in the final mapping</u>

A version of the PNF Map that included all tenures was not available, so the evaluation had to start with the Map_input_raw and replicate steps carried out by DPE to create the final PNF Map (Box 1). The steps described here are implemented to ensure that the product being evaluated here is as close to the PNF Map as possible. In making the PNF Map, the Map_input_raw was customized to exclude grid cells in vegetation not classed as "tree cover" or "tree cover matrix" (see Box 1). A grid cell size of 5m was used in that calculation, since the vegetation information was available at that grain. That was repeated here. Of the 3,184 Validation sites_final, 410 were in vegetation classes other than "tree cover" or "tree cover matrix" (as judged by the categorisation from the 5m native vegetation type raster). Of these, only 24 had a Map_input_raw value of 1. This was changed to 0, in line with the steps used to make the PNF Map (Box 1). The implication is that habitat suitability at those sites is not expected to be high.

Two other filters were used in making the PNF Map:

- small patches (less than 2ha in size) were removed, because these would not be subject to regulation by the PNF code. This filter was not applied in this evaluation because the evaluation is about the accuracy of the prediction of high value koala habitat, not about what will be regulated.
- The Map_input was clipped to only regulated private land. This filter was not applied here, because it would reduce the dataset by over 50%, and because the aim is to evaluate the strength of the relationship between the Map_input and observed koala browse tree species (which should be tenure-independent).

The final set of 3,184 validation sites are distributed as shown in Fig. 4. Note that 235 of these sites have "NA" for climate suitability (most being along the coast, where the climate suitability layer did not completely reach the true coastline). This lack of data is acceptable since the modelling method being used in this analysis can infer values for missing predictors. At these 3,184 sites, Map_input values were: 2,331 sites: "0" and 853 sites "1".

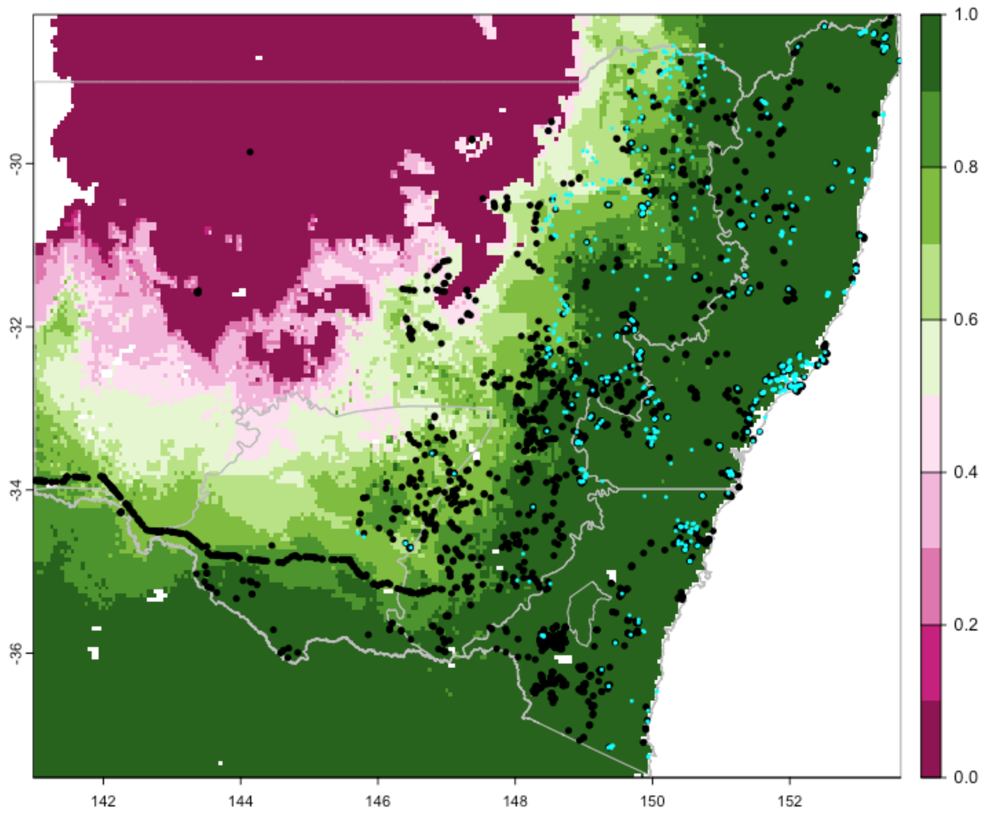


Fig 4. Final set of sites used for validation (3,184 sites; black for Map_input = 0; aqua for Map_input = 1) shown against koala climate suitability. Pale grey lines show PNF regions.

c. *Presence and cover of browse tree species at sites*

As mentioned in the section "Data used for the analysis", the full floristic data available at the 3,184 validation sites can be used to count how many browse tree species were observed at each site. The three browse tree lists described earlier (C1 / C2 species, PNF1 / PNF2 species, KTS species) were not identical. The full datasets are available in Attachment 1.zip (see files beginning "valid_dat" in the "output" folder, and code and inputs for their construction). C1 and C2 species are lists for the whole state, and include 31 C1 and 50 C2 species. PNF primary (PNF1) and secondary (PNF2) lists vary by PNF region (Northern, Southern, River Red Gum, Cypress & Western Hardwoods), with 8 to 13 PNF1 tree species per region and 20 to 50 PNF2 species per region. KTS species are specified separately for each of the nine KMR regions (Fig.3), with number of species varying from 5 to 21, with a mean of 13. For foliage cover, the sum of cover values for all relevant observed species was calculated at each site.

The different tree lists will be used in different evaluations (i.e. not at the same time) because they are alternative representations of trees suitable for koalas. To quantify the similarity between the counts of species and cover values across the different lists, Table 1 shows correlations between counts of number of species per site, across the 3,184 sites. It compares counts of browse tree species based on the different lists. For instance, in the top left cell the comparison is between the list of counts of C1 species at each of the 3,184 sites, and the list of counts of PNF1 species at each site, summarized as a Spearman correlation coefficient. The tabled comparisons include sums across lists from the same source (e.g. C1 + C2) because some lists (e.g. KTS) do not discriminate between these. All correlations are positive. Whilst some correlations are quite weak (e.g. C2/PNF2 (0.39), C2/KTS (0.25)) others are strong (e.g. PNF1+PNF2/KTS (0.71), C1+C2/ PNF1+PNF2 (0.83)). The variation in the strengths of these correlations demonstrates that the lists have some similarities but also notable differences. Hence one would expect the relationships between the Map inputs and the number of observed koala browse tree species to vary depending on the list used. Table 2 shows comparable data for the sum of foliage cover at each site. Overall, similar patterns were seen in correlations based on the cover data, though the correlations with the KTS lists were consistently lower for the cover data compared with those from the counts of species.

Table 1. Correlations (Spearman correlation coefficient) between counts of number of browse tree species recorded at the 3184 validation sites, for different tree species lists (row and column headers)

	PNF1	PNF2	KTS	PNF1+PNF2
C1	0.54		0.68	
C2		0.39	0.25	
PNF1			0.63	
PNF2			0.52	
C1+C2			0.66	0.83
PNF1+PNF2			0.71	

Table 2. Correlations (Spearman correlation coefficient) between sums of cover of browse tree species recorded at the 3184 validation sites, for different tree species lists (row and column headers)

	PNF1	PNF2	KTS	PNF1+PNF2
C1	0.48		0.65	
C2		0.37	0.14	
PNF1			0.56	
PNF2			0.43	
C1+C2			0.59	0.83
PNF1+PNF2			0.62	

Table 3 summarises for all lists both the counts of number of tree species across sites, plus the range and mean of foliage cover values. The summary shows that most sites have no browse tree species observed at them (44% to 76%), but 22% to 39% of sites have 1 browse tree species (primary or secondary), and 2% to 13 % of sites have 2 browse tree species (primary or secondary). The maximum number of primary browse tree species per site is 4 (C1) or 3 (PNF1), and secondary is 4 (C2 & PNF2). The maximum number of KTS species per site is 4. Across the sites foliage cover of the relevant species varied from zero to 100%, with mean across sites always less than 10% (Table 3 and Fig 5). The foliage cover of browse trees was strongly positively correlated with the number of browse tree species, with Spearman correlations all above 0.88 (Table 4). Spearman correlations allow for non-linear relationships between the two datasets, so this shows that as number of tree species increases, their cover usually increases also.

Table 3. Proportion of the sites with 0, 1, 2 etc browse tree species recorded, plus range and mean of foliage cover values. Number of browse trees shown in top line (columns 1 to 6) plus cover estimates, datasets in left column.

dataset	0	1	2	3	4	5	6	min cover	max cover	mean cover
C1	0.509	0.369	0.108	0.013	0.001	0	0	0	80.00	6.97
C2	0.587	0.296	0.095	0.019	0.003	0	0	0	100.00	5.97
PNFprimary	0.759	0.216	0.022	0.003	0.000	0	0	0	80.00	2.69
PNFsecondary	0.440	0.392	0.133	0.030	0.005	0	0	0	100.00	8.83
KTS	0.512	0.346	0.124	0.016	0.002	0	0	0	80.45	6.82

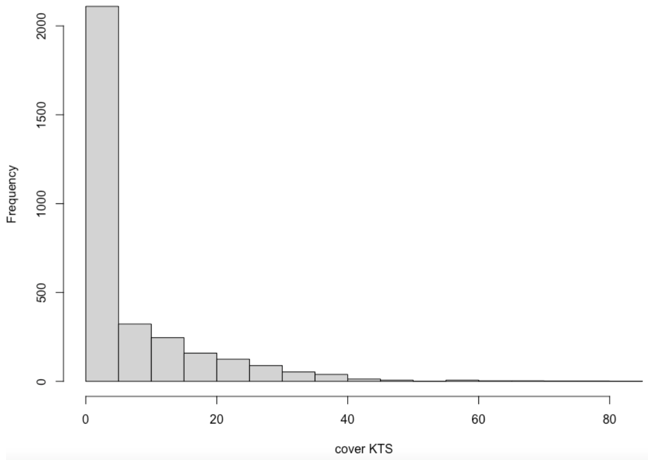


Fig 5. Example of the distribution of cover values across the 3184 sites, this example for cover of KTS species

Table 4. Spearman correlation coefficients between number of browse tree species and total cover of those species, across the 3184 sites.

Species list	Correlation between number of browse tree species and cover of those species
C1	0.92
C2	0.95
PNF1	0.99
PNF2	0.88
KTS	0.92

Finally, most climate suitability values in the dataset are high (>0.95; Fig. 6), also shown mapped in Fig 4 (see the darker green areas).

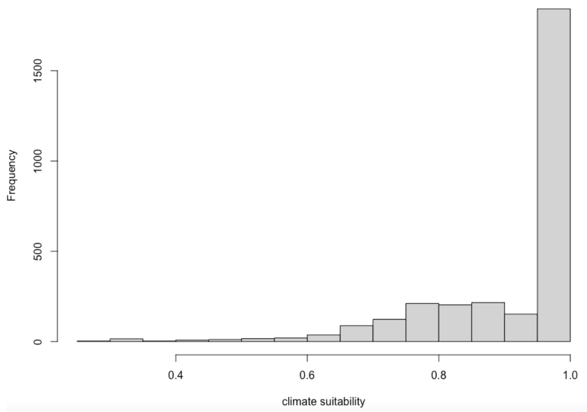


Fig 6. Distribution of climate suitability values in the validation data (3184 sites)

d. *<u>Modelling relationships between map inputs and number or cover of browse tree species.</u>*

We use a regression model to model this relationship:

Map_input (1/0) ~ number (or total foliage cover) of browse tree species + climate suitability

The response is to the left of the tilda (~) and the predictor or explanatory variables to the right. Often a regression model like this has an observed response – e.g. the presence or absence of a koala. Here the response is not observed, but modelled. The reason for fitting the model this way is that the focus is how much of the variation in the Map_input can be explained by predictors such as number of browse tree, and not the reverse (for instance, we expect there is variation in the floristic data that is not relevant to koalas, and that variation is not of interest here). The modelled Map_input aims to identify sites with high habitat suitability for koalas (Box 1). Hence the model we are fitting here is asking: when tested at sites not used in creation of the Map_input, does the Map_input bear an ecologically sensible and strong relationship with number or total cover of the browse tree species at the site, adjusted for the effect of climate suitability? Since across NSW one would expect browse trees and climate suitability to be major determinants of koala habitat suitability, this is a sensible question to ask. Indeed, the tree species index used in making the regional DPE models was always the most important predictor in those models (DPIE 2019 Appendix 8.1). The relationship fitted here for evaluation would not be expected to be without noise, since there are many other things (connectivity, past disturbances, nuances of local climate and soil type) that will also impact suitability for koalas.

Where lists differentiated between primary and secondary browse tree species, counts (or cover) of primary and secondary species are included separately (e.g. for the C lists based on presence or absence of trees, number of browse tree species becomes number of C1 species + number of C2 species). This enables the model to weight the contribution of the primary / secondary species separately. In any one model only one list of browse tree species (C or PNF or KTS) will be used, since each is independently trying to define the important tree species for koalas. In some of the analyses number (or cover) of browse tree species will be replaced by the KTS index, which is predicted rather than observed, and – as mentioned in the Data section – is not independent of the response. Climate suitability is in red because the effect of leaving it out will be explored. The formula shows "+" signs which implies no interaction between the terms to the right of the ~, but in some models interactions will be allowed to explore whether the modelled relationships with number of browse trees varies with climate suitability.

In this analysis boosted regression trees (BRT) (Elith et al. 2008) are used, a machine-learning version of logistic regression which is useful for its ability to automatically find the shape of the relationship between the response (Map_input (1/0)) and the predictors, and to find interactions. It can also deal with missing data in the predictors. BRT models by default will fit functional forms that are as complex as are needed to explain the data – i.e., instead of being constrained to linear, quadratic etc fitted functions as in parametric models, BRT fitted functions can be highly complex, if the data show complexity. However, it is possible to constrain a fitted function to be monotonic (i.e. always increasing or always decreasing) if required. The main text of this evaluation focusses on models with an enforced monotonic response to climate, but alternatives are presented in an appendix. The logic for the monotonic response is: one expects that sites more suitable climatically for koalas (as modelled by the eco-physiological model of Briscoe et al. 2016) should provide a higher habitat suitability for koalas. Once that is accounted for, the other important axis of suitability for which evaluation data are available is browse tree availability. After controlling the response to climate in an ecologically sensible manner, the response to trees is allowed full flexibility.

Analyses were run in R (code supplied in Attachment 1.zip).

The main analysis focusses on relationships across the full suite of sites – i.e. statewide, within areas at least somewhat climatically suitable for koalas. Correlations between predictors in each model (i.e. between the browse tree measures and climate) are weak and will not impact interpretation of model results in any substantial way (Appendix A1). Analyses for individual PNF regions are in Appendix A4. The following results are presented:

- Model details and evaluation

- Number of trees – BRT models of the type used here are ensembles of many regression trees. Since BRT model building has a stochastic element, number of trees in the final model needs to be at least 1,000 for a stable BRT model (i.e. for a model that won't change much from run to run, on the same dataset). Hence number of trees is reported.
- Tree complexity – where it is >1, interactions are allowed.
- Monotonicity – this reports if the fitted function was constrained to be monotonic. This was applied to climate suitability in some models to compare results with an unconstrained model.
- AUC (abbreviation for area under the receiver operating characteristic curve). AUC quantifies the discrimination of the model – i.e. its ability to predict higher at sites where the response (Map_input value) is 1. AUC was estimated using a 10-fold cross-validation (i.e. on data held out from model fitting, thus avoiding false optimism about overfitted models). AUC ranges from 0 to 1, with values of 0.5 implying the model does no better than random. Values greater than 0.75 are viewed as good discrimination (Hanley and McNeil 1982)
- Percent deviance explained (%DevExpl) – this shows how much of the variation in the response is explained by the model. In the experience of the author, for binary ecological data, values above 30% would be viewed as reasonable explanation of the data. Again, this was estimated using a 10-fold cross-validation.
- Variable contributions – the extent to which each predictor helps in explaining the data. For each model, this will sum to ~ 100%, because it is estimating how much of the variation that is explained by the model is explained by each predictor

- Plots of fitted functions– these show the shape of the fitted functions in the models. The y axis shows the response, scaled so that lower is less probable, higher is more probable. For any one model, the y axis in each plot is on the same scale. X axes vary because they represent different predictor variables.

Table 5: Model details and evaluation for all models with a monotonic response to climate, and one or two predictors summarising observed (C, PNF, KTS) or predicted (KTSIndex) tree species. All models were fitted with a learning rate of 0.001 and a tree complexity of 3. Models are numbered A to G, and predictors specified in column 2 ("model"). For the observations of trees either number of species or total foliage cover are used (column 3). Number of trees (trees) and evaluation statistics are reported. The evaluation statistics are on cross-validated (cv) data, and AUC (cvAUC) and % deviance explained (cv%DevExp) are reported. Further details in the text.

ID	Model	number or cover	trees	cvAUC	cv%DevExp
A	C1 C2 climate	num	3,250	0.67	7.04
B		cover	4,100	0.69	8.36
C	PNF1 PNF2 climate	num	4,000	0.68	8.86
D		cover	4,200	0.68	8.82
E	KTS climate	num	3,450	0.72	11.14
F		cover	3,300	0.72	11.40
G	KTSIndex climate	NA	5,200	0.81	24.90

Table 6: Variable importance for each predictor in the models presented in Table 5. Numbers are percentages. Empty cells mean the variable was not in the model.

ID	Model	number or cover	Climate Suit'y	C1	C2	PNF1	PNF2	KTS	KTS Index
A	C1 C2 climate	num	61.3	13.9	24.8				
B		cover	47.3	25.6	27.1				
C	PNF1 PNF2 climate	num	60.3			30.0	9.7		
D		cover	53.7			32.0	14.3		
E	KTS climate	num	56.1					43.9	
F		cover	52.5					47.5	
G	KTSIndex climate	NA	26.1						73.9

Comments on results

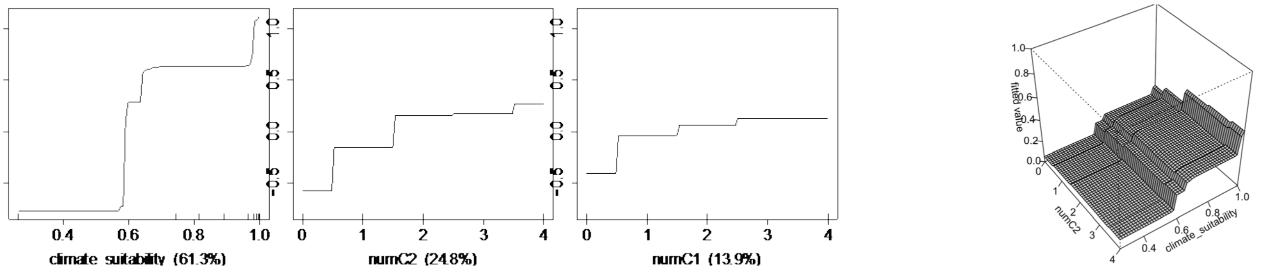
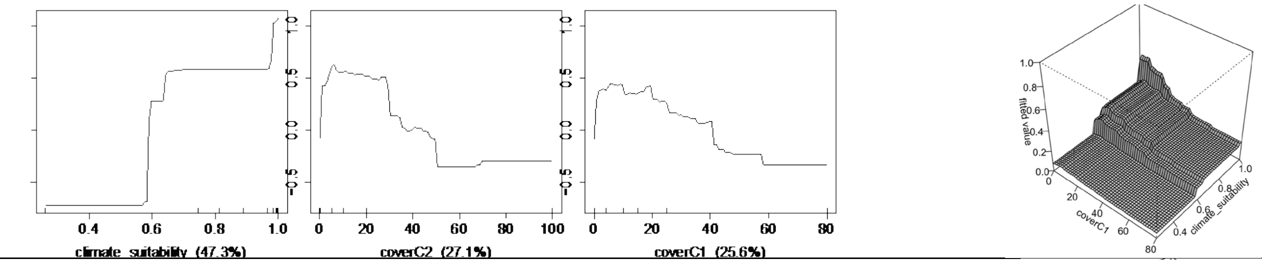
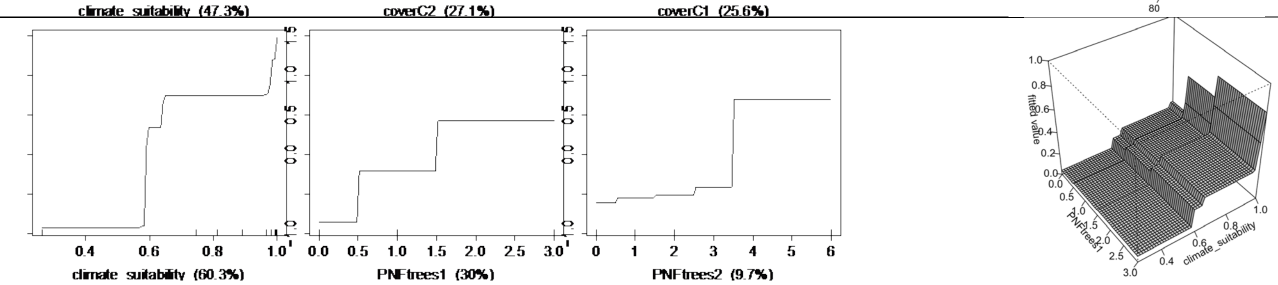
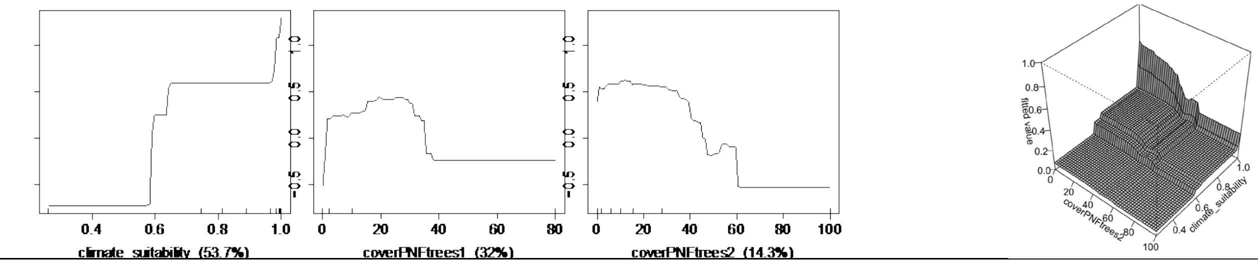
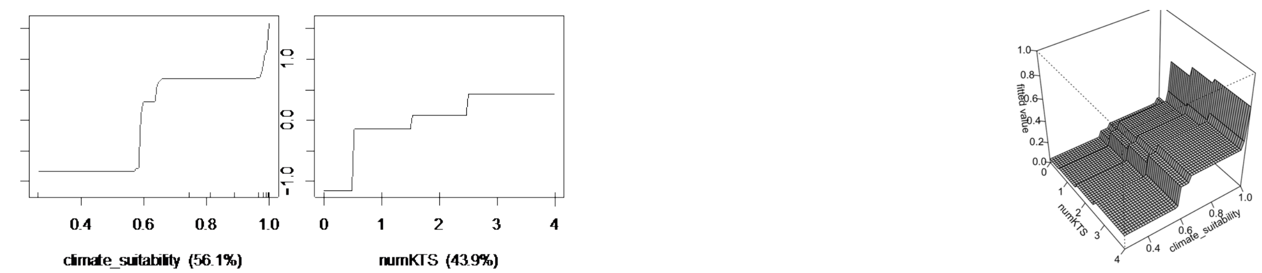
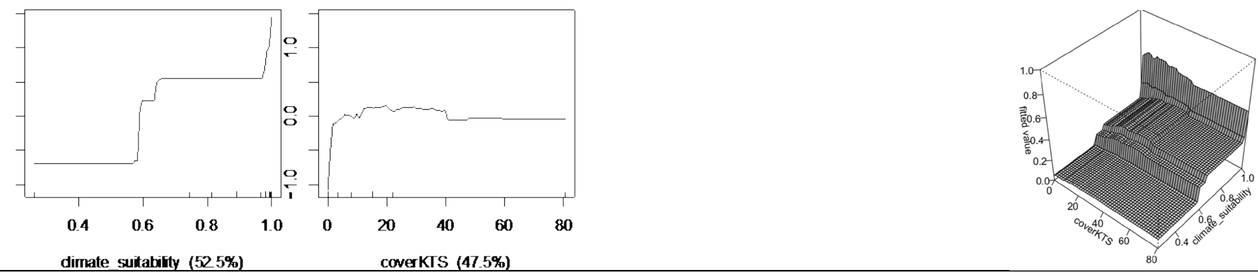
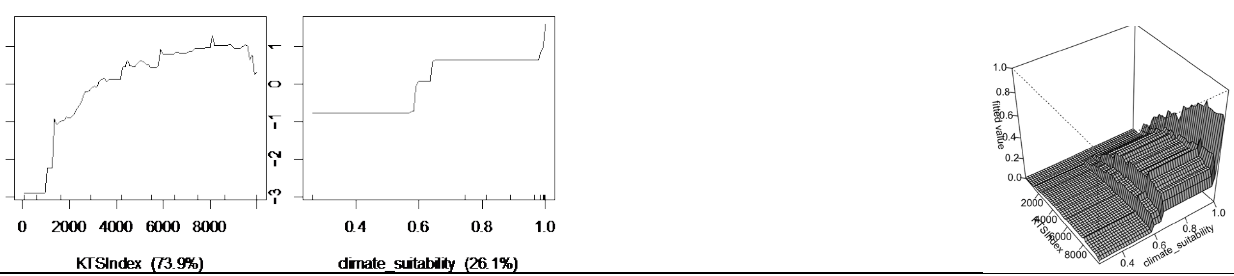
First considering models A to F, which are the models using observations of trees at the evaluation sites (rather than the modelled KTSIndex): model performance is similar across the different tree lists, and also for number vs total cover of browse tree species (Table 5). Discrimination (AUC) is moderate (0.67-0.72), meaning that the model based on trees plus climate tends to predict higher at sites where the Map_input is 1, but not strongly so. Only about 10% of the deviance is explained (7.0-11.1% for number of browse tree species, and 8.4 to 11.4% for cover). Results are similar for number of tree species and cover. The strongest trend is for better performance for the KTS tree list (Models E and F) which is understandable since this list informed the KTSIndex which was used in constructing the Map_input. Even with KTS tree lists, only about 11% of deviance was explained.

Climate suitability is the most important predictor in the models (56 – 61% variable importance for models based on number of tree species, and 47-54% for the cover-based ones, Table 6), and the response to climate suitability (constrained to be monotonic) is positive, with a threshold at ~ 0.6 and again at ~ 0.95 (Table 7, all models). The rest of the variation in the response is explained by the browse tree species. For number of tree species responses are ecologically sensible in that the Map_input tends to be higher at sites with more browse tree species. C2 species explain more of the response than C1 species, with the reverse true for PNF lists. The modelled responses for foliage cover of trees in the C and PNF lists (Models B & D, Table 7) is harder to interpret: the response first rises and tends to plateau (sensible) but then at cover values greater than about 20-30% tends to tail off. A similar response is seen in the models unconstrained to monotonic climatic responses (Table A3.5), so this is not a result of the imposed mono-climatic constraint. Given that most of the cover values are at the lower end of the range (Fig 5, and see rug plots in Table 7), this will not be explored further. The modelled response for cover of KTS species (Model F) is more as expected, not declining at higher cover values. Whilst the specific tree list used in this evaluation does influence the outcome, the difference in the broad meaning of the evaluation does not change: relationships are sensible except for the declining cover response, but leave a lot of variation unexplained.

The interactions (3 D plots in Table 7) either show not much effect of the interaction, or a stronger response to browse tree species at high climate suitabilities (~ 0.95+) than at lower ones (e.g. for models C, E, G) – this makes sense ecologically, since where climate is good, the distribution of browse trees will strongly impact suitability for koalas.

The KTS index was used in the DPE models of koala suitability and was the most important predictor in those koala habitat models (see Data section, point 4) and therefore will have had a strong effect on the Map_input. Hence model G, using KTS index, does not independently evaluate the Map_input, except to the extent that these 3,184 sites were not used in the model building for the PNF Map. Model G shows stronger discrimination and explanation than Models A to F, with AUC of 0.81 and 24.9% deviance explained (Table 5). This implies that there is increased consistency between a major input to the PNF Map and the final product, which is reassuring. KTSIndex is more important than climate in this model, with 74% variable importance. The response to KTSIndex is ecologically sensible: it is largely positive, levelling off at higher values of KTSIndex (Table 7). In Appendix A3 a model (model 12) only including KTSIndex as a predictor is presented, showing that it alone explains 18.7% of the deviance in the Map_input values. One might expect that a key predictor in the koala habitat suitability model (KHSM) underlying the PNF Map would explain more variation in the Map_input values. Possible reasons are that: (i) other variables are important and overshadow the tree index; (ii) the KHSM contained errors; (iii) spatial layers such as the native vegetation extent layers used to clip predictions contained errors (iv) the steps taken to convert the KHSM into the PNF Map introduced noise; or (v) perhaps the KTSIndex contained errors (see next section). These should all be kept in mind when the models are updated.

Table 7 : Modelled responses for models A to G. 3-D plots for some models show the responses for the strongest interaction identified in the model.

A	
B	
C	
D	
E	
F	
G	

Further investigations

Since the KTSIndex explains much more than the observed data it is worth understanding more about the relationship between the two. Analyses (logistic regression) in Appendix A2 and summarised below in Table 8 shows that the relationships between KTSIndex and observed tree species are positive, and (understandably) explain most variation when the actual species underlying the KTSIndex are considered. Note that these analyses only involved the counts of tree species, not cover, because the models underlying the KTSIndex were based on presence-absence tree species records, not measures of cover.

Table 8: Evaluation of the relationship between the KTSIndex and the counts of browse tree species, across the 3184 validation sites. Statistics estimated from a BRT model using 10-fold cross-validation

Tree list	Pearson correlation	% deviance explained
C1 & C2	0.38	12.3
PNF1 & PNF2	0.39	12.7
KTS	0.53	23.0

Focusing on the model of KTSIndex versus observed number of KTS species: one would expect a reasonably strong relationship between the KTSIndex and the observed KTS species if the models making up the KTSIndex were accurate descriptions of tree distribution. The fact that the models of tree distribution (like all models) will be imperfect, and that the analysis is comparing probabilities with counts from PA data means that there will be noise in the relationship. Since the KTSIndex reports the probability that at least one browse tree species is present, the strongest response should be for the step between zero and one trees occurring – this is observed in the models (Appendix A2). Correlations are moderate and 23% of deviance is explained (see final row, Table 8), which is reasonable but not strong explanation. The modelled response is as expected (Appendix A2, and see discussion there). A plot of the raw data (Fig 7) shows that at sites with no KTS species the KTSIndex had a median value of about 1,500 (note the probabilities of the index are multiplied by 10,000), and some had very high KTSIndex values. This will contribute to unexplained variance, and reflects a failure in some of the underlying models to accurately predict the distributions of some of the tree species in these validation data. This suggests that it would be worth revisiting the tree species models that make up the index, to see if they can be improved. As expected, relationships between KTSIndex and other tree lists were less strong (Table 8 and Appendix A2), and this will feed through to the poorer validation results for the C and PNF species lists.

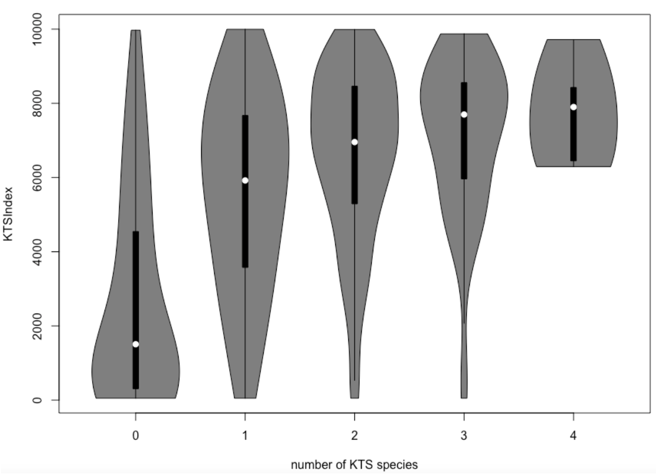


Fig 7: plot of the validation data showing the distribution of KTSIndex values for each distinct number of KTS species. Note that probabilities in the KTSIndex are multiplied by 10,000. The plots ("violin" plots) are kernel density estimates of the data. The white dots show the median value, and black rectangle, the interquartile range.

It is also of interest to know whether results are consistent across PNF regions. Results are presented in Appendix A3, though there are insufficient data for a meaningful evaluation in the River Red Gum region. In other regions the general trends noted above are also seen in these results. The Southern region shows strongest relationships between the Map_input and the various sets of predictors, except for the number of C1 and C2 species which does not predict well at all (presumably because the lists of C1 and C2 species do not align with the KTS species used in the modelling for the Map_input in that region). In the northern region performance of the models with monotonic climate responses are very similar to the averages across sites presented in this main section, except for the KTSIndex which did not predict as strongly. This is likely because in the northern region two koala habitat modelling products were merged. Relationships in the Cypress & Western Hardwoods region were less strong, and generally weaker than seen in the full dataset (except for the KTSIndex which performed reasonably well). These regional results probably largely reflect the regional variation in koala tree species lists across the C, PNF and KTS lists, though it is likely that regional factors also play into the difficulty of modelling koala habitat suitability.

Conclusions

This evaluation of the Map_input to the PNF Map shows that the Map_input has sensible, positive relationships with relevant independent data, but there is a lot of variation in the Map_input that cannot be explained by those data. This raises questions about why the relationship with independent data is not stronger. Since directly relevant data (koala presence-absence or abundance records) were not available for this evaluation, less direct measures of browse trees and climate suitability have been used for evaluation. That does introduce some uncertainty into the interpretation of results, since there is no way to establish how strong a relationship with those measures might be expected. As discussed above, the author did expect to see stronger relationships with browse trees and climate suitability, partly based on the knowledge that input models to the PNF Map showed that browse trees were the most important predictors of koala presence. There is also the issue of how the PNF Map was made. As seen in Box 1, many steps were taken in making the PNF Map including rescaling of predictions, breaking them into classes, choosing amongst overlapping models by taking the highest class, taking a median of values with patches of vegetation, and finally converting an output to binary values. It is not surprising that such steps would introduce noise into the map.

This validation suggests it is worth revisiting the steps in making the PNF Map, including the creation of tree distribution models, the summary of those models in an index, the modelling of koala distribution, and the conversion of those distribution maps to a final binary product that intends to represent high suitability koala habitat. Even when each step is scrutinised and re-run in line with accepted good practice it is possible that the relationship with independent relevant data will still be noisy. Koala habitat is difficult to model well, but re-visiting each step and using new data including recently improved spatial products may lead to meaningful improvements.

References

Briscoe, N.J., Kearney, M.R., Taylor, C.A. and Wintle, B.A. (2016). Unpacking the mechanisms captured by a correlative species distribution model to improve predictions of climate refugia. *Glob Change Biol*, 22: 2425-2439.

Department of Planning, Industry and Environment (DPIE) (2018). *BioNet Vegetation Information System*. Webpage accessed 8 June 2023.

DPIE (2019). *Koala Habitat Information Base Technical Guide*. Report prepared by the DPIE Environment, Energy and Science Division.

DPIE (2020). *Fire Extent and Severity Mapping (FESM) 2019/20*. Spatial dataset available on the NSW Government Sharing and Enabling Environment Data (SEED) portal.

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

Elith, J., S.J. Phillips, T. Hastie, M. Dudík, Y. En Chee, and C.J. Yates. A Statistical Explanation of MaxEnt for Ecologists. *Diversity and Distributions* 17, no. 1 (2011): 43–57.

Hanley, J. A. and McNeil, B. J. (1982). The meaning and use of the area under a Receiver Operating Characteristic (ROC) curve. *Radiology*, 143, 29–36.

Law, B., Caccamo, G., Roe, P., Truskinger, A., Brassil, T., Gonsalves, L., McConville, A. and Stanton, M. (2017). Development and field validation of a regional, management-scale habitat model: A koala *Phascolarctos cinereus* case study. *Ecology and Evolution*, 7(18), 7475–7489.

McNellie, M. J. and Etminaniesfarhani, A. (2023) *Floristic Data Audit and Preparation for Assessing and Transforming Vegetation Cover* (Draft). Unpublished report prepared by the Department of Planning and Environment.

Natural Resources Commission (NRC) (2022). *Advice on finalising draft Private Native Forestry Codes of Practice*. Report prepared for the NSW Government.

R Core Team (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Appendix A1: Correlations between predictors

Table A1.1 Based on PA data

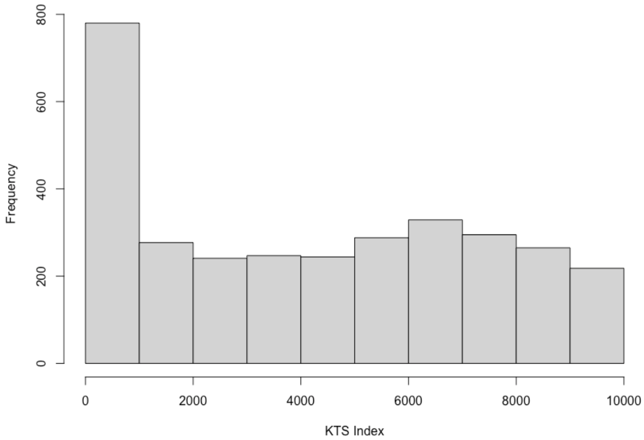
	Climate suitability	numC1	numC2	numPNF1	numPNF2	KTSIndex	numKTS
Climate suitability		0.03	0.35	0.14	0.22	0.08	0.09
numC1			-0.01				
numC2							
numPNF1					0.28		

Table A1.2 Based on cover data

	Climate suitability	coverC1	coverC2	coverPNF1	coverPNF2	KTSIndex	coverKTS
Climate suitability		-0.03	0.24	0.02	0.15	0.08	-0.02
coverC1			-0.09				
coverC2							
coverPNF1					0.07		

Appendix A2. Testing relationship between KTSIndex and number of observed species:

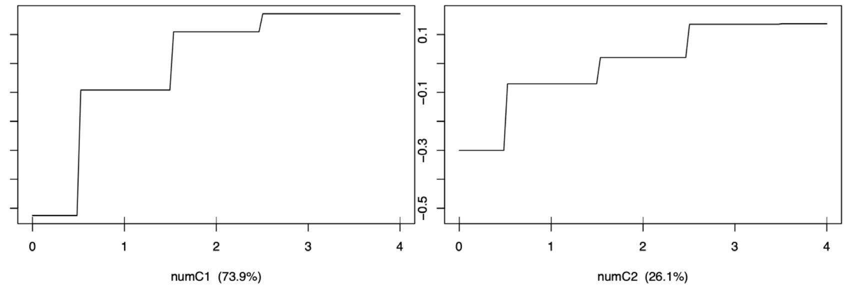
- Treat KTSIndex as a Poisson response because it behaves like a count and has no values below zero. This is the data distribution:



Again, using BRT, with no interactions

(a) KTSIndex ~ numC1 + numC2

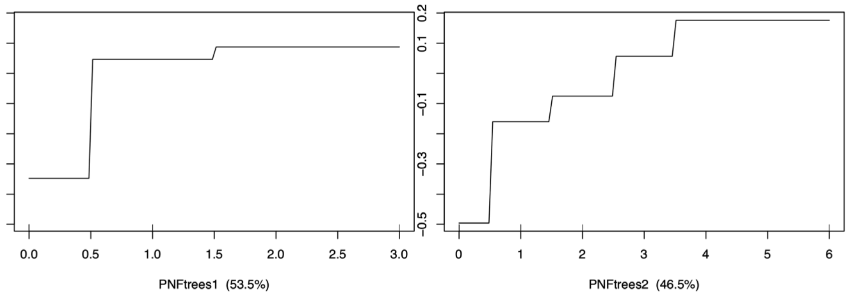
- 1650 trees, statistics from 10-fold cross-validation:
    - correlation between KTSIndex and the observed count of trees = 0.38
    - % deviance explained = 12.3%
    - Relationships are positive:



Here relationships are positive and sensible but there is a lot of variation in the KTSIndex not explained by these 2 predictors.

(b) KTSIndex ~ numPNF1 + numPNF2

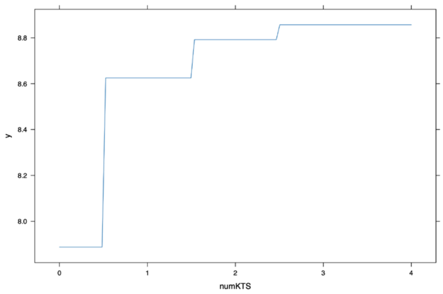
- 1,750 trees, statistics from 10 fold cross-validation:
    - correlation between KTSIndex and the observed count of trees = 0.39
    - % deviance explained = 12.7%
    - Relationships are positive:



Here relationships are positive and sensible but there is a lot of variation in the KTSIndex not explained by these 2 predictors.

(c) KTSIndex ~ numKTS

- This explores to what extent the predictions of KTSIndex at the validation sites are explained by the observed data for the species making up the index – here, the observations are number of KTS species at each site
- 1,100 trees, statistics from 10 fold cross-validation:
    - correlation between KTSIndex and the observed count of trees = 0.53
    - % deviance explained = 23.0%
    - Relationship is positive:



The correlation between KTSIndex and the observed count of trees is moderate (0.53) and the model explains ~ 23% of the deviance. Since the KTSindex reports the probability that at least one browse tree species is present, the strongest response should be for the step between zero and one trees occurring – this is the case, as seen in the figure above. Presumably the failure to explain more of the variation in the data reflects the failure in the underlying tree species models to predict high probabilities for the species actually present at these sites, and/or low probabilities for species not present.

Appendix A3 – All fitted BRT models

This section presents the results and commentary for all models explored – those with a fully flexible response to climate, those with a monotone response to climate (also included in the main document), and those without climate.

Table A3.1: Model details and evaluation for tree predictors based on the presence-absence data (i.e. on counts of tree species in each list). Models are numbered 1 to 12, and predictors in the model specified in column 2 ("model"). Learning rate (lr), tree complexity (tc) and number of trees (trees) are reported. The evaluation statistics are on cross-validated (cv) data, and AUC (cvAUC) and % deviance explained (cvDevExp) are reported.

model_number	model	lr	tc	trees	cvAUC	cvDevExp
1	numC1_C2/climate	0.010	3	4700	0.79	19.25
2	numC1_C2/climate_mono	0.001	3	3250	0.67	7.04
3	numC1_C2	0.001	1	3000	0.63	3.29
4	numPNF1_PNF2/climate	0.010	3	5250	0.81	20.56
5	numPNF1_PNF2/climate_mono	0.001	3	4000	0.68	8.86
6	numPNF1_PNF2	0.001	1	3000	0.58	3.10
7	numKTS/climate	0.010	3	6100	0.82	23.12
8	numKTS/climate_mono	0.001	3	3450	0.72	11.14
9	numKTS	0.001	1	2800	0.65	4.93
10	KTSindex/climate	0.010	3	3500	0.86	30.94
11	KTSindex/climate_mono	0.001	3	5200	0.81	24.90
12	KTSindex	0.001	1	4600	0.76	18.66

Table A3.2: Comparable to Table A3.1, but using tree predictors based on the total foliage cover of the relevant species at each site

model_number	model	lr	tc	trees	cvAUC	cvDevExp
13	coverC1_C2/climate	0.010	3	4950	0.80	19.73
14	coverC1_C2/climate_mono	0.001	3	4100	0.69	8.36
15	coverC1_C2	0.001	1	3550	0.63	3.34
16	coverPNF1_PNF2/climate	0.010	3	4100	0.79	19.19
17	coverPNF1_PNF2/climate_mono	0.001	3	4200	0.68	8.82
18	coverPNF1_PNF2	0.001	1	2450	0.60	2.77
19	coverKTS/climate	0.010	3	5050	0.81	21.11
20	coverKTS/climate_mono	0.001	3	3300	0.72	11.40
21	coverKTS	0.001	1	2350	0.64	4.73

Table A3.3: Variable importance for each predictor in the models presented in Table A3.1. Numbers are percentages.

model_number	model	climate_suitability	numC1	numC2	PNFtrees1	PNFtrees2	KTSIndex	numKTS
1	numC1_C2/climate	87.51	4.72	7.77	NA	NA	NA	NA
2	numC1_C2/climate_mono	61.32	13.91	24.77	NA	NA	NA	NA
3	numC1_C2	NA	10.27	89.73	NA	NA	NA	NA
4	numPNF1_PNF2/climate	85.59	NA	NA	9.90	4.52	NA	NA
5	numPNF1_PNF2/climate_mono	60.26	NA	NA	30.02	9.72	NA	NA
6	numPNF1_PNF2	NA	NA	NA	84.94	15.06	NA	NA
7	numKTS/climate	83.94	NA	NA	NA	NA	NA	16.06
8	numKTS/climate_mono	56.09	NA	NA	NA	NA	NA	43.91
9	numKTS	NA	NA	NA	NA	NA	NA	100.00
10	KTSindex/climate	45.35	NA	NA	NA	NA	54.65	NA
11	KTSindex/climate_mono	26.09	NA	NA	NA	NA	73.91	NA
12	KTSindex	NA	NA	NA	NA	NA	100.00	NA

Table A3.4: Comparable to Table A3.3, but using tree predictors based on the total foliage cover of the relevant species at each site

model_number	model	climate_suitability	coverC1	coverC2	coverPNFtrees1	coverPNFtrees2	KTSIndex	coverKTS
13	coverC1_C2/climate	72.09	12.46	15.45	NA	NA	NA	NA
14	coverC1_C2/climate_mono	47.30	25.57	27.13	NA	NA	NA	NA
15	coverC1_C2	NA	13.45	86.55	NA	NA	NA	NA
16	coverPNF1_PNF2/climate	75.07	NA	NA	14.92	10.01	NA	NA
17	coverPNF1_PNF2/climate_mono	53.67	NA	NA	32.01	14.32	NA	NA
18	coverPNF1_PNF2	NA	NA	NA	90.02	9.98	NA	NA
19	coverKTS/climate	74.66	NA	NA	NA	NA	NA	25.34
20	coverKTS/climate_mono	52.48	NA	NA	NA	NA	NA	47.52
21	coverKTS	NA	NA	NA	NA	NA	NA	100.00

Comments on results

a. *Models using the independent predictors (models 1 to 9 and 13 to 21)*

1. Models with an unconstrained response to climate suitability (models 1, 4, 7, 13, 16, 19):

These are the only models with good discrimination (AUC ~ 0.8) and reasonable but not strong explanation of the data (~ 20% deviation explained). Climate suitability was a very dominant predictor in these models (importance 84 to 88% for models based on number of browse species (i.e. presence-absence data) and 72-75% for models based on total cover of browse species).

Where the response to climate is allowed to take full flexibility the response is highest (but noisy) around climate values of 0.7, tending to decline until climate suitability reaches about 0.85 then rise again (see relevant plots in Table A3.5). Since the number of browse trees does not seem to explain that variation well (see their low variable importance), climate suitability is soaking up much of the variation. The climate suitability predictor summarises suitability of the climate for koalas from an eco-physiological perspective, and is best for identifying unsuitable climates (Dr N. Briscoe, University of Melbourne, pers. comm). Once climate is reasonably suitable, it is likely that other aspects impact koala distribution. Hence whilst these models with unconstrained responses to climate suitability are useful for their ability to show patterns in the data, it could be argued that the fitted function for climate is to some extent soaking up noise, and these models are not the most reliable for evaluating the Map_input. This is why the models with monotone responses to climate are the focus of the main text.

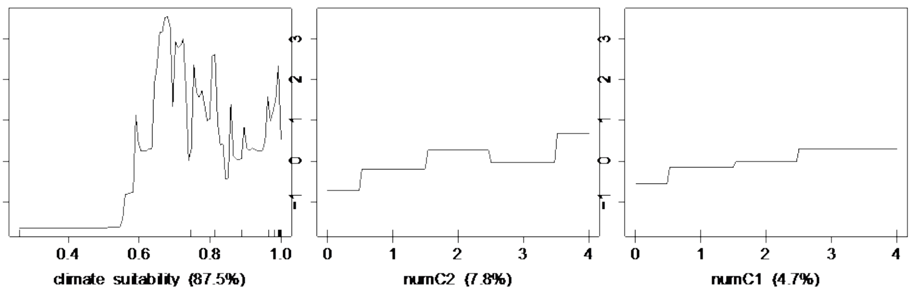
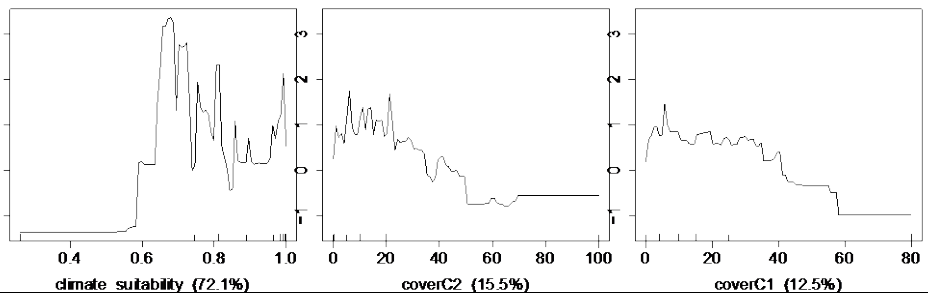
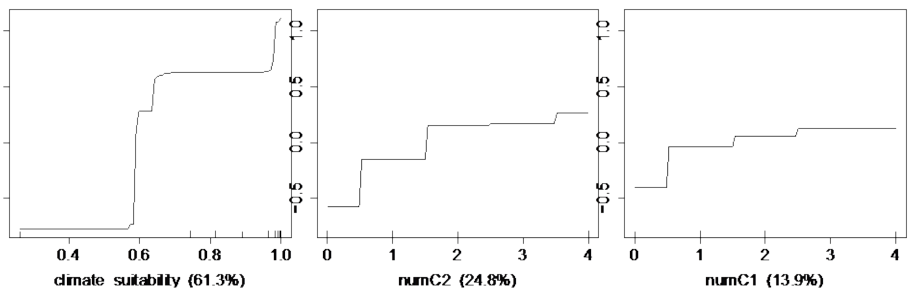
2. Models only with number or cover of browse tree species as predictors (models 3, 6, 9, 15, 18, 21):

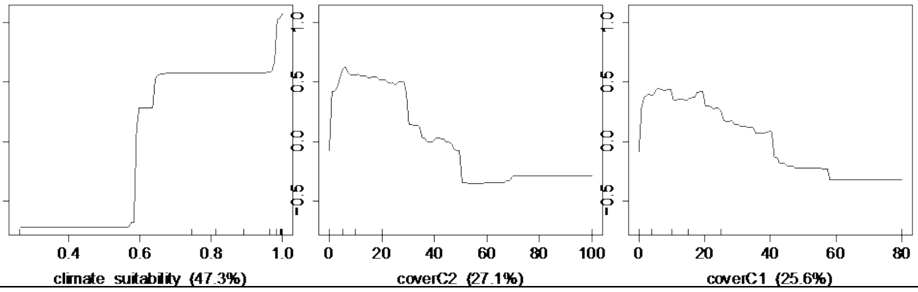
When climate suitability is excluded from the models, discrimination (AUC) drops to 0.58 to 0.65, and only 2.8 to 4.9% of deviance is explained (slightly more for number of species than cover). For C1 and C2 species, number or cover of C2 is much more important in the models than their C1 counterparts (Table A3.3 and A3.4). This is an unexpected result, and cannot be explained by correlations in the data (Tables 1 and 2). In contrast, for PNF lists PNF 1 is much more important than PNF 2. The relationships are sensible and positive for number (Table A3.5), but the results are counterintuitive for the C and PNF lists for cover, with suitability declining at higher cover (Table A3.5). Since there is so much variation left unexplained these models primarily show that number or cover of browse tree species does not alone explain much of the variation in the Map_input values.

b. *Models using KTSIndex instead of observations of the number of browse tree species (models 10-12)*

These models (10-12) showed stronger discrimination and explanation than any of their counterparts using independent data, with AUCs of 0.86, 0.81 and 0.76 for models 10 to 12 respectively, and % deviance explained of 31%, 25% and 19% (Table A3.5). As for the models with independent predictors (models 1-9 and 13 to 21) climate was important when included, but – in contrast to the other models - was never the most important predictor (climate suitability variable importance of 45% when the response was unconstrained, and 26% when constrained, Table 7). The response to KTSIndex was ecologically sensible: it was largely positive, levelling off at higher values of KTS index (Table A3.5). This is discussed further in the main text.

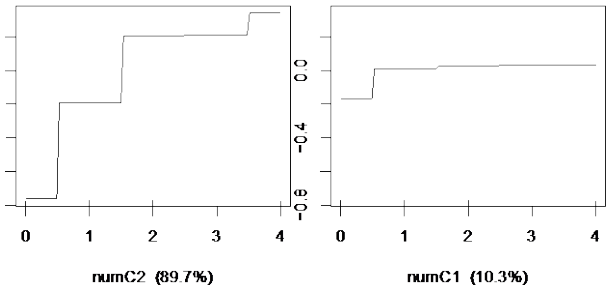
Table A3.5: Modelled responses, with models for number and cover presented together – model numbers as above. See x axes of plots for the tree data used in the model

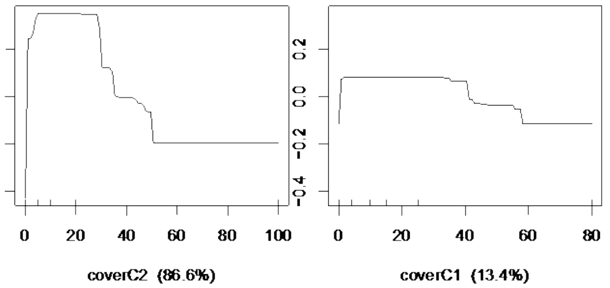
1 13	 
2 14	


climate suitability (47.3%)
coverC2 (27.1%)
coverC1 (25.6%)

3

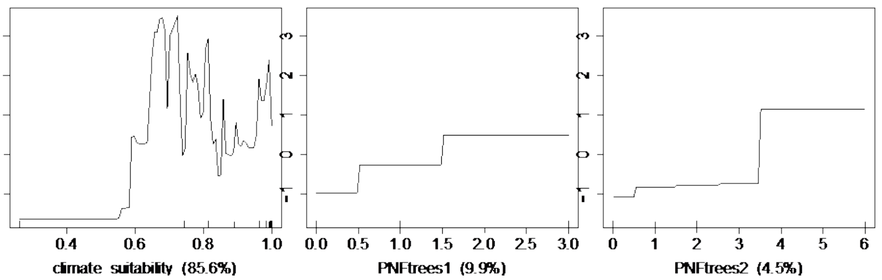
15

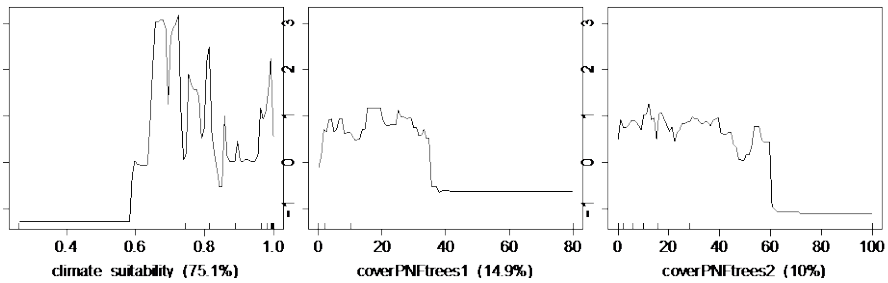

numC2 (89.7%)
numC1 (10.3%)


coverC2 (86.6%)
coverC1 (13.4%)

4

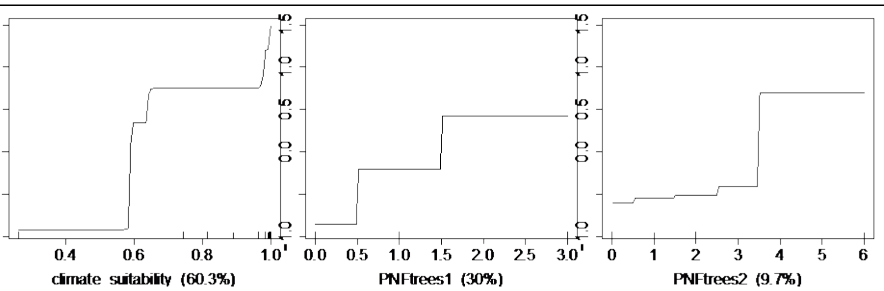
16

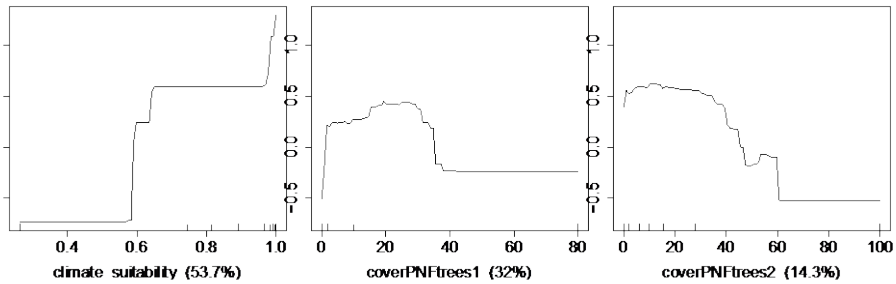
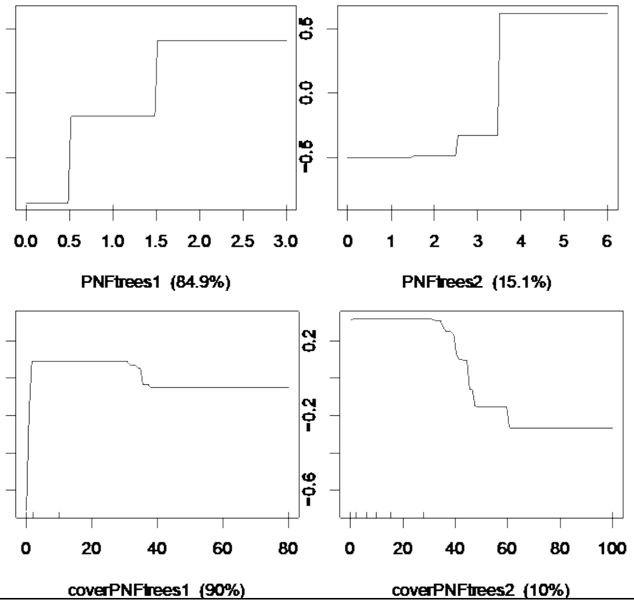

climate suitability (85.6%)
PNFtrees1 (9.9%)
PNFtrees2 (4.5%)

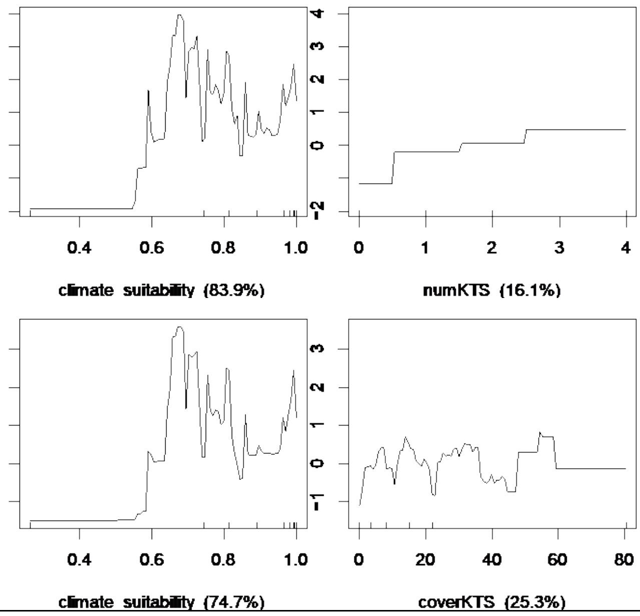
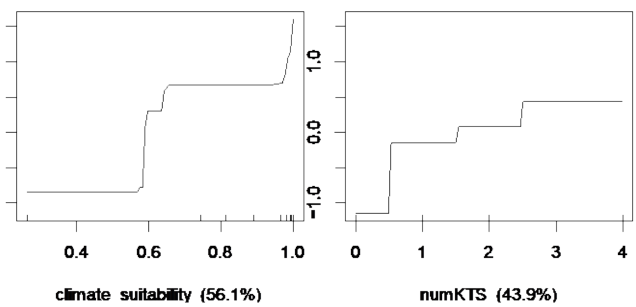

climate suitability (75.1%)
coverPNFtrees1 (14.9%)
coverPNFtrees2 (10%)

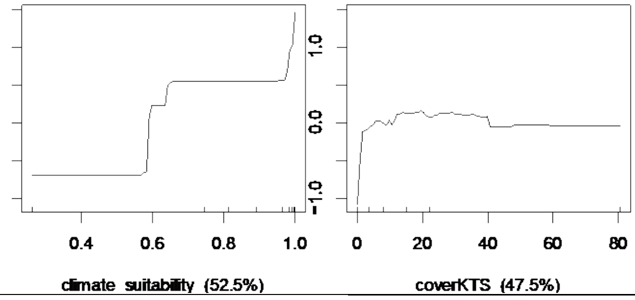
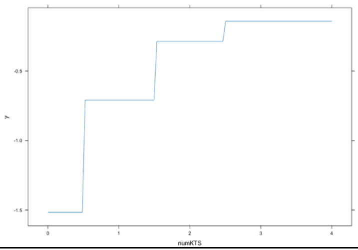
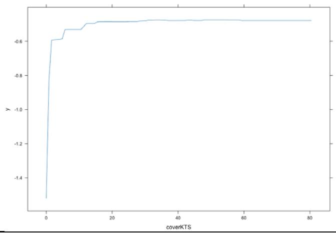
5

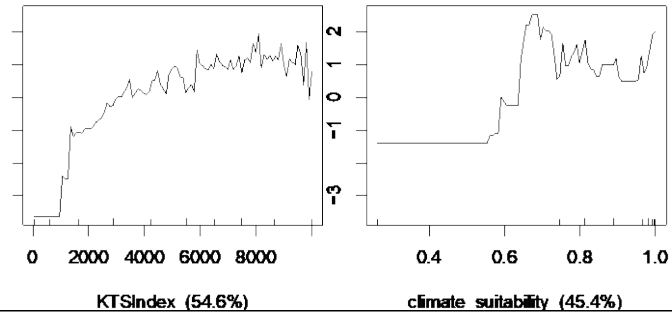
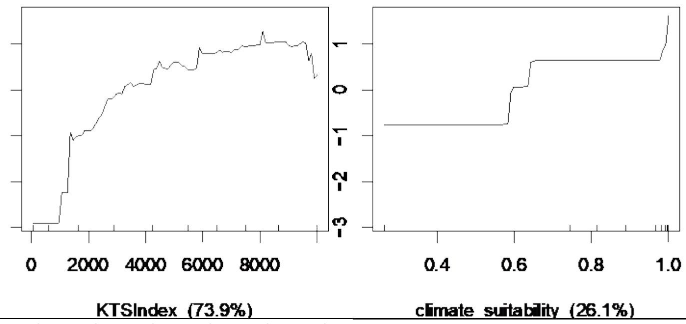
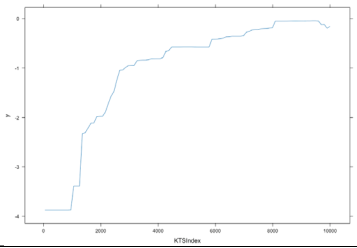
17


climate suitability (60.3%)
PNFtrees1 (30%)
PNFtrees2 (9.7%)

	
6 18	

7 19	
8 20	

	
9 and 21 (side by side)	 

10	
11	
12	

Appendix A4 – Analyses per PNF region.

Since it is of interest to understand how the relationships vary with PNF region, the set of 3,184 validation sites was divided by PNF region. Table A4.1 tabulates the Map index values per PNF region, showing that 3 regions have a reasonable amount of data for this analysis. The exception is the River Red Gum region, which only has 7 sites with a Map_input value (high koala habitat suitability) of 1. Hence the results for that region are unlikely to be reliable and models have not been run. Tables A4.2 and A4.3 present results comparable to those in Appendix A3, but this time on a regional basis.

Table A4.1: Map_input values for sites within each PNF region

Region / Map_input	Northern	Southern	Cypress West Hardwoods	River Red Gum
0	584	580	859	308
1	495	86	265	7

Table A4.2: Model details and evaluation statistics per region (results for monotonic response to climate highlighted)

A. Number of koala tree species.

```
$Northern
   model_number                        model    lr tc trees cvAUC cvDevExp
1             1            numC1_C2/climate 0.010  3  2600  0.75    14.03
2             2       numC1_C2/climate_mono 0.001  3  3100  0.68     7.46
3             3                    numC1_C2 0.001  1  1800  0.57     1.12
4             4        numPNF1_PNF2/climate 0.010  3  2850  0.75    13.51
5             5   numPNF1_PNF2/climate_mono 0.001  3  3650  0.69     8.13
6             6                numPNF1_PNF2 0.001  1  4150  0.61     2.22
7             7              numKTS/climate 0.010  3  3600  0.77    16.03
8             8         numKTS/climate_mono 0.001  3  2800  0.68     6.98
9             9                      numKTS 0.001  1  1150  0.55     0.31
10           10            KTSindex/climate 0.010  3  2400  0.78    17.98
11           11       KTSindex/climate_mono 0.001  3  3800  0.71    10.87
12           12                    KTSindex 0.001  1  2850  0.58     3.46
```

```
$Southern
   model_number                        model    lr tc trees cvAUC cvDevExp
1             1            numC1_C2/climate 0.001  3  5500  0.75    11.73
2             2       numC1_C2/climate_mono 0.001  3  1950  0.64     3.82
3             3                    numC1_C2 0.001  1  2100  0.62     2.12
4             4        numPNF1_PNF2/climate 0.001  3  7150  0.85    24.63
5             5   numPNF1_PNF2/climate_mono 0.001  3  3050  0.75    14.87
6             6                numPNF1_PNF2 0.001  1  4050  0.72    12.29
7             7              numKTS/climate 0.001  3  4250  0.82    23.65
8             8         numKTS/climate_mono 0.001  3  2300  0.76    16.15
9             9                      numKTS 0.001  1  3200  0.73    15.48
10           10            KTSindex/climate 0.001  3  4550  0.95    46.31
11           11       KTSindex/climate_mono 0.001  3  3650  0.94    44.18
12           12                    KTSindex 0.001  1  5000  0.93    44.44
```

```
$`Cypress West Hardwoods`
   model_number                        model    lr tc trees cvAUC cvDevExp
```

1	1	numC1_C2/climate	5e-03	3	3850	0.75	13.11
2	2	numC1_C2/climate_mono	1e-03	3	1650	0.53	2.16
3	3	numC1_C2	1e-05	1	1000	0.54	0.01
4	4	numPNF1_PNF2/climate	5e-03	3	6000	0.78	16.81
5	5	numPNF1_PNF2/climate_mono	1e-03	3	2900	0.57	3.01
6	6	numPNF1_PNF2	NA	NA	NA	NA	NA
7	7	numKTS/climate	5e-03	3	6650	0.78	16.82
8	8	numKTS/climate_mono	1e-03	3	2600	0.60	4.33
9	9	numKTS	1e-05	1	1000	0.59	0.04
10	10	KTSindex/climate	5e-03	3	1650	0.84	27.07
11	11	KTSindex/climate_mono	1e-03	3	3900	0.76	19.61
12	12	KTSindex	1e-03	1	4100	0.73	16.34

B. Cover of browse tree species

$Northern

	model_number	model	lr	tc	trees	cvAUC	cvDevExp
1	13	coverC1_C2/climate	0.010	3	2850	0.76	15.70
2	14	coverC1_C2/climate_mono	0.001	3	5000	0.70	9.14
3	15	coverC1_C2	0.001	1	1850	0.59	1.06
4	16	coverPNF1_PNF2/climate	0.010	3	2000	0.74	13.81
5	17	coverPNF1_PNF2/climate_mono	0.001	3	3350	0.68	7.89
6	18	coverPNF1_PNF2	0.001	1	3750	0.60	2.36
7	19	coverKTS/climate	0.010	3	2100	0.74	12.97
8	20	coverKTS/climate_mono	0.001	3	2950	0.68	7.40
9	21	coverKTS	0.001	1	1000	0.55	0.52

$Southern

	model_number	model	lr	tc	trees	cvAUC	cvDevExp
1	13	coverC1_C2/climate	0.001	3	6600	0.78	15.00
2	14	coverC1_C2/climate_mono	0.001	3	2250	0.65	5.61
3	15	coverC1_C2	0.001	1	5000	0.66	4.12
4	16	coverPNF1_PNF2/climate	0.001	3	5750	0.83	19.56
5	17	coverPNF1_PNF2/climate_mono	0.001	3	2300	0.72	8.02
6	18	coverPNF1_PNF2	0.001	1	5150	0.71	9.09
7	19	coverKTS/climate	0.001	3	4400	0.82	20.85
8	20	coverKTS/climate_mono	0.001	3	2500	0.75	15.69
9	21	coverKTS	0.001	1	4300	0.73	13.61

$`Cypress West Hardwoods`

	model_number	model	lr	tc	trees	cvAUC	cvDevExp
1	13	coverC1_C2/climate	5e-03	3	2800	0.74	11.83
2	14	coverC1_C2/climate_mono	1e-03	3	1800	0.56	2.12
3	15	coverC1_C2	1e-05	1	1000	0.54	0.01
4	16	coverPNF1_PNF2/climate	5e-03	3	4000	0.74	12.91
5	17	coverPNF1_PNF2/climate_mono	1e-03	3	2400	0.54	2.86
6	18	coverPNF1_PNF2	1e-06	1	1000	0.51	0.00
7	19	coverKTS/climate	5e-03	3	4600	0.78	15.82
8	20	coverKTS/climate_mono	1e-03	3	2550	0.64	4.32
9	21	coverKTS	1e-05	1	1000	0.61	0.04

Table A4.3: Variable importance per region (A. number of chosen tree species)

$Northern

	model_number	model	climate_suitability	numC1	numC2	PNFtrees1	PNFtrees2	KTSIndex	numKTS
1	1	numC1_C2/climate	84.32	5.92	9.76	NA	NA	NA	NA
2	2	numC1_C2/climate_mono	70.90	7.74	21.35	NA	NA	NA	NA
3	3	numC1_C2	NA	3.12	96.88	NA	NA	NA	NA
4	4	numPNF1_PNF2/climate	83.19	NA	NA	7.90	8.91	NA	NA
5	5	numPNF1_PNF2/climate_mono	65.17	NA	NA	19.49	15.33	NA	NA
6	6	numPNF1_PNF2	NA	NA	NA	50.00	50.00	NA	NA
7	7	numKTS/climate	89.45	NA	NA	NA	NA	NA	10.55
8	8	numKTS/climate_mono	79.47	NA	NA	NA	NA	NA	20.53
9	9	numKTS	NA	NA	NA	NA	NA	NA	100.00
10	10	KTSindex/climate	58.78	NA	NA	NA	NA	41.22	NA
11	11	KTSindex/climate_mono	46.48	NA	NA	NA	NA	53.52	NA
12	12	KTSindex	NA	NA	NA	NA	NA	100.00	NA

$Southern

	model_number	model	climate_suitability	numC1	numC2	PNFtrees1	PNFtrees2	KTSIndex	numKTS
1	1	numC1_C2/climate	75.97	13.57	10.46	NA	NA	NA	NA
2	2	numC1_C2/climate_mono	43.33	30.06	26.62	NA	NA	NA	NA
3	3	numC1_C2	NA	61.88	38.12	NA	NA	NA	NA
4	4	numPNF1_PNF2/climate	56.64	NA	NA	13.54	29.82	NA	NA
5	5	numPNF1_PNF2/climate_mono	28.93	NA	NA	20.84	50.23	NA	NA
6	6	numPNF1_PNF2	NA	NA	NA	31.03	68.97	NA	NA
7	7	numKTS/climate	53.10	NA	NA	NA	NA	NA	46.90
8	8	numKTS/climate_mono	27.08	NA	NA	NA	NA	NA	72.92
9	9	numKTS	NA	NA	NA	NA	NA	NA	100.00
10	10	KTSindex/climate	17.02	NA	NA	NA	NA	82.98	NA
11	11	KTSindex/climate_mono	10.07	NA	NA	NA	NA	89.93	NA
12	12	KTSindex	NA	NA	NA	NA	NA	100.00	NA

$`Cypress West Hardwoods`

	model_number	model	climate_suitability	numC1	numC2	PNFtrees1	PNFtrees2	KTSIndex	numKTS
1	1	numC1_C2/climate	92.72	5.10	2.18	NA	NA	NA	NA
2	2	numC1_C2/climate_mono	80.17	12.39	7.44	NA	NA	NA	NA
3	3	numC1_C2	NA	3.32	96.68	NA	NA	NA	NA
4	4	numPNF1_PNF2/climate	91.83	NA	NA	5.38	2.79	NA	NA
5	5	numPNF1_PNF2/climate_mono	76.57	NA	NA	11.49	11.94	NA	NA
6	6	numPNF1_PNF2	NA	NA	NA	NA	NA	NA	NA
7	7	numKTS/climate	90.64	NA	NA	NA	NA	NA	9.36
8	8	numKTS/climate_mono	64.15	NA	NA	NA	NA	NA	35.85
9	9	numKTS	NA	NA	NA	NA	NA	NA	100.00
10	10	KTSindex/climate	52.38	NA	NA	NA	NA	47.62	NA
11	11	KTSindex/climate_mono	20.56	NA	NA	NA	NA	79.44	NA
12	12	KTSindex	NA	NA	NA	NA	NA	100.00	NA

Table A4.3: Variable importance per region (B. cover of chosen tree species)

$Northern

	model_number	model	climate_suitability	coverC1	coverC2	coverPNFtrees1	coverPNFtrees2	KTSIndex	coverKTS
1	13	coverC1_C2/climate	62.60	16.80	20.59	NA	NA	NA	NA
2	14	coverC1_C2/climate_mono	47.65	25.29	27.06	NA	NA	NA	NA
3	15	coverC1_C2	NA	31.38	68.62	NA	NA	NA	NA
4	16	coverPNF1_PNF2/climate	66.08	NA	NA	15.96	17.96	NA	NA
5	17	coverPNF1_PNF2/climate_mono	55.92	NA	NA	21.93	22.15	NA	NA
6	18	coverPNF1_PNF2	NA	NA	NA	46.95	53.05	NA	NA
7	19	coverKTS/climate	75.50	NA	NA	NA	NA	NA	24.50
8	20	coverKTS/climate_mono	66.53	NA	NA	NA	NA	NA	33.47
9	21	coverKTS	NA	NA	NA	NA	NA	NA	100.00

$Southern

	model_number	model	climate_suitability	coverC1	coverC2	coverPNFtrees1	coverPNFtrees2	KTSIndex	coverKTS
1	13	coverC1_C2/climate	52.12	22.98	24.91	NA	NA	NA	NA
2	14	coverC1_C2/climate_mono	20.75	40.40	38.85	NA	NA	NA	NA
3	15	coverC1_C2	NA	66.33	33.67	NA	NA	NA	NA
4	16	coverPNF1_PNF2/climate	51.35	NA	NA	28.66	19.99	NA	NA
5	17	coverPNF1_PNF2/climate_mono	31.05	NA	NA	47.15	21.80	NA	NA
6	18	coverPNF1_PNF2	NA	NA	NA	74.99	25.01	NA	NA
7	19	coverKTS/climate	47.74	NA	NA	NA	NA	NA	52.26
8	20	coverKTS/climate_mono	24.42	NA	NA	NA	NA	NA	75.58
9	21	coverKTS	NA	NA	NA	NA	NA	NA	100.00

$`Cypress West Hardwoods`

	model_number	model	climate_suitability	coverC1	coverC2	coverPNFtrees1	coverPNFtrees2	KTSIndex	coverKTS
1	13	coverC1_C2/climate	81.81	11.72	6.47	NA	NA	NA	NA
2	14	coverC1_C2/climate_mono	64.34	17.04	18.62	NA	NA	NA	NA
3	15	coverC1_C2	NA	8.14	91.86	NA	NA	NA	NA
4	16	coverPNF1_PNF2/climate	82.90	NA	NA	7.85	9.25	NA	NA
5	17	coverPNF1_PNF2/climate_mono	58.64	NA	NA	20.87	20.49	NA	NA
6	18	coverPNF1_PNF2	NA	NA	NA	97.29	2.71	NA	NA
7	19	coverKTS/climate	80.94	NA	NA	NA	NA	NA	19.06
8	20	coverKTS/climate_mono	49.86	NA	NA	NA	NA	NA	50.14
9	21	coverKTS	NA	NA	NA	NA	NA	NA	100.00