The influence of non-climate predictors at local and landscape resolutions depends on the autecology of the species

DONNA B. HARRIS,[1]* STEPHEN D. GREGORY,[2] BARRY W. BROOK,[1]
EUAN G. RITCHIE,[3] DAVID B. CROFT,[4] GRAEME COULSON[5] AND
DAMIEN A. FORDHAM[1]

[1]*School of Earth and Environmental Sciences, University of Adelaide, Adelaide, SA 5005, Australia
(Email: donnabharris@googlemail.com);* [2]*Game and Wildlife Conservation Trust, River Laboratory,
Wareham, Dorset, UK;* [3]*School of Life and Environmental Sciences, Deakin University, Burwood,*
[5]*Department of Zoology, The University of Melbourne, Melbourne, Victoria, and* [4]*School of Biological,
Earth & Environmental Sciences, University of New South Wales, Sydney, New South Wales, Australia*

Abstract Species distribution models have come under criticism for being too simplistic for making robust future forecasts, partly because they assume that climate is the main determinant of geographical range at large spatial extents and coarse resolutions, with non-climate predictors being important only at finer scales. We suggest that this paradigm might be obscured by species movement patterns. To explore this we used contrasting kangaroo (family Macropodidae) case studies: two species with relatively small, stable home ranges (*Macropus giganteus* and *M. robustus*) and three species with more extensive, adaptive ranging behaviour (*M. antilopinus*, *M. fuliginosus* and *M. rufus*). We predicted that non-climate predictors will be most influential to model fit and predictive performance at local spatial resolution for the former species and at landscape resolution for the latter species. We compared residuals autocovariate – boosted regression tree (RAC-BRT) model statistics with and without species-specific non-climate predictors (habitat, soil, fire, water and topography), at local- and landscape-level spatial resolutions (5 and 50 km). As predicted, the influence of non-climate predictors on model fit and predictive performance (compared with climate-only models) was greater at 50 compared with 5 km resolution for *M. rufus* and *M. fuliginosus* and the opposite trend was observed for *M. giganteus*. The results for *M. robustus* and *M. antilopinus* were inconclusive. Also notable was the difference in inter-scale importance of climate predictors in the presence of non-climate predictors. In conclusion, differences in autecology, particularly relating to space use, may contribute to the importance of non-climate predictors at a given scale, not model scale per se. Further exploration of this concept across a range of species is encouraged and findings may contribute to more effective conservation and management of species at ecologically meaningful scales.

Key words: home range, kangaroo, land cover, scale, species distribution model.

INTRODUCTION

Climate has been consistently argued to be the main determinant of species distributions at large spatial extents and coarse resolutions, while non-climate predictors (such as topography and habitat) are more important at smaller scales (Pearson & Dawson 2003; Heikkinen *et al*. 2006; Luoto *et al*. 2007). It is therefore common to build coarse-resolution species distribution models (SDM) to characterize species geographic extents and spatial patterns of occurrence using only climate predictors (Araújo & Peterson 2012). This approach, often referred to as bioclimatic envelope modelling, has been criticized, in part,

because the explanatory power (i.e. model fit) and predictive performance of SDMs is often improved by incorporating non-climate predictors, at both small (e.g. Ritchie *et al*. 2008; Tingley & Herman 2009; Titeux *et al*. 2009) and large scales (e.g. Luoto & Heikkinen 2008; Ritchie *et al*. 2008; Bertrand *et al*. 2012; Gillingham *et al*. 2012). Non-climate predictors may affect species distributions by influencing demographic rates in conjunction with (e.g. topography that offers shelter) or independently from (e.g. anthropogenic habitat destruction) climate (e.g. Opdam & Wascher 2004; Luoto & Heikkinen 2008; Sormunen *et al*. 2011). As a result, failure to consider non-climate predictors in SDMs can result in poorly characterized estimates of current distributions and substantial bias in range-change forecasts under future climates (e.g. Bertrand *et al*. 2012; Fordham *et al*. 2012). In consequence, sensitivity analyses of SDM to

both non-climate and climate predictors is increasingly recommended and there is a growing realization that the influence of non-climate predictors may be related more to autecology than to model scale per se (Austin & Van Niel 2011; Bertrand *et al.* 2012). More specifically, we propose that the paradigm of decreasing importance of non-climate predictors with spatial resolution might be obscured by species movement patterns.

Species-environment relationships vary with spatial scale and there will be a particular spatial resolution that best captures the main environmental influences affecting a species vital rates. As the spatial resolution exceeds this optimum, environmental heterogeneity will increase and the signal relating the species to its environment will degrade (Austin & Van Niel 2011). For mobile species, species-environment relationships may be most apparent when the model resolution matches the functional grain of resource patches, which shapes the movements of individuals in the landscape (Titeux *et al.* 2009). In large herbivores, for example, it has been suggested that the primary influence on home range occupation patterns is the spatiotemporal availability of resources (Mueller & Fagan 2008). Stable, sedentary home ranges are promoted where resources are consistently abundant and predictable over time (Mueller & Fagan 2008). In such cases, we hypothesize that non-climate predictors will be most influential at the scale of the home range. Such logic could arguably be applied to any species distribution model but we believe that the concept of the home range is probably too simplistic for species which include irregular and temporary environmental dispersal in their behavioural repertoire (Norbury *et al.* 1994). Where the location of food resources is unpredictable (e.g. influenced by local rainfall events), individuals may make long range excursions to access resources or a nomadic strategy may be optimal (Norbury *et al.* 1994; Mueller & Fagan 2008; McAllister *et al.* 2009) and where resources are predictably found in different places at particular times of the year seasonal migration tends to occur (Mueller & Fagan 2008). For such species, we hypothesize that non-climate predictors will be most influential at coarser/landscape resolutions as the area of environmental influence on a species vital rates will be greater.

In this paper we present a case study on kangaroos (family Macropodidae) to investigate whether the contribution of ecologically relevant non-climate predictors to species distribution models at contrasting spatial resolutions is consistent with the traditional rule of thumb (i.e. climate predictors are more important at coarse resolutions and vice versa for non-climate predictors) or varies according to knowledge of species-specific movement patterns.

The kangaroos are herbivorous marsupials and we selected five species, all with extensive distributions,

for our case study: the eastern grey kangaroo *Macropus giganteus*, common wallaroo *M. robustus*, western grey kangaroo *M. fuliginosus*, red kangaroo *M. rufus* and the antilope wallaroo *M. antilopinus*. The sedentary eastern grey kangaroo, *M. giganteus*, inhabits the mesic eastern seaboard of Australia. It occupies stable home ranges of 0.3–1.6 km² in southern Victoria to areas roughly 10 times larger in semiarid western New South Wales with males travelling at least 3 km per day. The higher habitat productivity across most of its distribution may explain why, on average, home ranges are smaller than those of species inhabiting arid and semi-arid areas (Clancy & Croft 1990; Viggers & Hearn 2005; Coulson 2008). The common wallaroo *M. robustus* has a much broader arid to mesic distribution but a tolerance of high temperatures and a low water requirement, plus an ability to subsist on poor quality vegetation, may buffer the species from variation in resource availability, permitting small (1–3 km²), predominantly stable home ranges (Dawson & Denny 1969; Dawson *et al.* 1975; Clancy & Croft 1990, 2008; Croft 1991). The western grey kangaroo *M. fuliginosus* inhabits arid and semi-arid regions across much of southern Australia. Average home ranges increase with harsher climates varying in size from 0.4 to 0.7 km² in the temperate south-west of Western Australia to 10 times larger in semiarid western New South Wales (Coulson 2008), and long-range movements have been reported in this species, e.g. one male moved 85 km (Priddel *et al.* 1988b). The red kangaroo *M. rufus* prefers the open plains of arid, inland Australia. Home ranges can be large (up to 36 km², Norbury *et al.* 1994) and unpredictable temporal and spatial variation in resource availability encourages long distance, opportunistic forays of around 50 km (Priddel *et al.* 1988a,b; Croft & Clancy 2008). The antilopine wallaroo *M. antilopinus* occupies the monsoonal tropical woodlands of Northern Australia. The species is little-studied, but as observations suggest that it tracks fresh grass growth arising from localized rain (Ritchie 2008) or fire (Hirst 2006) we assume that it makes long-range movements.

Based on this autecological synopsis of the ranging behaviour of these different kangaroo species, we defined two resolutions to test our hypotheses: local resolution was defined as 5 km to approximate the home range sizes of *M. giganteus* and *M. robustus*, and landscape resolution was defined as 50 km to approximate the movement patterns of *M. fuliginosus*, *M. rufus* and *M. antilopinus*. We predicted, *a priori*, that non-climate predictors will be most influential to SDM fit and predictive performance at local resolutions (5 km grid cells) for *M. giganteus* and *M. robustus* and most influential at landscape resolutions (50 km grid cells) for *M. fuliginosus*, *M. rufus* and *M. antilopinus*.

METHODS

We built bioclimatic envelope boosted regression tree (BRT) SDMs for each kangaroo species at local (5 km) and landscape (50 km) resolution for mainland Australia and compared their model fit (% deviance explained) and predictive performance (Kappa) to the same models extended to include non-climate predictors. To account for spatial autocorrelation we included a residuals auto-covariate (RAC) layer as a predictor in our models (Crase *et al.* 2012), that is, RAC-BRT models.

All statistical analyses were done in program R v2.15 (http://www.r-project.org) and specific R packages are mentioned.

Synchronizing data

As both climate and non-climate (e.g. land use) predictors vary over time, it is important to synchronize (approximately) the years of species occurrence record collection with the period of predictor data collection. Failure to do so can degrade the quality of SDMs (Roubicek *et al.* 2010). We therefore restricted the kangaroo occurrence data to records collected over the past 17 years (1995–2011), because this period encompasses the non-climate predictor mapping and corresponds well with the 'current-day' climate averages (1976–2005; focused on 1990) for which we have reliable Australia-wide data.

Species occurrence data and modelling extents

We compiled kangaroo occurrences on mainland Australia (number of occurrences in parenthesis) for *M. giganteus* (7312), *M. robustus* (10 330), *M. fuliginosus* (2058), *M. rufus* (4550) and *M. antilopinus* (907) from a variety of sources (see Appendix S1 in Supporting Information). Data were screened by the following criteria: observations had to be of live individuals or of recently dead animal remains (e.g. road kill with soft tissue) that would have been alive within the period 1995–2011; locations of observations had to be accurate to within 2.5 km, and species identification had to be certain (meaning some aerial occurrence records were excluded). Due to a low number of recent occurrence records for the Western Australian desert regions, 10 location records (4 *M. robustus*, 6 *M. rufus*) from 1961–1980 were included (the only additional records we could find that were accurate to within 2.5 km and for which species identification was certain). Since the habitat of Western Australian deserts remains relatively unaffected by anthropogenic impact (B. Prince, pers. comm., 2011), the benefits of including older desert kangaroo occurrences outweighed any potential problems with the asynchrony of these records with our current-day predictor data.

To delineate species-specific extents for model training and prediction, three of us with expertise in macropod ecology (G. Coulson, D. Croft and E. Ritchie) blindly selected (i.e. without the aid of the occurrence data) Interim Biogeographic Regionalisation of Australia (IBRA) regions in which the species could feasibly be found. In the next step,

these experts judged the plausibility of any occurrence data outside of nominated extents and we appended any additional IBRA regions accordingly. Each species-specific subset of IBRA polygons (vector data) was then converted to a 5 km raster in ArcGIS 9 using the maximum combined area cell assignment type. The 5 km raster was then cropped and resampled to a standardized extent before aggregating by mode to 50 km resolution using package raster (Hijmans 2013). The number of cells in the IBRA extent rasters was (5 km, 50 km): *M. giganteus* 114 327, 1333; *M. robustus* 254 115, 2760; *M. fuliginosus* 131 754, 1490; *M. rufus* 238 183, 2589 and *M. antilopinus* 46 318, 586. For each species, the final set of point locations was converted to a 5 km resolution presence-only raster (i.e. unoccupied cells were classified as 'no data') of standardized extent. To standardize interspecific comparisons, the ratio of presence : extent cells was calculated for each species and presences randomly subsampled so that all species had ratios equal to that of the species with the lowest ratio (1:164 for *M. antilopinus*). The final point locations used are plotted over the IBRA regions in Figure 1. The 5 km rasters were then aggregated to 50 km resolution by mode. The number of occurrence cells in the 5 and 50 km rasters respectively was: *M. giganteus* (719, 286), *M. robustus* (1544, 431), *M. fuliginosus* (805, 230), *M. rufus* (1447, 476) and *M. antilopinus* (282, 93).

Climate predictors

Climate spatial layers (0.05 × 0.05° latitude/longitude) describing long-term monthly averages (1976–2005) of precipitation (total monthly, mm) and temperature (mean, maximum and minimum monthly, °C), centred on 1990, were sourced from the Australian Bureau of Meteorology (http://www.bom.gov.au). These data were processed to produce 13 climate predictors considered important in defining kangaroo species ranges and used in previous models of kangaroo distributions (Caughley *et al.* 1987; Walker 1990; Skidmore *et al.* 1996; Skidmore 1998; Ritchie *et al.* 2008): annual mean temperature (*tann*), maximum temperature of the warmest period (*maxtwarmp*), minimum temperature of the coldest period (*mintcoldp*), mean temperature of the wettest quarter (*twetq*), mean temperature of the driest quarter (*tdryq*), annual temperature range (*trange*), annual mean precipitation (*pann*), annual precipitation range (*prange*), coefficient of variation in precipitation (*cvprecip*), precipitation of the driest period (*pdryp*), precipitation of the driest quarter (*pdryq*), precipitation of the wettest period (*pwetp*), precipitation of the wettest quarter (*pwetq*). Period refers to a month and quarter is three months.

Non-climate predictors

Species-specific non-climate predictors were identified from a literature review of factors affecting distribution or abundance of the five focal species (see Appendix S2 in Supporting Information): *habitat* (grassland or non-grassland), *landuse* (grazing land or non-grazing land), *soil* (high productivity or low productivity), *latefire* and *earlyfire* (number of
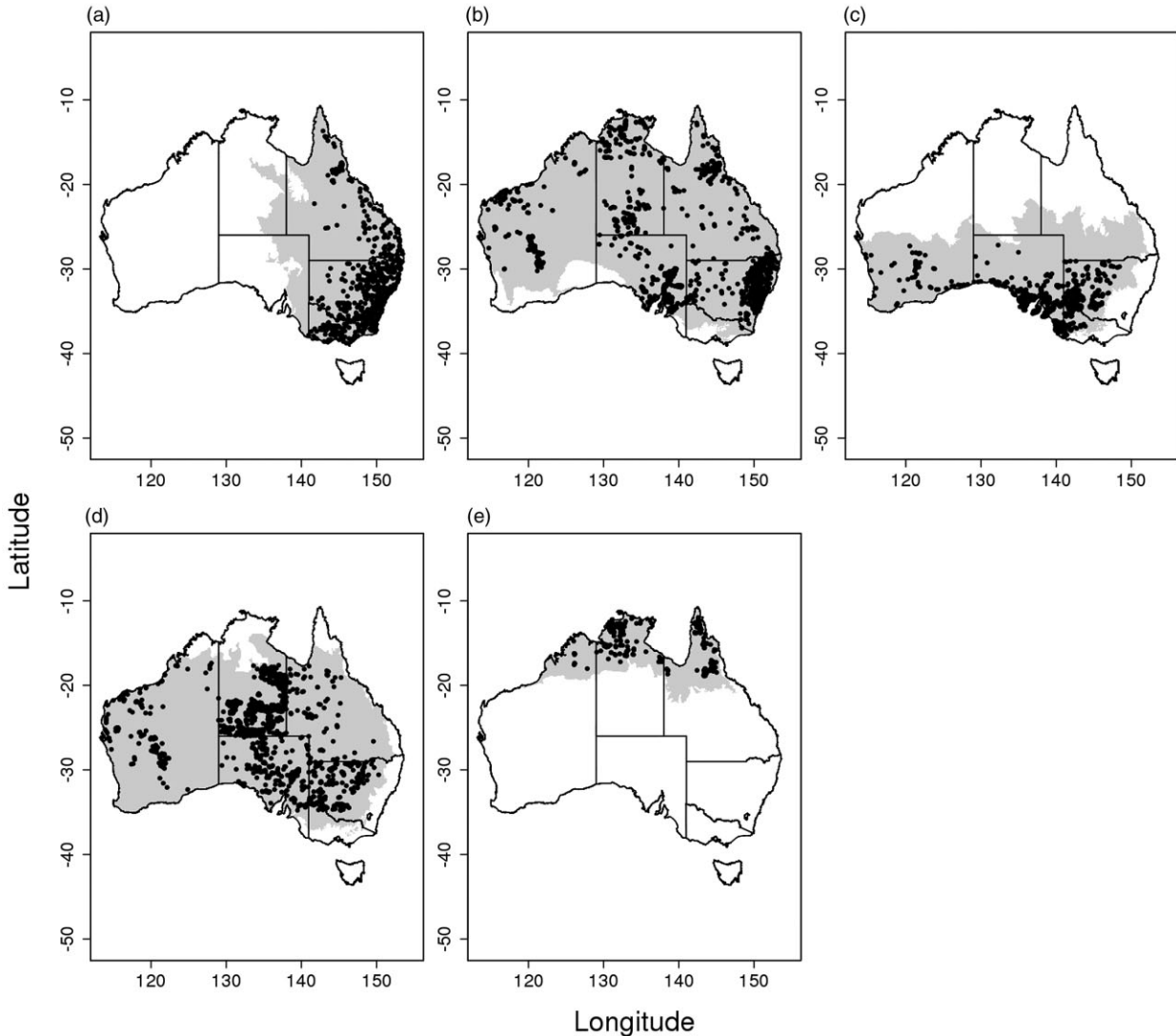
Fig. 1. Point locations for (a) *Macropus giganteus*, (b) *M. robustus*, (c) *M. fuliginosus*, (d) *M. rufus* and (e) *M. antilopinus* plotted on species-specific model extents derived from Interim Biogeographic Regionalisation of Australia (IBRA) regions.

years land burned between 1997–2010 in the late dry season or early dry season, respectively), *pwater* (distance to closest permanent water, in hundreds of km), *rock* (distance to closest rock shelter: caves, cliffs and outcrops; in hundreds of km) and *slope* (in degrees from horizontal). Data sources and processing methods are summarized in Appendix S3 of Supporting Information.

To minimize the number of climate and non-climate predictors, and so retain a relatively parsimonious model set, those parameters that were highly correlated (Spearman's Rank $r \geq \pm 0.7$), across a given species-specific extent were excluded from further analysis in favour of representative predictors that were inter-correlated with the most others or, where all else was equal, were more general or easier to interpret (e.g. *pdryq* retained instead of *pdryp*). Spearman's Rank r exceeded 0.7 only between climate predictors so it was not necessary to remove any non-climate predictors.

Data analysis

To test our predictions, we compared the model fit and predictive performance of SDMs with and without non-climate predictors (i.e. an inter-model comparison). We did this at the local and landscape resolutions separately and contrasted the relative improvement in explanatory and predictive power between these resolutions graphically. We repeated this for each species separately and contrasted the findings based on their autecology.

The SDMs were built using boosted regression trees (BRTs). A BRT model can be understood as an additive regression model in which individual terms are simple trees (models that relate a response to their predictors by recursive binary splits), fitted in a forward, stagewise fashion (Elith *et al*. 2008). This method was preferred for several reasons: (i) to avoid the need to transform data or eliminate outliers,

(ii) because BRTs can fit nonlinear relationships, (iii) because the hierarchical structure of a tree means that interactions between predictors are automatically modelled, (iv) because they readily accommodate continuous and binary variables, and (v) because regression trees have performed relatively well in previous models of kangaroo distributions (Walker 1990; Skidmore *et al.* 1996; Elith *et al.* 2008). Although BRTs can fit complex functions, the implementation in gbm prevents overfitting by applying a constant (the learning rate [*lr*]) to the cost function to reduce the speed that it moves down the error surface. In addition, overfitting can be further avoided by restricting the tree complexity (*tc*; synonymous with the level of interactions), which we did (see below). BRTs require both presence and absence data, but because Australia has not been systematically and extensively surveyed for kangaroos in recent years, reliable absence data are unavailable. To address this shortcoming we generated random 'pseudoabsences' from non-occurrence cells within the species-specific IBRA extent or background environment. Using IBRA regions to constrain background points, thus limiting the pseudoabsence sampling pool, reduces the likelihood of inflated estimates of model performance and over-prediction of distribution area (VanDerWal *et al.* 2009). As model performance measures are affected by prevalence (occurrences/[occurrences + pseudoabsences]) we used a prevalence of 0.5 across all models (Elith *et al.* 2008; Tingley & Herman 2009).

Climate-only and climate + non-climate BRTs were built for each species at 5 and 50 km resolutions using package gbm (Ridgeway 2013) and the helper functions in package dismo (Hijmans *et al.* 2013). For each species/resolution combination, the climate-only model was built using the same presence/absence dataset as for the climate + non-climate model. All combinations of settings for tree complexity (*tc*, the number of nodes in a tree) 1, 2, 3, 4, 5 and learning rate (*lr*, which determines the contribution of each tree to the growing model) 0.01, 0.005, 0.001, 0.0005 were tested. We selected the largest *lr* and smallest *tc* at which number of trees (*nt*) > 1000 and 10-fold cross-validated holdout model deviance was minimized. As stochasticity improves predictive performance (Elith *et al.* 2008), bag fraction (*bf*) – which specifies the proportion of data to be selected at each step – was retained at 0.75 (default in function gbm.step), which gives good results for presence-absence responses (Elith *et al.* 2008). The model's structural fit (% deviance explained), model predictive performance (Kappa) and relative contribution of predictors (% importance) were calculated. To estimate Kappa, continuous model output must first be converted to binary predictions using a threshold probability of occurrence value. The threshold was that at which the sum of sensitivity and specificity was maximized (Liu *et al.* 2005). The percentage importance measure relates to the number of times the variable is selected during the BRT fitting procedure and is scaled across model predictors so that the sum is 100, with higher numbers indicating stronger influence on the response variable. Partial dependence function plots show the effect of a predictor on the response (probability of occurrence) after accounting for the average effects of all other variables in the model (Elith *et al.* 2008). Fitted functions can be interpreted as follows: $y = 0$ there is no partial dependence on the predictor for the prediction into class 1

(presence) and $y > 0$ or < 0 indicate a positive or negative influence, respectively (Friedman & Meulman 2003). Fitted functions were plotted as local polynomial regression (loess) smooths with span set to 0.4, as given as an option in gbm.plot.

Boosted regression tree is more robust to autocorrelation in the response variable than more general linear models but the residuals should still be examined for any sign of autocorrelation. To inspect the structure of spatial autocorrelation (SAC) in the data we plotted correlograms of model residuals using package ncf (Bjornstad 2013). These revealed that SAC was present in the residuals of most models. To reduce SAC, we adopted an approach recently proposed by Crase *et al.* (2012), which when tested extensively with simulated data, was shown to reduce SAC whilst still preserving the direction and shape of the relationship between the response and predictor variables. We calculated a residual autocovariate (RAC) layer for each model to capture SAC by averaging the residuals in a first order neighbourhood (3×3 cells). This RAC layer was then included as a predictor in the corresponding BRT model. For a more detailed description of this procedure see Crase *et al.* (2012). The model settings from the original BRT models were applied unless $nt < 1000$, in which case *lr* was decreased until $nt > 1000$. The RAC-BRT residuals were plotted in correlograms and compared with the BRT correlograms to assess the reduction in residual SAC. Where first order neighbourhood averaging did not reduce the SAC below that in the original BRT model we compared second (5×5 cells), third (7×7 cells) and fourth (9×9 cells) order neighbourhoods and selected the neighbourhood that minimized sum of squares to the zero line in the corresponding correlogram.

The RAC layer will absorb an unknown amount of variance that might otherwise be explained by the predictor variables, complicating the isolation of their influence on the model fit and predictive performance at contrasting spatial resolutions. To address this important issue we ran a 10-fold cross-validation in which RAC was included in the training dataset but removed from the test dataset by setting the RAC predictor to zero. This approach allowed us to compare the influence of non-climate predictor addition on predictive performance (predicted deviance and Kappa) at contrasting spatial resolutions while controlling for the potentially confounding effects of the RAC layer. Predicted deviance was averaged across the 10 held-out test datasets. The threshold probability of occurrence value (at which the sum of sensitivity and specificity was maximized) and corresponding Kappa value were calculated for each test dataset and then Kappa values ($n = 10$) averaged. Where the cross-validation analyses supported the inter-model (climate *vs.* climate + non-climate) trends between local and landscape resolutions from the complete dataset analyses (e.g. non-climate predictors increased model predictive performance more at 50 km compared with 5 km using the complete and cross validation datasets) we assumed our inter-model comparisons to be robust to the influence of RAC, that is, we were able to conclude that the trends were attributable to the influence of non-climate predictors. In such cases we examined the importance scores of pertinent predictors from the models built using the complete datasets. Partial dependence plots were also interpreted to elucidate the ecological relevance of the models.

RESULTS

Importance of non-climate predictors at contrasting scales

The trends in non-climate predictor contribution to model fit and predictive performance between local and landscape resolutions when using the complete datasets were supported by the cross-validation results for three species *M. rufus*, *M. fuliginosus* and *M. giganteus*. The addition of non-climate predictors to the bioclimatic envelope SDMs had greater influence on model fit and predictive performance (% deviance explained and Kappa) at a 50 km grid cell resolution compared with a 5 km resolution for *M. rufus* and *M. fuliginosus*. The opposite effect was observed for *M. giganteus* with non-climate predictors having more of an influence at 5 km. Results for these species are discussed further below. Results for *M. robustus* and *M. antilopinus* were inconclusive as the inter-model trend was not maintained at cross-

validation suggesting that the variance explained by RAC confounds the inter-resolution comparison of non-climate predictor contribution (Fig. 2: calculated from values in Table 1). It is also interesting to note that the addition of non-climate predictors affected the relative importance of climate predictors at contrasting scales (Table 2).

M. giganteus

Model fit and predictive performance were improved by the addition of non-climate predictors at 5 km relative to 50 km resolution and this trend was consistent across both the RAC-BRT built on the complete and the cross-validation datasets (Fig. 2). This suggests that the differences in model fit and predictive performance (Fig. 2a,b) can be linked to the influence of the non-climate predictors. The partial responses for *M. giganteus* for the five most influential variables in the 5 km climate + non-climate model indicate a
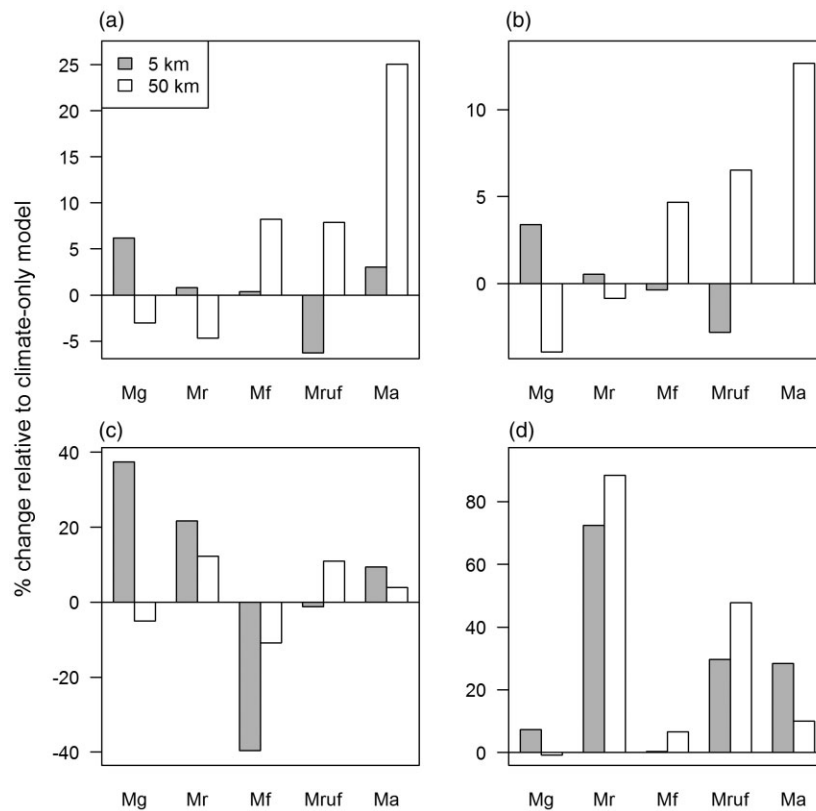

Fig. 2. Percent difference in explanatory and predictive metrics for climate + non-climate relative to climate-only residuals autocovariate – boosted regression tree (RAC-BRT) models, that is, {[(metric value for climate + non-climate RAC-BRT – metric value for climate-only RAC-BRT)/metric value for climate-only RAC-BRT] × 100}, for five kangaroo (*Macropus*) species at 5 km resolution (grey bars) and 50 km resolution (white bars). Model fit is measured as % deviance explained (a) and predictive performance as Kappa (b). Corresponding bar charts for predicted deviance (c) and Kappa (d) from the 10-fold cross validation (with RAC in the test folds set to zero) are also presented. Note: Y axis scales differ between plots. Species abbreviations are: Mg, *M. giganteus*; Mr, *M. robustus*; Mf, *M. fuliginosus;* Mruf, *M. rufus* and Ma, *M. antilopinus*.

Table 1. Residuals autocovariate – boosted regression tree (RAC-BRT) model fit (% deviance explained) and predictive performance (Kappa) with optimal parameters (*lr* – learning rate, *tc* – tree complexity, *nt* – number of trees) followed by 10-fold cross-validated predictive performance (predictive deviance and Kappa) for each combination of species, predictor set and spatial resolution

Species	Res	Model	%DE	Kappa	*lr*	*tc*	*nt*	CV_PredDev	CV_kappa
M. giganteus	5 km	clim	76.9	0.861	0.005	3	1650	1.686	0.579
		clim + non-clim	81.7	0.890	0.005	4	1650	1.055	0.620
M. robustus		clim	96.9	0.986	0.005	4	1700	1.273	0.338
		clim + non-clim	97.7	0.992	0.01	4	1200	0.997	0.583
M. fuliginosus		clim	97.1	0.990	0.005	5	1100	1.119	0.566
		clim + non-clim	97.4	0.986	0.005	5	1150	1.563	0.568
M. rufus		clim	87.6	0.933	0.01	4	1200	1.293	0.310
		clim + non-clim	82.1	0.907	0.005	5	2000	1.308	0.401
M. antilopinus		clim	93.8	0.986	0.001	3	4050	1.314	0.348
		clim + non-clim	96.6	0.986	0.005	3	1150	1.191	0.446
M. giganteus	50 km	clim	84.6	0.920	0.005	1	2900	0.655	0.773
		clim + non-clim	82.0	0.883	0.005	1	3200	0.687	0.767
M. robustus		clim	66.8	0.805	0.001	5	4300	1.305	0.242
		clim + non-clim	63.7	0.798	0.001	3	4450	1.144	0.455
M. fuliginosus		clim	67.9	0.822	0.001	1	6700	1.102	0.513
		clim + non-clim	73.5	0.860	0.001	4	3450	1.222	0.547
M. rufus		clim	48.6	0.687	0.0025	4	1600	1.271	0.334
		clim + non-clim	52.4	0.732	0.0025	5	1950	1.133	0.493
M. antilopinus		clim	50.9	0.772	0.0005	3	6250	1.282	0.453
		clim + non-clim	63.7	0.870	0.001	4	3250	1.232	0.499

CV_PredDev is the mean deviance between the predicted and the observed response variable calculated as $[y\_i \times \log(u\_i)] + [(1 - y\_i) \times \log(1 - u\_i)]$ where y_i are the observed responses and u_i are the predicted responses.

species occurring in relatively cool, wet, unburnt environments (Fig. 3a). Both late and early dry season fire frequencies as well as temperature of the driest quarter (reflected also in the climate-only model) were more pertinent at 5 km relative to 50 km resolution (Table 2).

M. fuliginosus

The trend in non-climate predictor influence on model fit and predictive performance at the two spatial resolutions was consistent across both the RAC-BRT built on the complete and the cross-validation datasets for *M. fuliginosus* (Fig. 2). The partial responses for *M. fuliginosus* for the five most influential variables in the 50 km climate + non-climate model (Fig. 3b) indicate a species occurring in relatively arid environments with moderate average temperature and warm dry seasons, preferably close to permanent water. It is associated with areas with a low intra-annual precipitation range. All of the top five most important predictors in the 50 km climate + non-climate model were more important at 50 km relative to 5 km resolution but distance to permanent water and average annual precipitation (greater magnitude than climate-only comparison) increased the most in importance (Table 2).

M. rufus

Model fit and predictive performance were improved by the addition of non-climate predictors at 50 km relative to 5 km resolution and this trend was consistent across both the RAC-BRT built on the complete and the cross-validation datasets (Fig. 2). The partial responses for *M. rufus* for the five most influential variables in the 50 km climate + non-climate model (Fig. 3c) indicate a species occurring in hot, arid areas with a large intra-annual temperature range but close to permanent water and with an absence of early dry season fire. All of the top five most important predictors in the 50 km climate + non-climate model were more important at 50 km relative to 5 km resolution but early dry season fire frequency, average annual temperature (greater magnitude than climate-only comparison) and distance to permanent water were the most obvious (Table 2).

M. robustus and *M. antilopinus*

The trend in non-climate predictor influence on model fit and predictive performance at the two spatial resolutions was not consistent across both the RAC-BRT built on the complete and the cross-validation datasets for *M. robustus* and *M. antilopinus* (Fig. 2).

Table 2. Predictor importance scores at contrasting resolutions (5 km and 50 km) for climate + non-climate residuals autocovariate – boosted regression tree (RAC-BRT) models and climate-only RAC-BRTs for the five most important 5 km climate + non-climate predictors for *M. giganteus* and 50 km climate + non-climate predictors for *M. fuliginosus* and *M. rufus*

	Climate + non-climate models			Climate-only models		
	5 km	50 km	% difference	5 km	50 km	% difference
M. giganteus						
tann	5.97	14.52	143.16	7.47	17.16	129.78
pann	4.93	8.61	74.65	3.91	9.50	142.86
tdryq	1.95	1.47	−24.73	3.01	1.65	−45.23
latefire	0.26	0.00	−98.95			
earlyfire	0.23	0.05	−79.94			
RAC	86.13	74.48		85.61	71.69	
Other	0.54	0.88		0	0	
M. fuliginosus						
tann	0.60	5.53	827.27	0.50	9.54	1792.34
pann	0.17	2.98	1646.37	0.27	1.01	273.91
pwater	0.09	1.69	1759.62			
tdryq	0.26	1.40	443.98	0.23	0.63	175.12
prange	0.14	1.13	679.21	0.24	0.34	44.44
RAC	98.57	86.38		98.62	88.03	
Other	0.17	0.90		0.14	0.44	
M. rufus						
pann	2.07	9.34	351.68	1.84	9.66	424.33
pwater	1.18	6.35	438.89			
tann	0.94	5.82	516.98	1.04	4.63	346.37
trange	1.84	4.98	170.96	1.53	2.31	51.66
earlyfire	0.52	4.14	696.92			
RAC	91.45	63.80		93.93	78.28	
Other	2.00	5.57		1.67	5.11	

The RAC predictor importance and sum of all other predictors' importance scores ('other') are also presented. Predictors are earlyfire, early fire frequency; latefire, late fire frequency; pann, annual mean precipitation; prange, annual precipitation range; pwater, distance to closest permanent water; tann, annual mean temperature; tdryq, mean temperature of the driest quarter; trange, annual temperature range.

This suggests that the variation explained by the RAC layer confounds the comparisons between local and landscape resolutions. We do not, therefore, present any further results for these species.

Spatial autocorrelation

The correlograms of residuals from the BRT models and the RAC-BRT models show that the addition of the RAC predictor substantially reduced first-order spatial autocorrelation in the model residuals (Fig. 4).

DISCUSSION

Recent studies have highlighted the concept of non-climate predictor influence at a scale most relevant to a species ecology (Austin & Van Niel 2011; Bertrand *et al.* 2012). Other studies have compared non-climate predictor importance in SDMs at contrasting spatial scales (Luoto *et al.* 2007 – birds; Gillingham *et al.* 2012 – beetles) and at least one study has directly contradicted Pearson and Dawson's (2003) paradigm of increasing non-climate predictor pertinence at smaller spatial scales (Gillingham *et al.* 2012). Our paper is the first attempt, however, to consider whether autecological knowledge (of movement patterns) might be a better determinant, than model scale per se, of non-climate predictor importance in SDMs at contrasting spatial scales.

The results of our study were consistent with our predictions for three of the five species examined. Our results supported predictions for *M. fuliginosus* and *M. rufus*, that is, that non-climate predictors would be more influential at landscape resolutions, compared with local resolutions, for species that exhibit extensive, adaptive long range movements. Our results also supported our prediction for *M. giganteus* that non-climate predictors would be more influential at local resolutions, compared with landscape resolutions for a species inhabiting small,
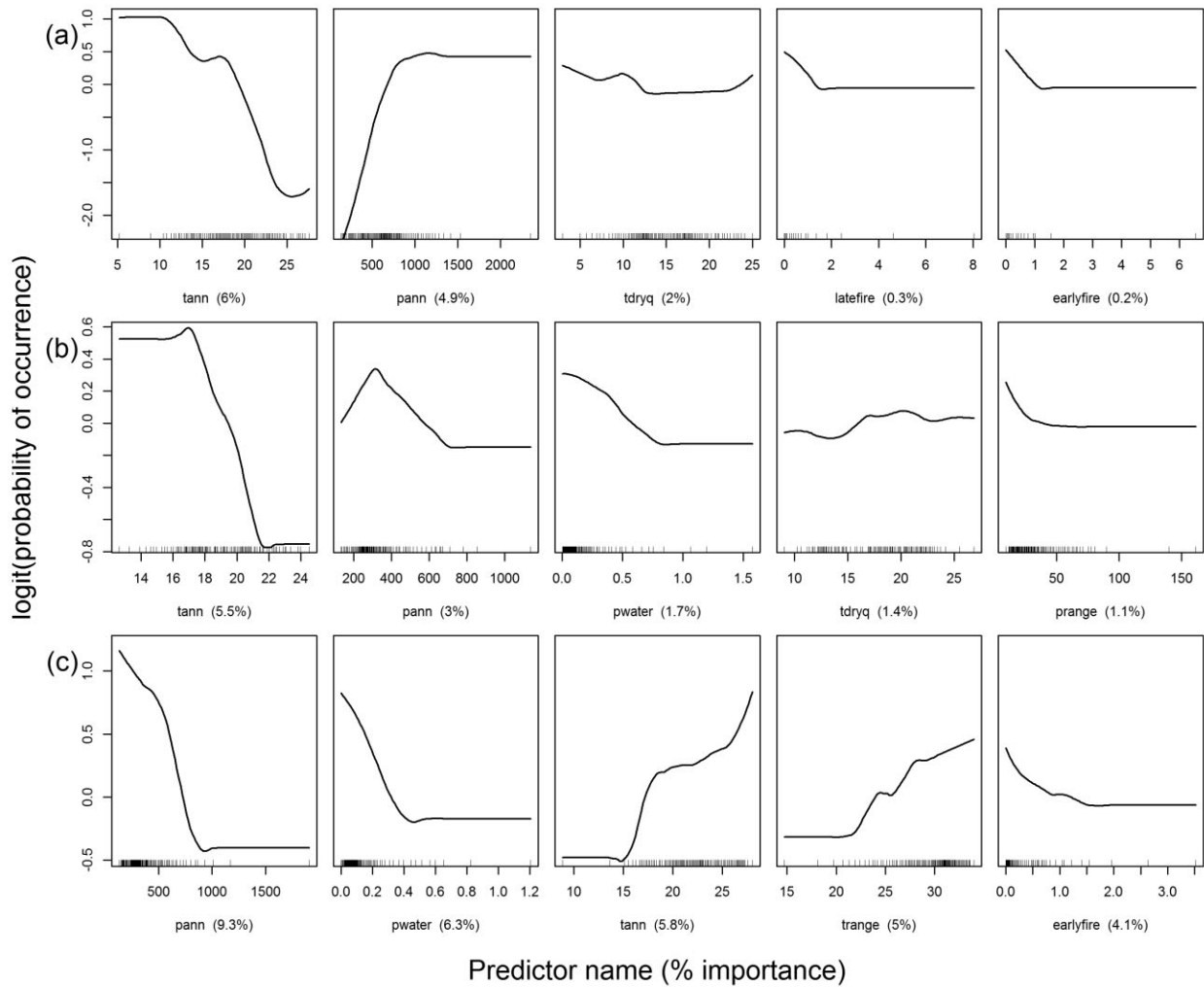
Fig. 3. Partial dependence functions for (a) *M. giganteus* from the 5 km climate + non-climate residuals autocovariate – boosted regression tree (RAC-BRT) models and (b) *M. fuliginosus* and (c) *M. rufus* from the 50 km climate + non-climate RAC-BRT models. Plots show the effect of the five most influential predictors on the response variable after accounting for the average effects of all other variables in the model. Predictors are earlyfire, early fire frequency; latefire, late fire frequency; pann, annual mean precipitation; prange, annual precipitation range; pwater, distance to closest permanent water; tann, annual mean temperature; tdryq, mean temperature of the driest quarter; trange, annual temperature range.

stable home ranges. Our findings were inconclusive for *M. robustus* and *M. antilopinus*.

Based on the species with definitive results, it appears that differences in species-specific movements better explain the importance of non-climate predictors at contrasting scales than model scale per se.

Ecological interpretation

The species ecological niches described by the RAC-BRT model fitted functions were consistent with the autecological literature.

The results for *M. giganteus* reflect its largely mesic preferences and predominant occurrence in the cooler eastern seaboard of Australia (Caughley *et al*. 1987;

Coulson 2008). In contrast to field studies (e.g. Southwell & Jarman 1987) our results suggest that the species prefers unburnt locations. It is possible that mesic conditions promote sufficient green grass over most of its range without the potentially destructive assistance of fire.

The plots for *M. fuliginosus* confirm the profile of a species that occupies a zone of low average annual temperature, where rainfall is winter dominated or uniform (hence low precipitation range) with an annual average (303 millimetres) half that for sites with *M. giganteus* (Caughley *et al*. 1987; Coulson 2008). Proximity to permanent water is more influential at the landscape resolution suggesting that *M. fuliginosus* may travel long distances (>5 km) to access water (Priddel *et al*. 1988a,b). The increased
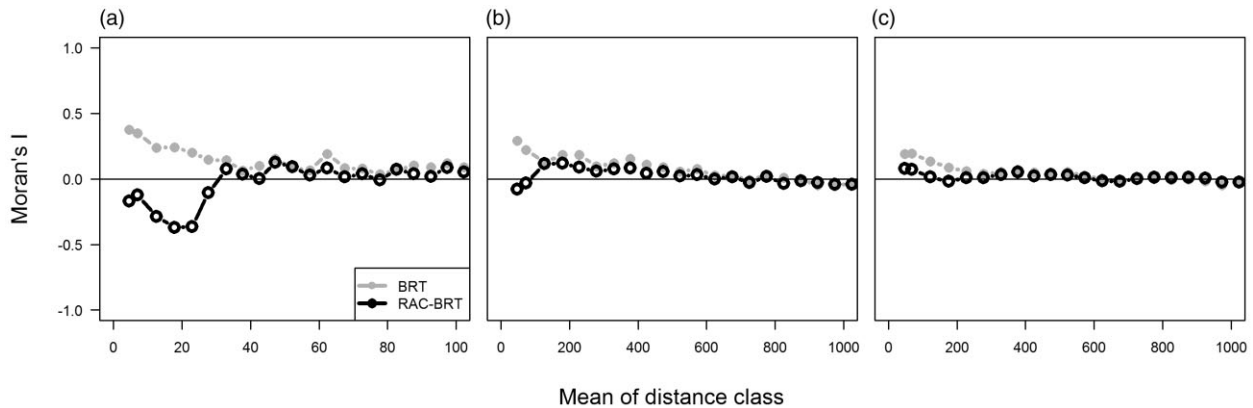
Fig. 4. Correlograms showing the reduction in first-order spatial autocorrelation of residuals autocovariate – boosted regression tree (RAC-BRT) model residuals compared with BRT model residuals for (a) *M. giganteus* 5 km climate + non-climate models, (b) *M. fuliginosus* 50 km climate + non-climate models and (c) *M. rufus* 50 km climate + non-climate models.

influence of average annual precipitation, possibly through interaction with non-climate predictors (resources) at landscape *versus* local resolution is also consistent with ecological knowledge as home ranges in the harsher climates of semiarid parts of the range span up to 7 km², ten times larger than in temperate regions (Coulson 2008).

The partial responses for *M. rufus* for the most influential variables confirm that it is predominantly restricted to the arid interior of Australia (Caughley *et al.* 1987; Croft & Clancy 2008). There is a preference for unburnt locations, at least with respect to early dry season fire. Opportunistic forays of around 50 km may facilitate location of green grass arising from localized rain (Croft & Clancy 2008) and as the weather becomes hotter and drier, individuals move further to seek remnants of green pasture resulting in a positive correlation between home range size and temperature. This explains the landscape resolution importance of temperature and its likely interactions with non-climate predictors/resources (Priddel *et al.* 1988a). The landscape resolution importance of permanent water may be an artefact of preferred habitats (hosting succulent forbs and grasses) being associated with permanent water, or could be directly linked to drinking-water availability, as individuals will travel further than 5 km to drink (Newsome 1975; McAlpine *et al.* 1999; Montague-Drake & Croft 2004; Fukuda *et al.* 2009).

The finding that non-climate predictors did not improve model fit or predictive performance at 50 km resolution for *M. giganteus* corresponds with our prediction of a local functional grain of resource use for this species. In other words, at 50 km the optimal spatial resolution has been surpassed and the relationship between non-climate predictors, climate predictors and the response variable (probability of occupancy) has diminished (Austin & Van Niel 2011).

In contrast, the finding that non-climate predictors, in conjunction with climate predictors, improved model fit and predictive performance more at 50 km than at 5 km resolution for *M. rufus* and *M. fuliginosus* is consistent with our prediction that a landscape resolution of 50 km is more representative of the radius of environmental influence on vital rates (Austin & Van Niel 2011). This is likely a consequence of a harsher climate and more variable productivity (in contrast to *M. giganteus*).

Ecologically informed model building

The focus of this paper is on the *relative* percent improvement not *absolute* values of model fit and predictive performance after the addition of non-climate predictors at two spatial resolutions. Nevertheless, it is noticeable that the model fit and predictive performance of models for most species/predictor set combinations was higher at 5 km resolution, a possible artefact of the larger sample size at finer resolution (Heikkinen *et al.* 2006). However, this tendency was not always seen in the predictive performance statistics from cross-validation suggesting that the variance explained by RAC may have been higher at finer resolutions in the original analyses. Further investigation, beyond the scope of this current paper, would provide further clarification. Depending on the purpose of modelling, 5 km resolution models may be suitable but model choice based on model fit or predictive performance per se could have serious implications for certain applications. For example, if comparing surface water management strategies and their effect on kangaroo occurrence, SDMs at 5 km resolution would not detect the relationship between *M. rufus* and permanent water distribution because individuals travel further than 5 km to drink before returning to

their original location to exploit optimal forage (Fensham & Fairfax 2008; Fukuda *et al.* 2009).

In addition to highlighting the contribution of non-climate predictors to model fit and predictive performance at ecologically relevant scales, our study also flags the potential importance of interactions between non-climate predictors, climate predictors and scale. For example, average annual precipitation had three times the importance at 50 km *versus* 5 km resolution in the climate-only model for *M. fuliginosus* but when non-climate predictors were added it became 16 times more important (Table 2). Exclusion of non-climate predictors may therefore bias predictions of current distribution and, potentially, forecasts of range change under future climates because effects of climate predictors on species distributions are modified by their interaction with non-climate predictors. Erroneous forecasts may be incurred, for example, by failed detection of refugia that buffer against the effects of a changing climate, preventing local extinctions that would otherwise be predicted using climate-only models. Surface water availability may provide such a refuge for kangaroos as survival may depend on artificial water points when climatic conditions deteriorate (Fukuda *et al.* 2009).

Our approach provides an opportunity to bridge the gap between recent progress in understanding movement behaviour of individuals and the emergent spatial dynamics at the population level (Mueller *et al.* 2011). Our case study focused on macropods, a group for which space use and social organization are closely correlated with variation in habitat productivity (Fisher & Owens 2000), but our approach could be applied to any mobile species. Expert opinion and/or literature review should advise on selection of non-climate predictors that directly or indirectly represent limiting factors (e.g. food, water, mate, competitor, predator or den/nest site distributions) and the resolution/s at which they might be most influential, based on space use or resource use patterns of the species. It would be interesting to see whether the scale-dependence of non-climate predictors can be predicted from species autecology for a greater number of species.

ACKNOWLEDGEMENTS

REFERENCES

Araújo M. B. & Peterson A. T. (2012) Uses and misuses of bioclimatic envelope modelling. *Ecology* **93,** 1527–39.

Austin M. P. & Van Niel K. P. (2011) Improving species distribution models for climate change studies: variable selection and scale. *J. Biogeogr.* **38,** 1–8.

Bertrand R., Perez V. & Gégout J.-C. (2012) Disregarding the edaphic dimension in species distribution models leads to the omission of crucial spatial information under climate change: the case of *Quercus pubescens* in France. *Glob. Chang. Biol.* **18,** 2648–60.

Bjornstad O. N. (2013) ncf: Spatial nonparametric covariance functions. R package version 1.1-5. [Cited 18 Dec 2013.] Available from URL: http://CRAN.R-project.org/package =ncf

Caughley G., Short J., Grigg G. C. & Nix H. (1987) Kangaroos and climate – an analysis of distribution. *J. Anim. Ecol.* **56,** 751–61.

Clancy T. F. & Croft D. B. (1990) Home range of the common wallaroo, *Macropus robustus erubescens*, in far western New South Wales. *Aust. Wildl. Res.* **17,** 659–73.

Clancy T. F. & Croft D. B. (2008) Common wallaroo, *Macropus robustus*. In: *The Mammals of Australia*, 3rd edn (eds S. Van Dyck & R. Strahan) pp. 346–8. New Holland Publishers, Sydney.

Coulson G. (2008) Eastern Grey Kangaroo, *Macropus giganteus*. In: *The Mammals of Australia*, 3rd edn (eds S. Van Dyck & R. Strahan) pp. 335–8. New Holland Publishers, Sydney.

Crase B., Liedloff A. C. & Wintle B. A. (2012) A new method for dealing with residual spatial autocorrelation in species distribution models. *Ecography* **35,** 879–88.

Croft D. B. (1991) *Home range of the euro*, Macropus robustus erubescens. *J. Arid Environ.* **20,** 99–111.

Croft D. & Clancy T. (2008) *Red Kangaroo*, Macropus rufus. In: *The Mammals of Australia*, 3rd edn (eds S. Van Dyck & R. Strahan) pp. 352–4. New Holland Publishers, Sydney.

Dawson T. J. & Denny M. J. S. (1969) A bioclimatological comparison of the summer day microenvironments of two species of arid-zone kangaroo. *Ecology* **50,** 328–32.

Dawson T. J., Denny M. J. S., Russell E. M. & Ellis B. (1975) Water usage and diet preferences of free-ranging kangaroos, sheep and feral goats in the Australian arid zone during summer. *J. Zool. (Lond.)* **177,** 1–23.

Elith J., Leathwick J. R. & Hastie T. (2008) A working guide to boosted regression trees. *J. Anim. Ecol.* **77,** 802–13.

Fensham R. J. & Fairfax R. J. (2008) Water-remoteness for grazing relief in Australian arid-lands. *Biol. Conserv.* **141,** 1447–60.

Fisher D. O. & Owens I. P. F. (2000) Female home range size and the evolution of social organization in macropod marsupials. *J. Anim. Ecol.* **69,** 1083–98.

Fordham D. A., Akcakaya H. R., Araujo M. B. *et al.* (2012) Plant extinction risk under climate change: are forecast range shifts alone a good indicator of species vulnerability to global warming? *Glob. Chang. Biol.* **18,** 1357–71.

Friedman J. H. & Meulman J. J. (2003) Multiple additive regression trees with application in Epidemiology. *Stat. Med.* **22,** 1365–81.

Fukuda Y., Mccallum H. I., Grigg G. C. & Pople A. R. (2009) Fencing artificial waterpoints failed to influence density and distribution of red kangaroos (*Macropus rufus*). *Wildl. Res.* **36,** 457–65.

Gillingham P. K., Palmer S. C. F., Huntley B., Kunin W. E., Chipperfield J. D. & Thomas C. D. (2012) The relative

importance of climate and habitat in determining the distributions of species at different spatial scales: a case study with ground beetles in Great Britain. *Ecography* **35**, 831–8.

Heikkinen R. K., Luoto M., Araujo M. B., Virkkala R., Thuiller W. & Sykes M. T. (2006) Methods and uncertainties in bioclimatic envelope modelling under climate change. *Prog. Phys. Geogr.* **30**, 751–77.

Hijmans R. J., Phillips S., Leathwick J. & Elith J. (2013) dismo: Species distribution modeling. R package version 0.9-3. [Cited 18 Dec 2013.] Available from URL: http://CRAN .R-project.org/package=dismo

Hijmans R. J. (2013) raster: Geographic data analysis and modeling. R package version 2.2-5. [Cited 18 Dec 2013.] Available from URL: http://CRAN.R-project.org/package =raster

Hirst S. (2006) *The Antilopine Wallaroo: An Unusual 'Roo.* National Wildlife Rehabilitation Conference, Darwin.

Liu C. R., Berry P. M., Dawson T. P. & Pearson R. G. (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* **28**, 385–93.

Luoto M. & Heikkinen R. K. (2008) Disregarding topographical heterogeneity biases species turnover assessments based on bioclimatic models. *Glob. Chang. Biol.* **14**, 483–94.

Luoto M., Virkkala R. & Heikkinen R. K. (2007) The role of land cover in bioclimatic models depends on spatial resolution. *Glob. Ecol. Biogeogr.* **16**, 34–42.

McAllister R. R. J., Smith D. M. S., Stokes C. J. & Walsh F. J. (2009) Patterns of accessing variable resources across time and space: desert plants, animals and people. *J. Arid Environ.* **73**, 338–46.

McAlpine C. A., Grigg G. C., Mott J. J. & Sharma P. (1999) Influence of landscape structure on kangaroo abundance in a disturbed semi-arid woodland of Queensland. *Rangel. J.* **21**, 104–34.

Montague-Drake R. & Croft D. B. (2004) Do kangaroos exhibit water-focused grazing patterns in arid New South Wales? A case study in Sturt National Park. *Aust. Mammal.* **26**, 87–100.

Mueller T. & Fagan W. F. (2008) Search and navigation in dynamics environments – from individual behaviors to population distributions. *Oikos* **117**, 654–64.

Mueller T., Olson K. A., Dressler G. *et al.* (2011) How landscape dynamics link individual to population-level movement patterns: a multispecies comparison of ungulate relocation data. *Glob. Ecol. Biogeogr.* **20**, 683–94.

Newsome A. E. (1975) An ecological comparison of the two arid-zone kangaroos of Australia, and their anomalous prosperity since the introduction of ruminant stock to their environment. *Q. Rev. Biol.* **50**, 389–424.

Norbury G. L., Norbury D. C. & Oliver A. J. (1994) Facultative behaviour in unpredictable environments: mobility of red kangaroos in arid Western Australia. *J. Anim. Ecol.* **63**, 410–18.

Opdam P. & Wascher D. (2004) Climate change meets habitat fragmentation: linking landscape and biogeographical scale levels in research and conservation. *Biol. Conserv.* **117**, 285–97.

Pearson R. G. & Dawson T. P. (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Glob. Ecol. Biogeogr.* **12**, 361–71.

Priddel D., Shepherd N. & Wellard G. (1988a) Home ranges of sympatric red kangaroos *Macropus rufus*, and western grey kangaroos *M. fuliginosus*, in western New South Wales. *Aust. Wildl. Res.* **15**, 405–11.

Priddel D., Wellard G. & Shepherd N. (1988b) Movements of sympatric red kangaroos. *Macropus rufus*, and western grey kangaroos, *M. fuliginosus* in western New South Wales. *Aust. Wildl. Res.* **15**, 339–46.

Ridgeway G. (2013) gbm: Generalized Boosted Regression Models. R package version 2.1. [Cited 18 Dec 2013.] Available from URL: http://CRAN.R-project.org/package=gbm

Ritchie E. G. (2008) Antilopine wallaroo, *Macropus antilopinus*. In: *The Mammals of Australia*, 3rd edn (eds S. Van Dyck & R. Strahan) pp. 325–6. New Holland Publishers, Sydney.

Ritchie E. G., Martin J. K., Krockenberger A. K., Garnett S. & Johnson C. N. (2008) Large-herbivore distribution and abundance: Intra-and interspecific niche variation in the tropics. *Ecol. Monogr.* **78**, 105–22.

Roubicek A. J., VanDerWal J., Beaumont L. J., Pitman A. J., Wilson P. & Hughes L. (2010) Does the choice of climate baseline matter in ecological niche modelling? *Ecol. Model.* **221**, 2280–6.

Skidmore A. K. (1998) Nonparametric classifier for GIS data applied to kangaroo distribution mapping. *Photogramm. Eng. Rem. S.* **64**, 217–26.

Skidmore A. K., Gauld A. & Walker P. (1996) Classification of kangaroo habitat distribution using three GIS models. *Int. J. Geogr. Inf. Syst.* **10**, 441–54.

Sormunen H., Virtanen R. & Luoto M. (2011) Inclusion of local environmental conditions alters high-latitude vegetation change predictions based on bioclimatic models. *Polar Biol.* **34**, 883–97.

Southwell C. J. & Jarman P. J. (1987) Macropod studies at Wallaby Creek III. The effect of fire on pasture utilization by macropodids and cattle. *Aust. Wildl. Res.* **14**, 117–24.

Tingley R. & Herman T. B. (2009) Land-cover data improve bioclimatic models for anurans and turtles at a regional scale. *J. Biogeogr.* **36**, 1656–72.

Titeux N., Maes D., Marmion M., Luoto M. & Heikkinen R. K. (2009) Inclusion of soil data improves the performance of bioclimatic envelope models for insect species distributions in temperate Europe. *J. Biogeogr.* **36**, 1459–73.

VanDerWal J., Shoo L. P., Graham C. & William S. E. (2009) Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know? *Ecol. Model.* **220**, 589–94.

Viggers K. L. & Hearn J. P. (2005) The kangaroo conundrum: home range studies and implications for land management. *J. Appl. Ecol.* **42**, 99–107.

Walker P. A. (1990) Modeling wildlife distributions using a geographic information system – kangaroos in relation to climate. *J. Biogeogr.* **17**, 279–89.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1.   Data sources for species occurrences.
Appendix S2.   Factors influencing the distribution or abundance of the five kangaroo species.
Appendix S3.   Data sources and processing methods for non-climate predictors.