Compositional patterns in terrestrial fauna and wetland flora and fauna across the Pilbara biogeographic region of Western Australia and the representativeness of its conservation reserve system

L.A. Gibson[1,*], K.J. Williams[2], A.M. Pinder[1], T.D. Harwood[2], N.L. McKenzie[1], S. Ferrier[2], M.N. Lyons[1], A.H. Burbidge[1] and G. Manion[3]

[1] Department of Parks and Wildlife, Locked Bag 104, Bentley Delivery Centre, Western Australia 6983, Australia.

[2] CSIRO Land and Water Flagship, GPO Box 1600, Canberra, Australian Capital Territory 2601, Australia.

[3] New South Wales Office of Environment and Heritage, University of New England, New South Wales 2351, Australia.

* Corresponding author: email: Lesley.Gibson@dpaw.wa.gov.au

Abstract – A biological survey of the Pilbara biogeographic region was undertaken between 2002 and 2007 to provide a regional perspective on biodiversity patterns as a contribution to nature conservation planning. During this survey, 304 sites were sampled for small ground-dwelling mammals, birds, reptiles, spiders, ants, beetles and scorpions. A further 98 sites were sampled for wetland invertebrates, aquatic macrophytes and fringing riparian vegetation. Data for these two groups of sites were aggregated separately (i.e. terrestrial fauna and wetland biodiversity) and models of turnover in species composition within each data set were developed using generalised dissimilarity modelling (GDM). A wide range of environmental variables was assessed as predictors of compositional turnover – biotic (vegetation cover indices), climate, landform, hydrologic, regolith (soil and geology) and geographic distance. Generally, predictors associated with regolith were the most strongly supported in both the terrestrial fauna and wetland biodiversity models, followed by combined landform/ hydrologic variables, then climate/biotic variables. Geographic distance between sites was retained in the terrestrial fauna model only. The final GDM models explained 46.1% and 58.5% of the deviance in the compositional turnover of terrestrial fauna and wetland biodiversity, respectively. Spatial representation of the coverage of survey sites showed that a large proportion of the core study area was well represented for both terrestrial fauna and wetland biodiversity. However, gaps in the proportional representation of both groups within the 2011 conservation reserve system were evident, particularly in the coastal region of the Pilbara (Roebourne subregion) and the Fortescue River valley (Fortescue subregion). With the addition of proposed reserves (in 2015) within these two subregions, the representation of terrestrial fauna and wetland biodiversity was substantially improved.

Keywords – biodiversity survey, compositional turnover, gap analysis, GDM, Pilbara, reserve representativeness

INTRODUCTION

Due to its rich mineral reserves, the Pilbara region of Western Australia is of major national economic importance. The region is also known for its visually striking landscapes and it attracts many thousands of visitors every year. Along with its economic and aesthetic values, the immense biodiversity values of the Pilbara are just being fully recognised. Largely in response to mining proposals, there have been numerous localised surveys in the Pilbara. To place these surveys into context, there was a clear need for broad-scale, systematically collected data to provide a regional perspective on nature conservation priorities (McKenzie *et al.* 2009). This first comprehensive biological survey of the Pilbara region was conducted by the Western Australian Department of Parks and Wildlife (formerly Department of Environment and Conservation) between 2002 and 2007 and uncovered a wealth of information on Pilbara biota. The survey was designed to sample

L.A. Gibson, K.J. Williams, A.M. Pinder, T.D. Harwood, N.L. McKenzie, S. Ferrier, M.N. Lyons, A.H. Burbidge and G. Manion

516

a wide range of organisms to provide a spatially explicit biodiversity model for the region as a basis for assessing gaps in the existing conservation reserve system (McKenzie *et al.* 2009).

During the survey, 304 sites were sampled for small ground-dwelling mammals, birds, reptiles, ground-dwelling spiders, ants, beetles and scorpions. A further 98 sites, centred on water bodies, were sampled for aquatic invertebrates, aquatic macrophytes (charophytes and vascular plants) and the fringing riparian vegetation. The sites were positioned across the geographical extent of the region in a stratified array to sample the main climatic gradients, geological formations, landforms, river catchments, soils and vegetation types. The survey findings, published as a series of papers in a dedicated Supplement of the *Records of the Western Australian Museum*, describe the relationship between the region's physical environment and components of its biota (Gibson and McKenzie 2009, McKenzie *et al.* 2009; Burbidge *et al.* 2010; Durrant *et al.* 2010; Guthrie *et al.* 2010; Heterick *et al.* 2010; Pinder *et al.* 2010; Volschenk *et al.* 2010; Doughty *et al.* 2011). Each of these taxon-specific papers related biodiversity survey data to environmental attributes at the site level. Here, we combine the data and develop community-level spatially explicit models of biodiversity patterns by relating the point-based biological data to continuous mapped environmental data using generalised dissimilarity modelling (GDM, Ferrier *et al.* 2007).

GDM is a statistical technique for modelling the compositional dissimilarity (i.e. spatial turnover of species composition) between pairs of geographical locations, as a non-linear multivariate function of summed environmental distances and, optionally, geographic distance, between these locations. The compositional dissimilarity between a given pair of locations can be thought of as the proportion of species occurring at one location that do not occur at the other location (averaged across the two locations) – ranging from '1' if the two locations have exactly the same species through to '0' if they have no species in common. GDM effectively weights and transforms the environmental variables such that distances between locations in this transformed multidimensional environmental space now correlate, as closely as possible, with observed biological compositional dissimilarities (see Ferrier *et al.* 2007). GDM performs best with comprehensive biological survey data on species presence and absence representatively sampled at sites across the region of interest – the type of data provided by the Pilbara Biodiversity Survey.

Once a GDM has been fitted to the biological data from the sampled locations, it can be used to predict compositional dissimilarity values for sites lacking biological data, based on their mapped environmental attributes. This predictive capacity of GDM provides a foundation for various subsequent spatial analyses such as an appraisal of the representativeness of conservation reserve networks (Ferrier *et al.* 2004; Overton *et al.* 2009; Thomassen *et al.* 2011), estimating biodiversity loss (Allnutt *et al.* 2008), survey gap analyses (Ferrier 2002; Funk *et al.* 2005), conservation prioritisation (Arponen *et al.* 2008; Thomassen *et al.* 2010) and climate-change impacts (Fitzpatrick *et al.* 2011; Mackey *et al.* 2012; Prober *et al.* 2012).

Here, we describe the application of GDM to develop models of species compositional turnover for the Pilbara, and the results of applying these models to assess the representativeness of the region's conservation reserve network and gaps in biological survey coverage as a measure of model predictability.

METHODS

Study area

The Pilbara biogeographic region (~179,000 km^2 in area) comprises four geomorphically distinctive subregions: undulating granite and basalt plains including significant areas of basaltic ranges (Chichester subregion), alluvial plains and river frontages (Fortescue Plains subregion), mountainous sedimentary ranges and plateaux dissected by gorges (Hamersley subregion), and alluvial and older colluvial coastal and sub-coastal plains (Roebourne subregion) (Figure 1; see McKenzie *et al.* 2009 for detailed descriptions). Wetland sampling also extended into the Ashburton and Augustus subregions of the Gascoyne IBRA region so as to include the upper reaches of the Ashburton Catchment (Figure 2).

The climate of the Pilbara region is dominated by annual and inter-annual cycles of wetting and drying. January, February and March are the wettest months while September and October are the driest (McKenzie *et al.* 2009). There is considerable variation in rainfall between years due to cyclones that occasionally cross the coast. Rainfall intensity can be high with thunderstorms and cyclones generating high runoff volumes, and fluvial patterns of erosion and deposition are apparent in the landform. Monthly maximum temperatures range from an average of 25.3ºC in July to 37.8ºC in January, and minimum temperatures from an average of 11.8ºC in July to 25.2ºC in January (Leighton 2004). The Pilbara straddles two bioclimatic regions (Beard 1990). The higher rainfall areas inland (Hamersley Plateau) and the relatively cooler areas near the coast have a semi-desert tropical climate. The other bioclimatic region has a desert climate with generally higher temperatures.
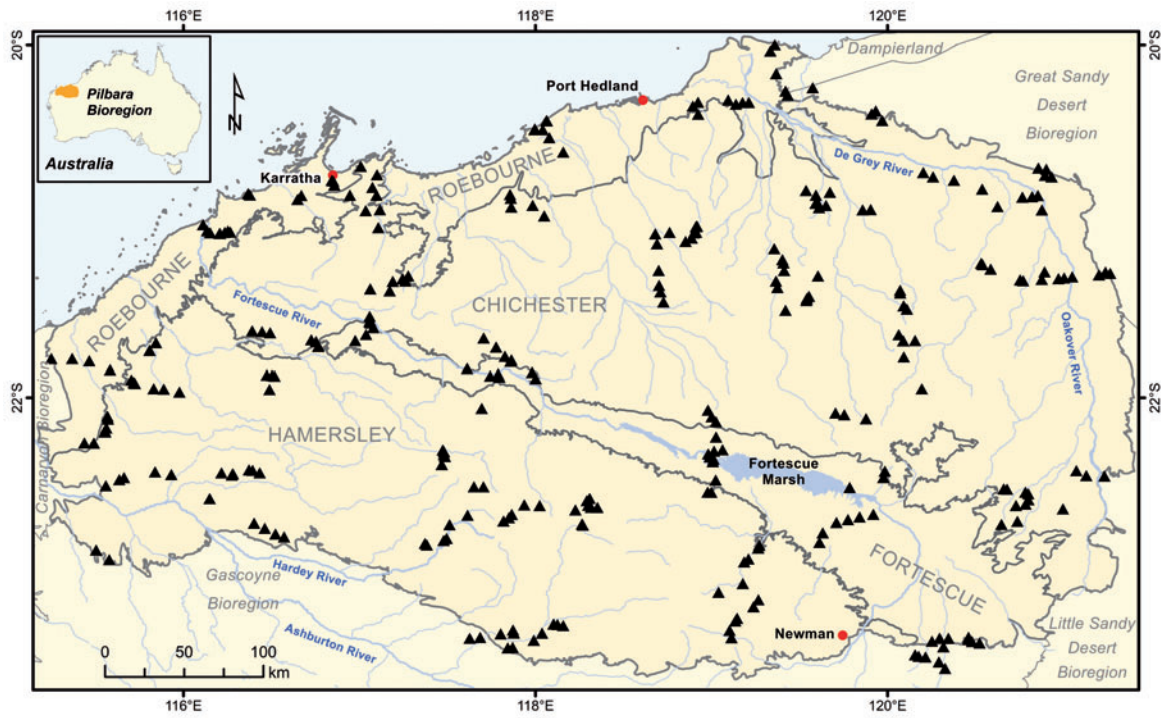
Figure 1 Study area and analysis domain used in developing generalised dissimilarity models for terrestrial fauna in the Pilbara biogeographic region showing terrestrial survey sites (black triangles), subregions (dark grey lines), major towns (red circles) and main rivers (blue lines). Surrounding bioregions are labelled: Carnarvon, Gascoyne, Little Sandy Desert, Great Sandy Desert and Dampierland.
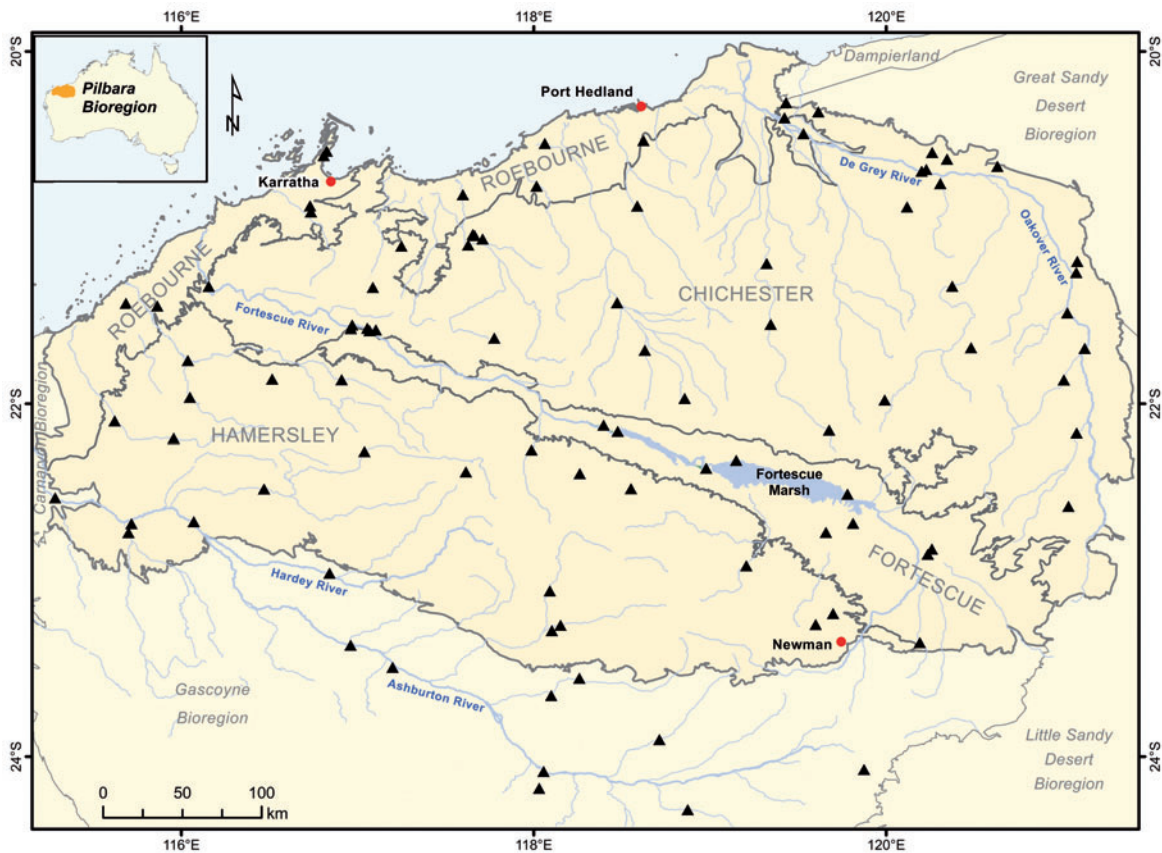

Figure 2 Study area and analysis domain used in developing a generalised dissimilarity model for wetland biodiversity in the Pilbara biogeographic region showing aquatic survey sites (black triangles), subregions (dark grey lines), major towns (red circles) and main rivers (blue lines). Surrounding bioregions are labelled: Carnarvon, Gascoyne, Little Sandy Desert, Great Sandy Desert and Dampierland.

L.A. Gibson, K.J. Williams, A.M. Pinder, T.D. Harwood, N.L. McKenzie, S. Ferrier,
M.N. Lyons, A.H. Burbidge and G. Manion

518

The soils of the Pilbara region are generally skeletal and either derived *in situ* or deposited as colluvium or alluvium; their colours reflect the underlying parent material (McKenzie *et al.* 2009).

Wetland habitats include pools within otherwise dry river channels, which range from temporary to permanently filled, temporary streams, mostly permanent springs and temporary floodplain wetlands (mostly claypans but also some isolated billabongs). The largest of the lentic wetlands is the 100 km long Fortescue Marsh formed by the Goodiarie Hills creating a barrier to flow within the Fortescue River Valley (shown in Figure 2). This marsh is fresh when episodically filled and becomes saline as it dries.

For the GDM analysis, the study domain was extended to include the surrounding areas. It encompassed all terrestrial and wetland survey sites and incorporated environmental gradients across the area. The spatial analysis unit is a 9-second (9sec) geographic grid (~250 m) in GDA94, approximating the size of the site for field survey sampling; ca. 1 ha for most zoological groups (McKenzie *et al.* 2009).

Biological data

The biological data (i.e. species presence/absence) had two distinct components: 1) the 304 terrestrial fauna sites (Figure 1), and 2) the 98 wetland sites (Figure 2). Data on the plant taxa recorded from terrestrial sites during the Pilbara Biodiversity Survey were not available at the time of this analysis.

Terrestrial fauna data

Of the 304 terrestrial sites, 296 were included in the analysis. The other eight were burnt in wildfires, stripped by floods or inundated by storm tides before sampling could be completed. Survey sites were located in habitat examples that were the least disturbed. Data for two sites (DRW01 and DRW02) that fell within the same 9sec grid cell were aggregated, resulting in 295 unique grid cells (Table 1). The terrestrial fauna data comprised 1324 species belonging to three vertebrate groups (birds, reptiles and mammals) and four invertebrate groups (ants, beetles, scorpions and spiders) (Table 1).

Wetland data

The aquatic macrophytes, aquatic invertebrates and riparian flora were sampled at the same water bodies. Thus, to model turnover in species composition, the species lists were aggregated to form a single site by species matrix (summarised in Table 2). Three clustered sites were excluded from the analysis because they were geographically isolated (i.e. 100 km east of the Pilbara), leaving 95 sites for the GDM analysis.

Environmental data

Environmental data were compiled according to a general conceptual model of the relationship between species diversity and habitat (Williams *et al.* 2012). For terrestrial diversity, the general model assumes responses to both physical and biological components of their environment (e.g. McKenzie *et al.* 2000b, 2004; George *et al.* 2011; McKenzie and Bullen 2012). The physical environment can be described by facets of climate, regolith, hydrology and landform and the biological environment by vegetation patterns or indirectly by the physical environmental correlates of vegetation patterns. We collated spatial environmental data for the Pilbara within these broad classes or their proxies (Appendix 1). Geographic distance between sampling localities was included because it may be independently associated with metapopulation range and dispersal processes that isolate populations, or be correlated with unexplained environmental variation. Additional variables (Appendix 1) were assembled to describe wetland environments based on a scientific understanding of aquatic systems in the Pilbara region (Pinder *et al.* 2010).

Model fitting and variable selection strategy

The GDM procedure firstly involves the calculation of Sørensen dissimilarity (Sørensen 1948) between pairs of sites to produce a site by site dissimilarity matrix as the response variable. The response variable is then fitted to the environmental variables (each as a distance matrix) using a generalised linear model as described by Ferrier *et al.* (2007). We used the .NET General Dissimilarity Modeller software (Manion 2014) to fit a GDM to the Pilbara Biodiversity Survey data (functions have now been incorporated into a R-package, Manion *et al.* 2015).

The variable selection strategy followed the stage-wise process outlined by Williams *et al.* (2010a, 2012) where different groups of correlated variables are initially tested for redundancy before combining the retained variables from each group and then removing relatively insignificant variables using a backward elimination procedure that tests the contribution of each variable. Preliminary models explored the effectiveness of a large number of candidate predictors grouped into climate (20 variables), biotic (4), regolith (32), hydrologic (6) and landform (15) (Appendix 1). The more marginal of the remaining retained predictors were successively removed using a stopping criterion of 0.05% partial deviance explained. This value was determined to be a reasonable trade-off in parsimony between the number of predictors included in the model and cumulative reduction in deviance explained by the model. Each predictor was evaluated using five splines defined by their quantile data positions

Table 1 Number of species, sites and records in each component included in the terrestrial fauna analysis (n.b. two quadrats were aggregated into a single grid).

Group	Number of species	Number of 9sec grid cells (sites)	Number of species by grid cell records	Number of grid cell 'site'-pairs representing group
Ants	245	295	5045	43,365
Beetles	427	295	3421	43,365
Scorpions	20	235	434	27,495
Spiders	376	293	3772	42,778
Birds	128	295	5557	43,365
Reptiles	108	295	3081	43,660
Mammals	20	294	1038	43,071
All species	1324	295	22,348	43,365

Table 2 Number of species, sites and records for each dataset included in the wetland biodiversity analysis. The water column assemblages include relatively speciose lists of algae and protozoans (Pinder *et al.* 2010).

Group	Number of species	Number of 9sec grid cells (sites)	Number of species by grid cell records	Number of grid cell 'site'-pairs representing group
Water column sites (aggregated)	1075	98	12,876	4753
Riparian flora sites (aggregated)	454	98	2679	4753

(0, 25, 50, 75, and 100%) to capture non-linear variation between compositional turnover and each environmental gradient.

Three broad groupings of the 77 candidate variables (biotic/climate, regolith and hydrologic/landform) were evaluated for their unique and shared statistical contributions to explained deviance using variance partitioning calculations (Borcard *et al.* 1992; Jones *et al.* 2013). Proportional set diagrams were developed using *eulerAPE* software by Micallef and Rodgers (2014).

Scaling dissimilarity between plot and spatial resolution

To transform our site-based compositional dissimilarity values so that they were relevant to the 9sec grid cell resolution, we applied the biodiversity scaling method in Mokany *et al.* (2013). Briefly, this approach applies the species-area power model ($S = c\,A^z$) to scale both species richness and compositional dissimilarity from small sample areas to larger areas. Under this approach, the species richness of a grid cell (S) is predicted from the observed richness of the local sample area (c) and the area of the grid cell relative to the survey area. To scale pair-wise compositional dissimilarity from the point-based surveys to the grid cells they occurred within, we first predict the number of species in common between the two grid cells i and j ($S_{com,ij}$) from the observed number of species in common between the two local sample areas ($c_{com,ij}$) using the species area power relationship ($S_{com,ij} = c_{com,ij}A^{zcom}$). We then calculate the predicted Sørensen's compositional dissimilarity between the two grid cells ($\beta_{ij} = 1 - [2S_{com,ij}\, / \,(S_i + S_j)]$) using the predicted species richness of each grid cell (S_i, S_j) and the predicted number of species in common between the two grid cells ($S_{com,ij}$). We applied a single scaling factor for species richness ($z = 0.25$) and for the number of species shared between a pair of local sample areas ($z_{com} = 0.42$), as derived previously (Mokany *et al.* 2013). This scaling approach retains the underlying gradients in species richness and compositional dissimilarity observed across the survey plots, but scales the site-based absolute values so that they better represent those of the larger grid cells being modelled.
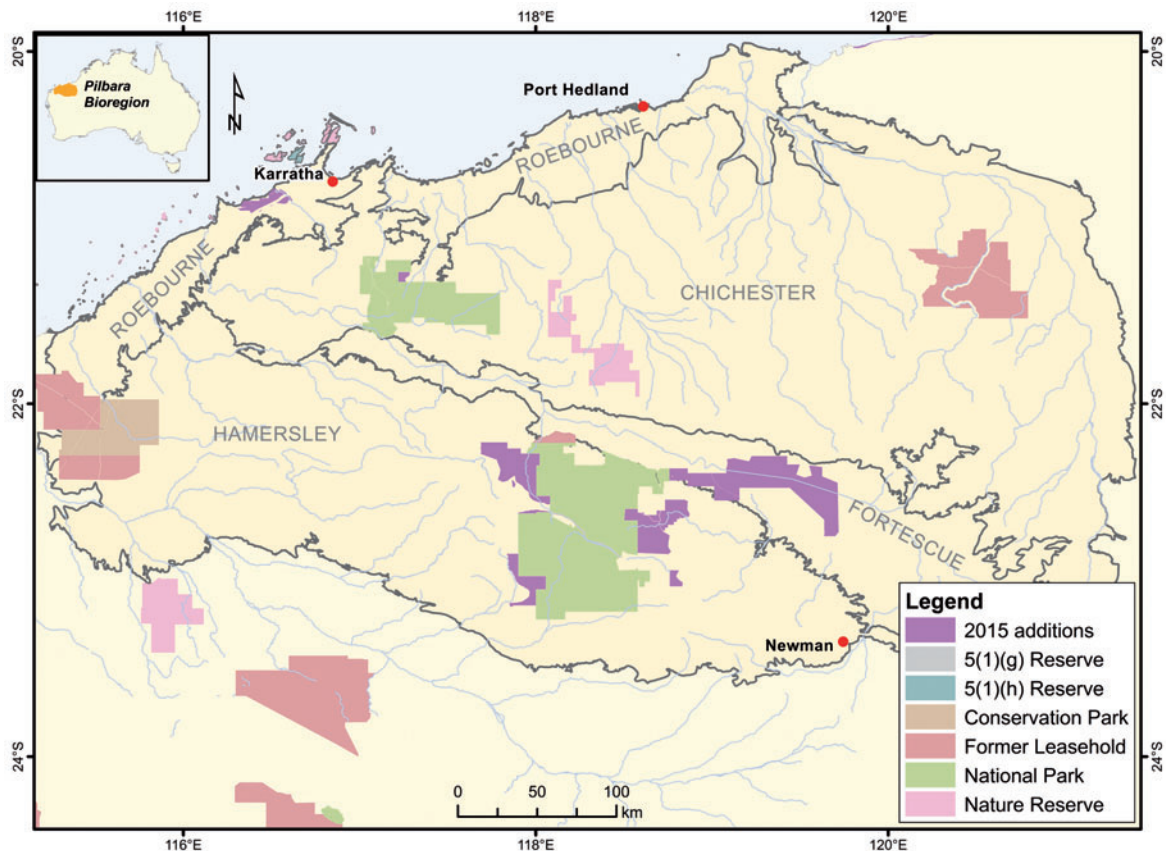
L.A. Gibson, K.J. Williams, A.M. Pinder, T.D. Harwood, N.L. McKenzie, S. Ferrier,
M.N. Lyons, A.H. Burbidge and G. Manion

520

Figure 3 Location of conservation tenures managed by the Western Australian Department of Parks and Wildlife (current as at June 2011) in the Pilbara biogeographic region. The proposed 2015 addition boundaries (2015 additions in legend) are indicative only as they are undergoing refinement. Subregions, major towns and main rivers are also shown.

Classification of compositional turnover

To visualise the spatial structure of compositional turnover, we clustered a sample of 50,000 grid-cells evenly spread in geographic space across the study area, and derived an agglomerative hierarchical classification (UPGMA; Sneath and Sokal 1973) of predicted Sørensen dissimilarity (Ferrier *et al.* 2007). We used 300 classes (clusters of grid cells) to represent patterns of terrestrial fauna compositional turnover and 100 classes for wetland biodiversity. Each class was coloured based on multi-dimensional scaling of predicted similarity, whereby the three primary axes are assigned red-green-blue values, such that similar colours are predicted to have similar biological composition (Belbin *et al.* 1983).

The spatial analysis domain, which extends beyond the sampled locations, leads to model extrapolation in environmental space and this can influence the assignment and colouring of classes. The start and end shapes of the predictor fitted function and its relative importance in the model influence the degree of extrapolation. The GDM software linearly extrapolates the fitted function based on the last 10% of data at either end, as applicable. The extrapolation index, derived as the absolute sum for each predictor, provides an objective means by which to map these areas. The classification was therefore applied both where the model interpolates and where extrapolation was minimal. The threshold used in each case was determined from the first of a four-class geometric interval histogram containing the majority of grid cells with minor extrapolation.

Assessing gaps in the conservation reserve system

The existing conservation reserve system in the Pilbara study area is managed by the Western Australian Department of Parks and Wildlife (Parks and Wildlife). The reserve network comprises various land tenures (Figure 3) with differing IUCN classifications (IUCN 1994). A number of additions to the current conservation estate (as at June 2011) have been proposed up to 2015. These areas are shown as '2015 additions' in Figure 3, although their boundaries are indicative only.

Consequently, we analysed gaps in the reserve system, 1) managed by Parks and Wildlife in June 2011, and 2) post-2015.

Using existing approaches, we derived a continuous metric describing the spatial distribution of reserve representativeness for both the terrestrial and wetland systems (Ferrier *et al.* 2004; Allnut *et al.* 2008). The fitted GDM model predicts the compositional similarity of any 9sec grid cell to any other grid cell in the Pilbara, on a scale ranging from 0 (dissimilar) to 1 (compositionally identical). The compositional similarity (= 1-Sørensen) between each cell and every other cell in the region was calculated. The sum total of all these similarities for each cell provides a baseline scaled area of the habitat similar to each cell. Rare habitats will therefore have a low area contribution, and more widely distributed habitats will have a greater area contribution. This calculation may also be viewed as the scaled total area which could potentially be reserved across the study area. Following the approach used in the assessment of the representativeness of the National Reserve System under climate change (Ferrier *et al.* 2010), the total scaled area of habitat similar to each cell in the reserve system was also calculated, i.e. the area of habitat which would be available to each cell class if all non-reserved land were removed. By dividing the scaled area of habitat in reserves by the scaled area of habitat which could be reserved (everywhere) for each cell, a metric describing the fraction of similar habitat reserved can be calculated. This can be directly interpreted as the representativeness of each cell with respect to a given reserve system. This metric was derived for both the current reserve estate and the proposed future estate as per objectives 1) and 2) above (see Figure 3).

Using biological survey coverage as a measure of spatial uncertainty

Paralleling the approach to assessing the environmental coverage of reserves, we evaluated the representativeness of the terrestrial or wetland survey sites as a continuous fraction of the Pilbara's total area, again scaled by the GDM model predictions of environmental similarity. This analysis yielded an estimate of relative survey density for the environment associated with every grid cell in the study area. These estimates were then used to indicate spatial uncertainty in the fitted GDM, i.e. those parts of the environmental space across the Pilbara study area in which the reserve gap analysis will be least reliable.

RESULTS

Terrestrial fauna GDM

The GDM model fit, or observed *versus* predicted compositional dissimilarity, is presented in Figure 4 and the relative contribution of predictors in Figure 5. Of the 77 candidate environmental variables (Appendix 1), 34 were retained in the final GDM model (Figure 5, Appendix 2). Geographic distance was also found to be significant. The resulting model explained 46.1% of the deviance with an intercept of 0.29 and sum of predictor spline coefficients of 5.65, which is an indicator of the overall magnitude of compositional turnover predicted by the model. The partial contribution to the percent deviance explained for each predictor variable is shown in Appendix 2.

The four most important predictors contributing to compositional turnover of the terrestrial fauna (Figure 5) were the first principal component of surficial soil spectra (SOILspectra1), which mainly correlates with the distribution of highly weathered landscapes with soils that contain large amounts of hematite and which occur in association with kaolinite clay minerals (Viscarra Rossel and Chen 2011), percent abundance of illite clay minerals in the top soil (TopSOILillite), surficial soil clay content (SOILclay) and the area of woody vegetation cover (WOODYVEGcover). Overall, predictors associated with regolith appear to be the main contributors to explaining compositional turnover of terrestrial fauna in the Pilbara (3.12 points of overall 5.65 – Figure 5). Variance partitioning confirms that
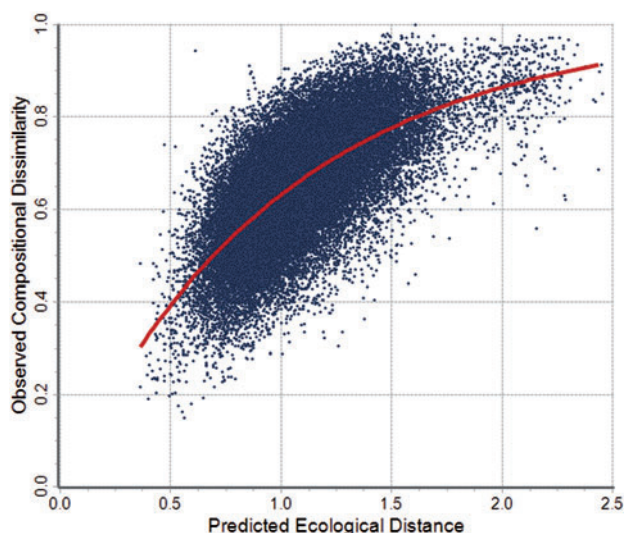

Figure 4   Overall fit of the final GDM to the terrestrial fauna data with the logit link function applied – observed v. linear predictor (predicted ecological distance).
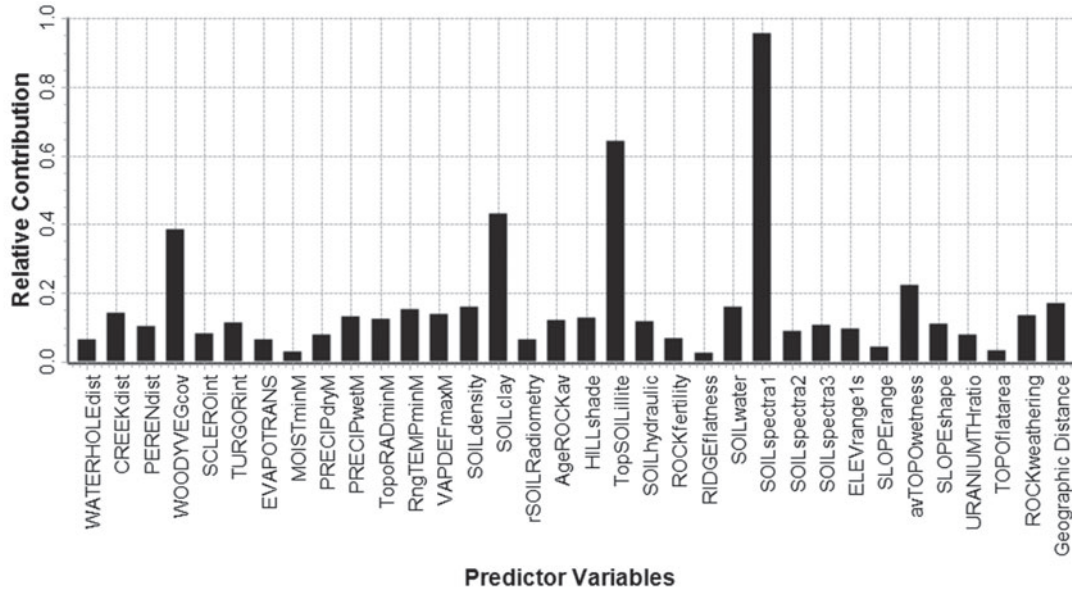
522

L.A. Gibson, K.J. Williams, A.M. Pinder, T.D. Harwood, N.L. McKenzie, S. Ferrier,
M.N. Lyons, A.H. Burbidge and G. Manion

Figure 5 Relative contribution (sum of all spline coefficient values) of each predictor in the final GDM of terrestrial fauna patterns. Summed across the coefficients for each variable, the overall relative contribution of predictors in each group is: regolith – 3.12, biotic and climate – 1.32, hydrologic and landform – 0.99, and geographic distance between site pairs – 0.17 (see group definitions in Appendix 1).
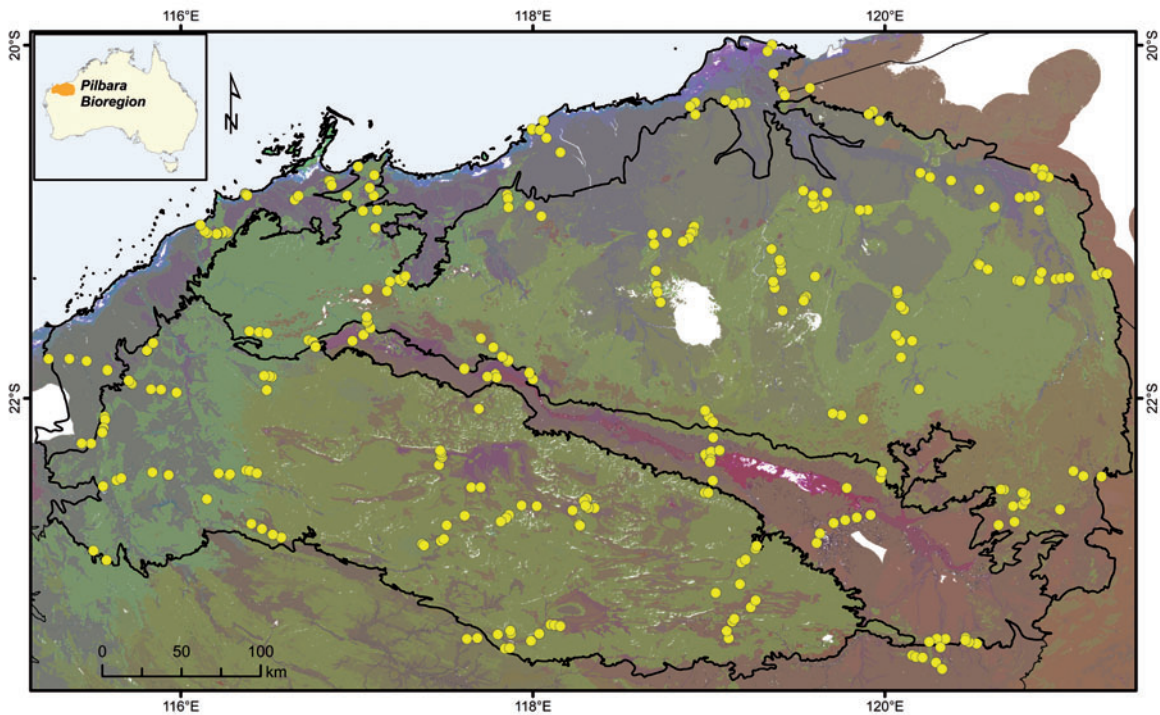

Figure 6 Visualisation of compositional turnover of terrestrial fauna species in the Pilbara, based on multi-dimensional scaling of predicted similarity, such that similar colours are predicted to have similar biological composition. Survey locations (yellow circles) and subregion boundaries (black lines) are shown. Note that white areas are extrapolation masks.
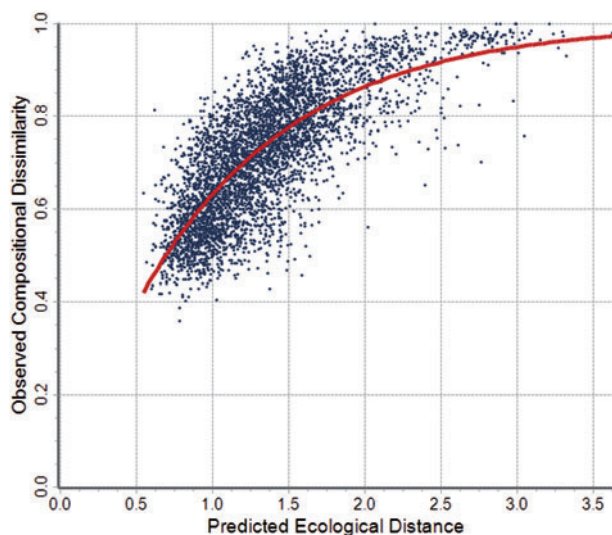
Figure 7 Overall fit of the final GDM to the wetland biodiversity data with the logit link function applied – observed v. linear predictor (predicted ecological distance).

regolith variables independently contribute the most to explaining patterns of compositional turnover in the Pilbara (15.1%) followed by landform/hydrologic variables (4.6%), then climatic/biotic variables (4.0%), and combinations with regolith largely contribute the remainder (Appendix 4A).

A classification of interpolated predicted compositional turnover of terrestrial fauna species provides a visual depiction of the fitted model using the full colour spectrum to show areas that are relatively similar or different (Figure 6). Areas that fall well outside the range of the data used in the fitted model have been excluded from this analysis.

Wetland biodiversity GDM

The final GDM model for wetland biodiversity explained 58.5% of the deviance with 31 environmental predictors retained in the model, and an intercept of 0.385 and sum of predictor spline coefficients of 7.85. Subsequent testing showed that geographic distance was not a significant predictor. The partial contribution to the percent deviance explained for each predictor variable is shown in Appendix 3.

The fit of the GDM model is shown in Figure 7, and the relative contribution of predictors in Figure 8. Again, regolith and landform appear to be more important predictors of compositional turnover of wetland biodiversity (5.61 points of overall 7.85) than regional climate or hydrologic variables (Figure 8). Variance partitioning also identified that a combination of regolith/soil and

landform/hydrologic variables contributes the most to compositional turnover in the Pilbara (12.1% independently and a further 34.5% in combination with other groups of variables) (Appendix 4B). The most important predictor is the third principal component of surficial soil spectra (SOILspectra3) representing soils with abundant amounts of smectite, a clay mineral which occurs primarily in low relief landscapes (Viscarra Rossel and Chen 2011). The first principal component is also important (SOILspectra1, interpreted above for the terrestrial model). The soil texture particle class (or stoniness of the soil) also ranks highly as an influential predictor (SOILcoarse), followed by the valley floor flatness landform predictor (VALLEYflatness) and surficial soil clay content (SOILclay). The minimum monthly diurnal temperature range is also important (RngTEMPminM) and only marginally more so than MODIS-derived evapotranspiration (EVAPOTRANS).

A classification of interpolated predicted compositional turnover within the wetland landscape provides a visual depiction of the fitted model using the full colour spectrum to show areas that are relatively similar or different (Figure 9). Areas that fall well outside the range of the data used in the fitted model, based on the extrapolation index, have been excluded from this analysis.

Representativeness of the conservation reserve network

The spatial representativeness of both the current reserve network in the Pilbara and the proposed future reserves (for 2015) for the terrestrial and wetland systems is shown in Figures 10a and b and 11a and b, respectively. The proportional representation of both terrestrial fauna and wetland biodiversity increased markedly when the proposed new reserves were included.

Biological survey coverage and spatial uncertainty

The spatial distribution of sampling representativeness (from comprehensive to sparse) associated with the terrestrial and wetland GDM models is shown in Figures 12 and 13, respectively. In both cases the core study area is well represented. Towards the southern edge, the wetland representation declines, probably due to a combination of reduced sample density and the larger area of similar habitat. Note that the wetland surveys are mostly concentrated on the waterways, resulting in a more variable level of representativeness. These results were developed as supporting information about a component of spatial uncertainty relevant to the interpretation of gaps in reserve representativeness.
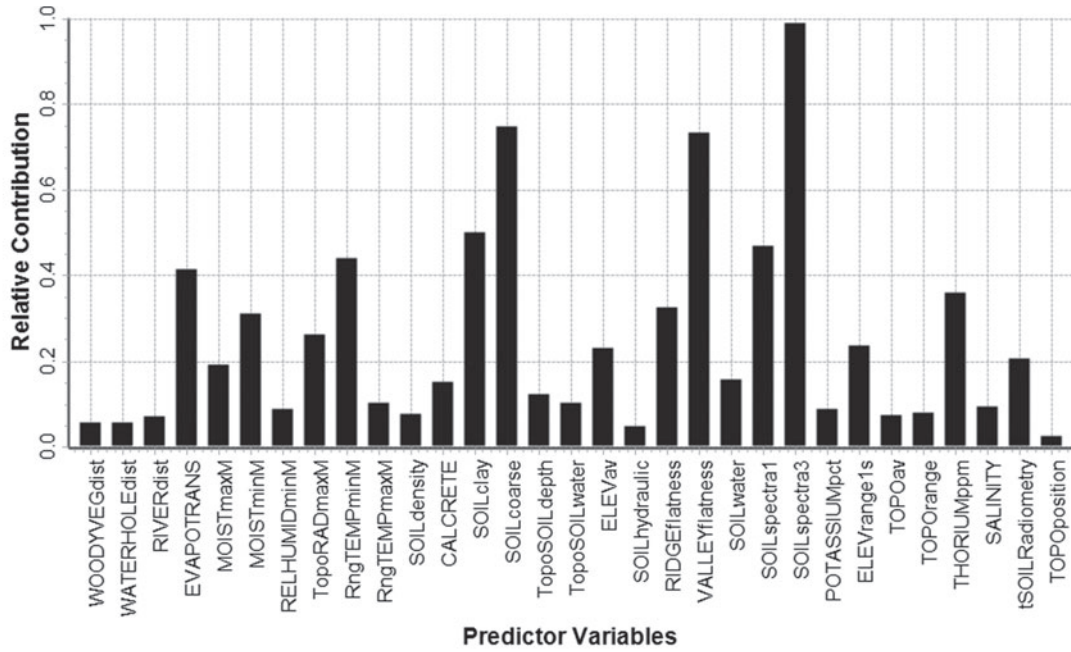
Figure 8   Relative contribution (sum of all spline coefficient values) of each predictor in the final GDM of wetland biodiversity patterns. Summed across the coefficients for each variable, the overall relative contribution of predictors in each group is: biotic and climate – 2.11, hydrologic and landform – 1.61, regolith – 4.13 (see group definitions in Appendix 1).
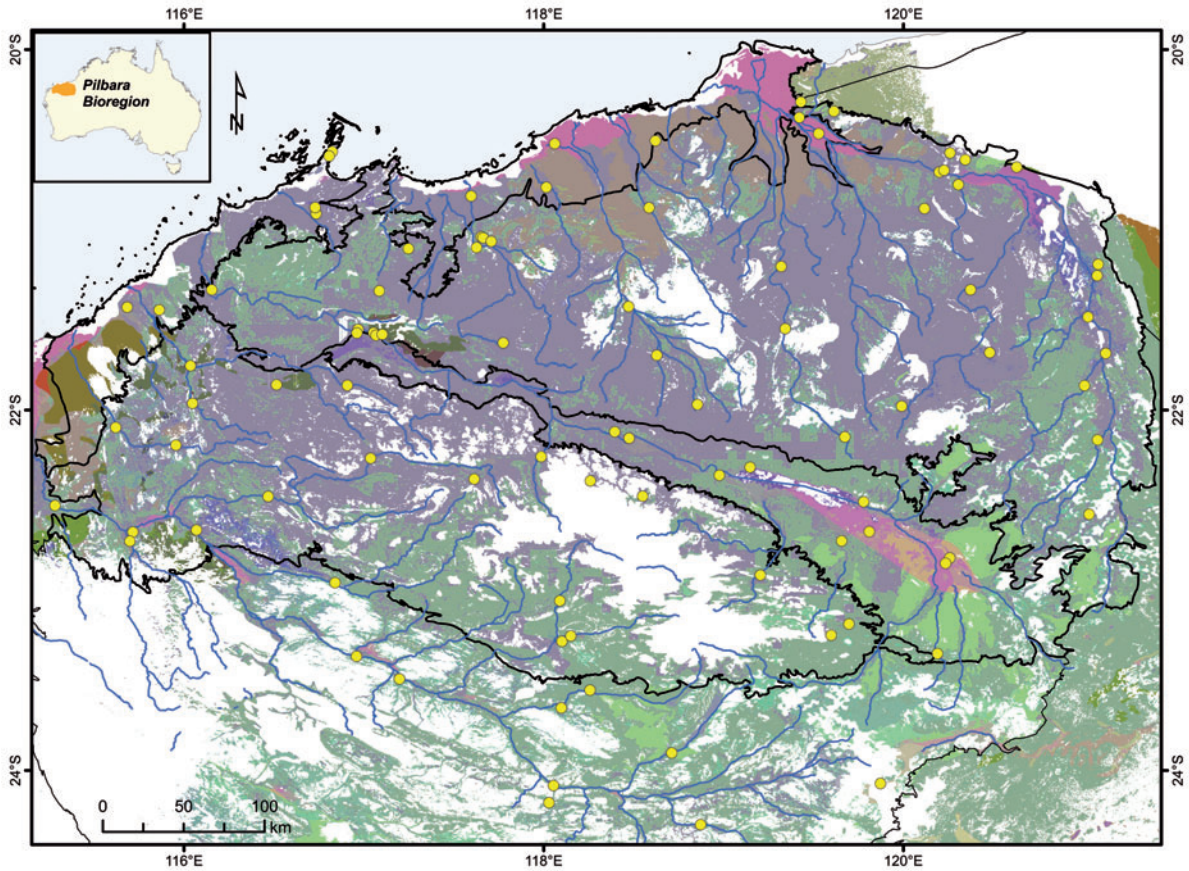

Figure 9   Visualisation of compositional turnover of wetland biodiversity in the Pilbara, based on multi-dimensional scaling of predicted similarity, such that similar colours are predicted to have similar biological composition. Survey locations (yellow circles), subregion boundaries (black lines) and major rivers (blue lines) are shown. Note that white areas are extrapolation masks.

DISCUSSION

Rapid advances in spatially explicit modelling techniques, in combination with increased accessibility of environmental data derived from high-resolution remote-sensed imagery (e.g. Guerschman *et al.* 2009) and predictive soil mapping (e.g. Viscarra Rossel and Chen 2011), have considerably improved the ability to assess biodiversity patterns across large geographic areas. In particular, community-level modelling approaches permit an appraisal of collective properties of biodiversity, such as changes in species composition (either spatially or temporally), which can incorporate large numbers of species, including those that are rare (Ferrier 2002; Ferrier *et al.* 2002, 2007; Arponen *et al.* 2008). The alternative, modelling at the species level, usually necessitates the exclusion of rare species, or those that are difficult to detect, due to insufficient data for statistical modelling (Ferrier *et al.* 2007). This means that a significant component of the biodiversity is under-represented in conservation assessments. Here, we have taken advantage of recent developments, specifically using generalised dissimilarity modelling (GDM), to relate changes in the composition (or turnover) of terrestrial fauna and wetland biodiversity to environmental gradients across the Pilbara biogeographic region. The output of the GDM models also enabled us to assess gaps in both the biological survey and the representativeness of the Pilbara's conservation reserve system.

Patterns in compositional turnover of the terrestrial fauna

Patterns in turnover of the terrestrial fauna were most strongly related to environmental variables associated with regolith, followed by landform/ hydrologic and then climate/biotic. This result is generally consistent with taxon-specific analyses of individual species relationships and multivariate analyses of the Pilbara Biodiversity Survey data (Gibson and McKenzie 2009; Burbidge *et al.* 2010; Durrant *et al.* 2010; Guthrie *et al.* 2010; Heterick *et al.* 2010; Pinder *et al.* 2010; Volschenk *et al.* 2010; Doughty *et al.* 2011). Among the taxonomic groups examined in these analyses, beetle, scorpion, bird, reptile and ground-mammal patterns all showed strong relationships with soil attributes. Similarly, three of the four most strongly supported variables in the GDM model of the aggregated data are all associated with surficial soil characteristics. Gollan *et al.* (2009) also identified soil parameters as important predictors in their study of invertebrate assemblages in the Pilbara, a relationship that has been shown in other arid areas as well (e.g. Bestelmeyer and Wiens 2001). Strong associations between the small ground-dwelling mammal and reptile assemblages, with substrate type (i.e. sand–clay–rockiness) have likewise been shown both within the Pilbara

(Gibson and McKenzie 2009, Doughty *et al.* 2011) and elsewhere (Woinarski *et al.* 1992; McKenzie *et al.* 2000b). Among the climate variables retained in the GDM model, almost all were related to a rainfall–soil moisture gradient, and hence productivity. Similarly the three biotic variables retained in the model, which included the variable that made the fourth largest contribution (woody vegetation cover), are likely to be surrogates for habitat productivity. Climate variables were weakly supported in the taxon-specific analyses, although precipitation variables were shown to marginally influence both ant and spider composition at sub-regional level (Durrant *et al.* 2010; Heterick *et al.* 2010). The landform variable that made the most contribution to the final model was average topographic wetness which is likely to influence habitat productivity and relate to high woody vegetation cover that occurs along riparian zones, foot slopes and other water-gaining environments in the Pilbara. Dispersal capacity also appears to play a role in the compositional patterns of the terrestrial fauna as indicated by geographic distance between pairs of sample locations being retained in the GDM model.

The classification of predicted compositional similarity of the terrestrial fauna shows that similarity in faunal composition largely conforms to bioregional boundaries. It also shows a change in composition from the area immediately adjacent to the Pilbara coast to areas inland, but little turnover in composition from the south-west to the north-east of the Pilbara, i.e. from the Hamersley subregion to the Chichester subregion. The Fortescue and Roebourne subregions are both compositionally distinct. This is most clearly evident within the Fortescue subregion with a distinct community corresponding to the Fortescue Marsh. This large, low-lying, ephemeral wetland, and associated alluvial plains, is situated between the substantially more rugged and rocky Chichester subregion to the north-east and the Hamersley subregion to the south-west. Compositional change from coastal to inland areas corresponds with a difference in soils and landforms as well as in regional climate, particularly the much higher humidity along the coastal strip, and is reflected by the variables retained in the GDM.

Patterns in compositional turnover of the wetland biodiversity

Turnover within the wetland biota was also strongly associated with regolith and landform, with seven climate, two hydrology and one biotic variable also included in the model. This result is consistent with the separate analyses of the major groups; aquatic invertebrates and riparian flora showed similar broad patterns and drivers. Pinder *et al.* (2010), on analysing patterns in the distribution of aquatic invertebrates in the
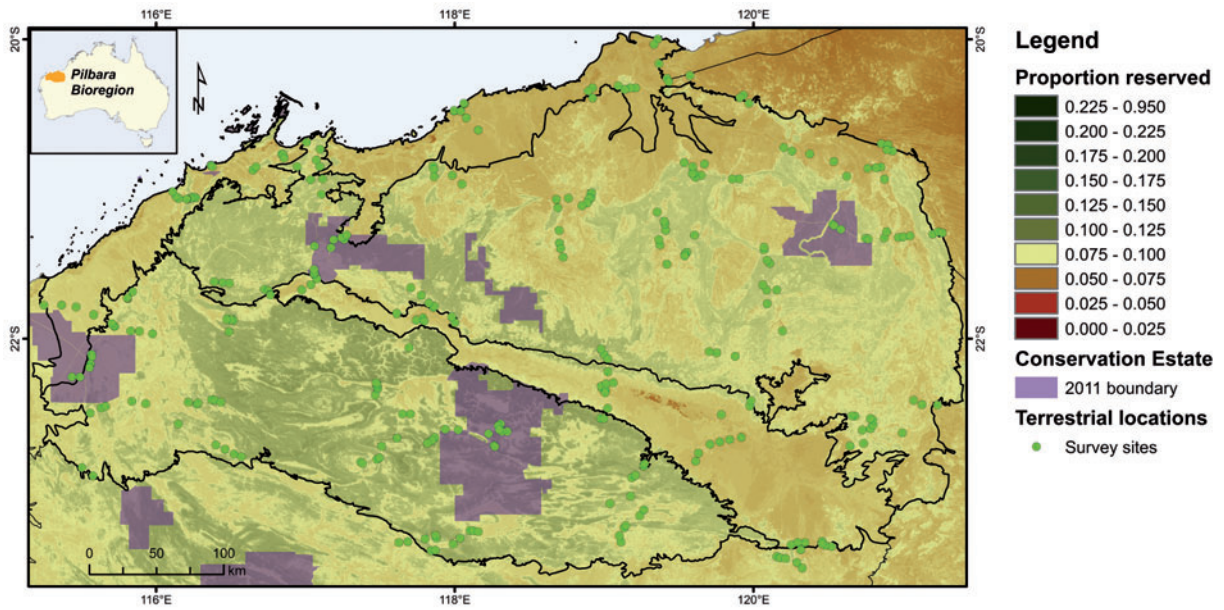
526

L.A. Gibson, K.J. Williams, A.M. Pinder, T.D. Harwood, N.L. McKenzie, S. Ferrier,
M.N. Lyons, A.H. Burbidge and G. Manion

Figure 10a Spatial distribution of the representation of terrestrial fauna biodiversity in conservation areas in the Pilbara biogeographic region, based on the 2011 conservation reserve estate (2011 boundary). Dark brown equals low proportional representation, through to dark green which equals high proportional representation. Note that the legend is shown in 10 classes for presentation but actual colouring and data are continuous.
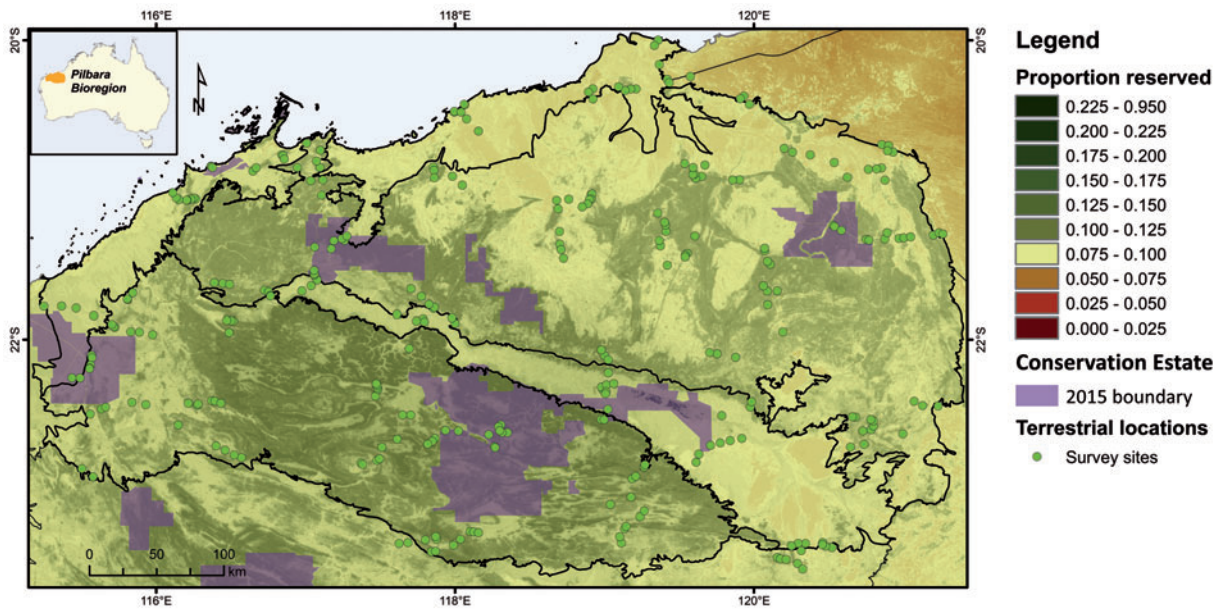

Figure 10b Spatial distribution of the representation of terrestrial fauna biodiversity in conservation areas in the Pilbara biogeographic region, based on the proposed conservation reserve estate for 2015 (2015 boundary). Dark brown equals low proportional representation, through to dark green which equals high proportional representation.

Pilbara, reported that variables describing hydrology (permanence and flow), turbidity, water chemistry (especially ionic composition and alkalinity) and sediments were important. Lyons (2015) also found that soil texture and hydrological setting were strongly associated with riparian flora composition. It was intended that the spatial variables selected for modelling would act as surrogates for these known local influences. In general terms, regolith, landform and climate are likely to influence the physical and chemical attributes of wetland (including riparian) habitats through their effects on hydrology, sedimentation and controls on geomorphology (Allan and Johnson 1997; Molnar *et al.* 2002).

Figure 11a Spatial distribution of the representation of wetland biodiversity in conservation areas in the Pilbara biogeographic region, based on the 2011 conservation reserve estate (2011 boundary). Dark brown equals low proportional representation, through to dark green which equals high proportional representation.
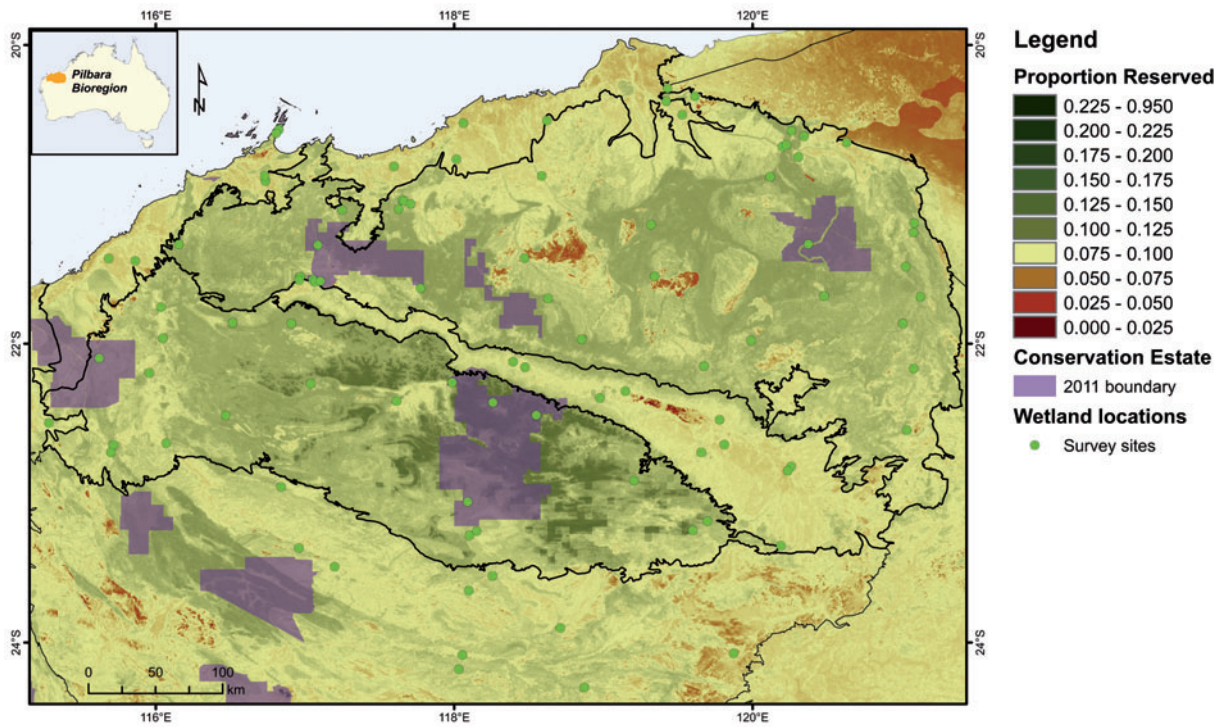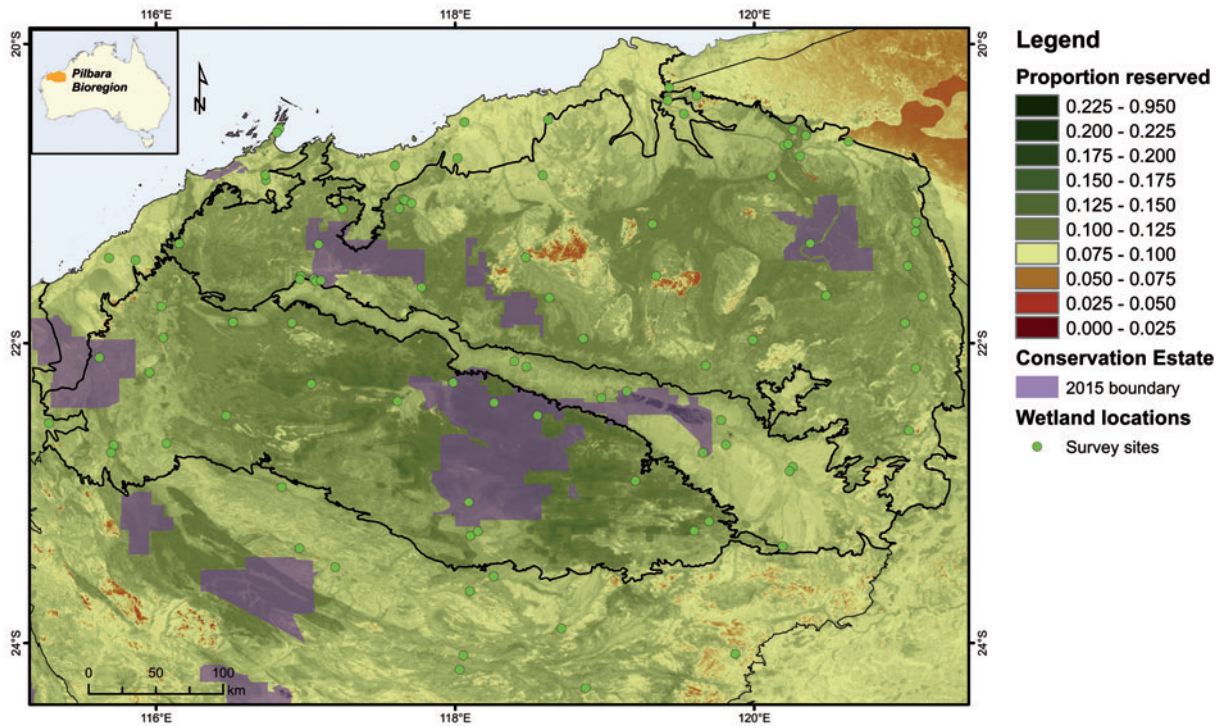

Figure 11b Spatial distribution of the representation of wetland biodiversity in conservation areas in the Pilbara biogeographic region, based on the proposed conservation reserve estate for 2015 (2015 boundary). Dark brown equals low proportional representation, through to dark green which equals high proportional representation.
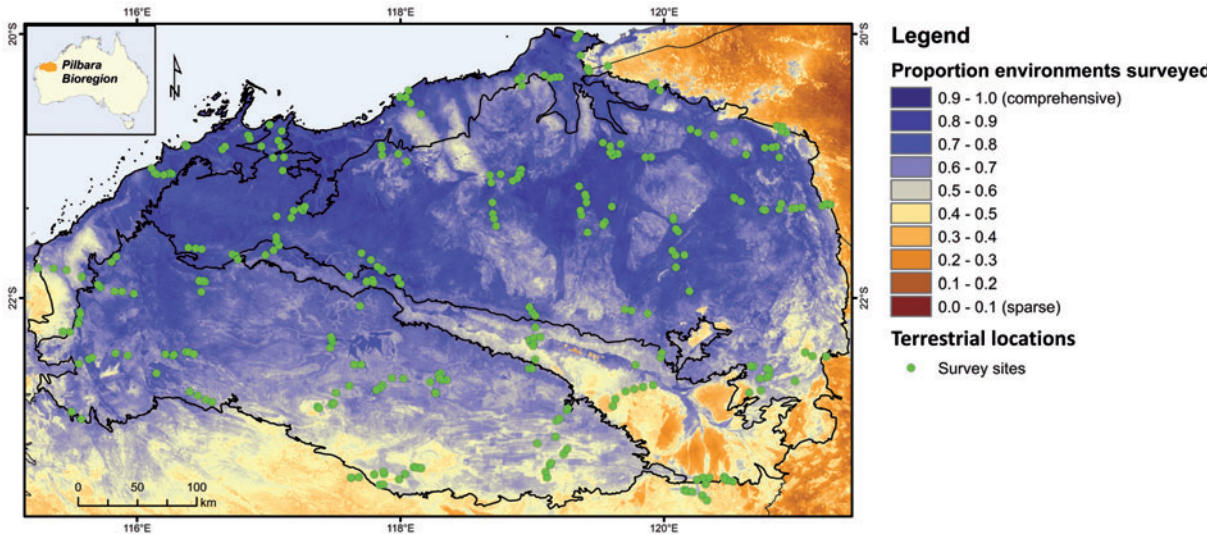
528

L.A. Gibson, K.J. Williams, A.M. Pinder, T.D. Harwood, N.L. McKenzie, S. Ferrier,
M.N. Lyons, A.H. Burbidge and G. Manion

Figure 12 Spatial distribution of the representativeness of the sample points for the terrestrial fauna GDM in the Pilbara biogeographic region. Dark blue areas are well sampled through to dark brown areas which are poorly sampled.
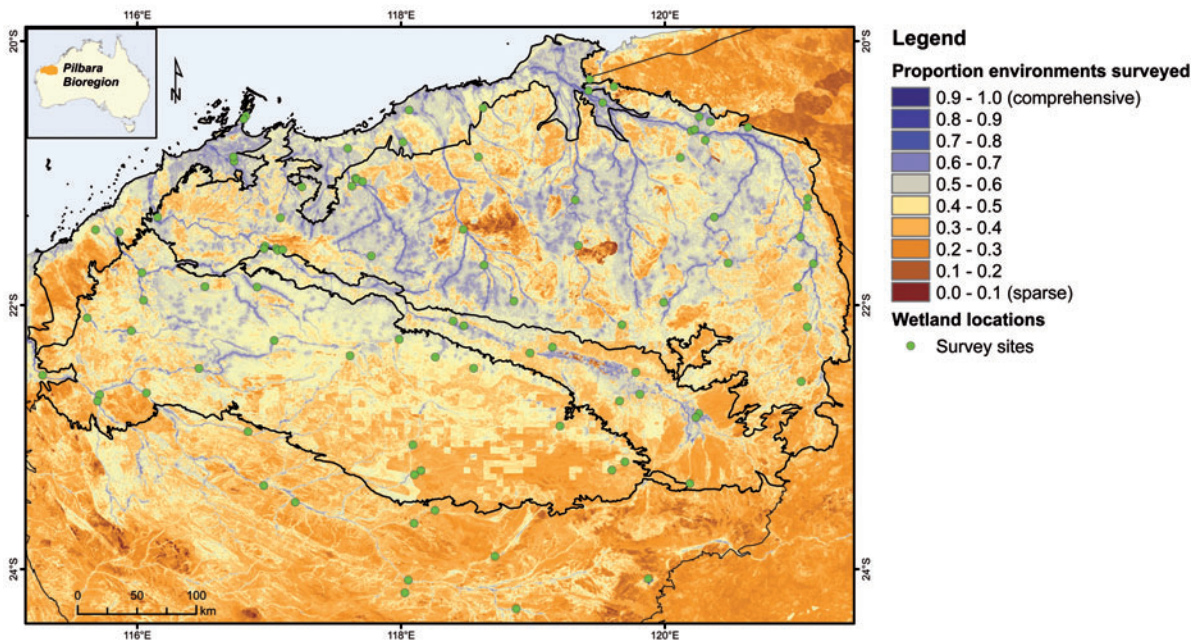

Figure 13 Spatial distribution of the representativeness of the sample points for the wetland biodiversity GDM in the Pilbara biogeographic region. Dark blue areas are well sampled through to dark brown areas which are poorly sampled.

Four soil attributes were included in the six most important predictors in the final GDM model. Two of these, one of which describes the stoniness of the soil and the other clay content, are likely to be associated with landscape influences on both turbidity and sedimentation of water bodies, and the composition of riparian soils. Two of the top-ranked variables in the model were soil spectra1 and 3, indicating weathered and depositional landscapes. Another soil

variable, CALCRETE, which influences alkalinity, is also likely to be a cipher for certain wetland habitats, including springs and spring-fed pools, and soil potassium (also included in the model) is likely to be an indicator of clay sediments. Pinder *et al.* (2010) found that dissolved potassium was a significant predictor of invertebrate composition.

The top-ranked landform variables in the final model describing ridge top and valley floor flatness,

in combination with other topographic landform variables, are likely to influence hydrology (including permanence and flow) through water runoff and retention patterns. The ranking of MODIS-derived evapotranspiration in the top seven most important variables is not surprising, given that the generation of this variable includes a term that implicitly represents precipitation interception losses and a remote-sensing index of surface moisture (Guerschman *et al.* 2009), and therefore a surrogate for water permanence. Other climate variables in the model, including measures of dry and wet season moisture, are also likely to reflect aspects of hydrological regime. Pinder *et al.* (2010) suggested that climate variables were acting as surrogates for landscape position, but the retention of both landform and climate variables in the wetland model suggests that these also have significant independent effects (as demonstrated by the variable partitioning results in Appendix 4B).

Salinity has been found to be a major driver of the composition of aquatic invertebrate communities (Pinder *et al*. 2005). Water table salinity was a minor variable in the model and may have been an indicator of surface water salinity at a few sites (although most sites were fresh) or this may have acted as a surrogate for less obvious aspects of landscape context (e.g. saline groundwater below the Fortescue Marsh). The minor effect of salinity in the GDM may also be attributed to its relatively coarse resolution, derived from national data compiled by the Bureau of Meteorology (2010).

Two of the six customised indices generated for the wetland biodiversity modelling were retained in the GDM (distance to drainage lines and to water points). Distance to drainage lines is probably a surrogate for channel v. off-channel wetlands (such as claypans which have a very different biota from river pools and springs). Distance to water bodies may indicate something about wetland size since it is mostly larger water bodies that are mapped. Unlike the terrestrial fauna, geographic distance between sites did not influence compositional turnover of this group, which is not surprising given that dispersal is partly determined by river network configuration and catchment topography (Morán-Ordóñez *et al.* in press) and that major floods are likely to have a homogenising effect within rivers (Thomaz *et al.* 2007).

The dominant spatial transition in wetland composition predicted by the GDM is south-east to north-west, separating the Hamersley Range, upper Fortescue River and upper Oakover River in the south and east (shades of green in Figure 9) from the Chichester and Roebourne Plains regions (brown/grey/purple). The centre of the saline and episodically flooded Fortescue Marsh, which lies near the boundary of these two zones, was largely excluded from Figure 9 due to under-sampling of this wetland type, but the model predicted that its fringes, which were sampled during the survey, support a biota unlikely to be replicated elsewhere in the region. The model also indicates that the broad alluvial landscapes south of, and upstream from, the marsh (blue/purple areas) support unique biotas within this southern area, but have biotas similar to floodplains associated with the Lower Oakover and De Grey rivers, parts of the eastern Roebourne Plains and the lower Ashburton south of Onslow (also purple in Figure 9).

The sharp change in predicted composition between the upper Fortescue valley (including the marsh) and the lower Fortescue reflects the very different wetland habitats above and below the Goodiarie Hills that isolate the upper and lower parts of this valley. The model also predicts that the Millstream area of the Fortescue River, with deep permanent pools and groundwater-fed springs, has a distinctive, unique biota and this is supported by Pinder *et al*. (2010) and Lyons (2015).

The model suggests that coastal parts of the study area have greater heterogeneity in composition than more inland areas. The model identified two small areas that may have distinct wetland biotas but which were not sampled. These are where the Hardey and Ashburton rivers meet to form a semi-confined floodplain with anabranching channels and areas upstream of the Oakover–Yilgalong confluence (both coloured rusty red in Figure 11). Finally, some unsampled areas east of the Pilbara (east of the lower De Grey and middle Oakover rivers) were predicted to be compositionally dissimilar to areas surveyed within the Pilbara, whereas there was little evidence of such transitions south and south-east of the Pilbara.

Variable selection strategy and model fit

While the increased availability of spatial environmental variables is advantageous, there is an associated problem of deciding which variables should be included in the model-building process. This is particularly a problem if the ratio of the number of environmental variables to sample size of the data set is large, as this may result in an over-fitted model, and reduce the accuracy of extrapolation into areas not sampled. Here, we followed the approach of Williams *et al.* (2010b, 2012) to first reduce the candidate variable set (by applying an ecological framework), and then apply a pruning procedure to remove correlated predictors that contributed little or nothing to the final fitted model – a trade-off between parsimony in the number of predictors included and reduced deviance explained following successive removal of predictors. We also explicitly mapped areas of extrapolation arising from the fitted model. Thus, of the 77 candidate predictors available, 34 were retained in the terrestrial model and 31 in the

wetland model. The deviance in turnover explained by these models (46.1% and 58.5% respectively) is comparable to, or greater than, those for various taxonomic groups modelled at the continental scale (15–52%) reported by Williams *et al.* (2010a). Other studies of compositional turnover report a deviance explained of 16.3% for freshwater fish, 18.5% for macroinvertebrates (Leathwick *et al.* 2011), 21.8% for Neotropical phytophagous arthropods (Bell *et al.* 2013), 40.9% for the ant genera of North America (Fitzpatrick *et al.* 2011) and 57% for land snails of New Zealand (Overton *et al.* 2009).

The level of unexplained variation in all these studies suggests that there are important factors influencing compositional patterns missing from the models. Such factors may relate to historical biogeography, competitive niche-differentiation and disturbance heterogeneity (Urban 2004; Armstrong 2005; Leibold *et al.* 2004, 2010), and other metacommunity processes that are difficult to capture with readily available spatial attributes (e.g. McKenzie and Bullen 2012). In addition, the poorer resolution of spatial predictors compared with actual processes influencing compositional turnover locally will limit predictive accuracy. The latter effect can be investigated where detailed physio-chemical analysis of site context data are collected along with the biotic observations. A GDM fitted using site context data and compared with spatial predictors may reveal how the two data sources provide complementary information. Geographic distance may be a reasonable surrogate for historical biogeography, and other missing or poorly represented factors, but our approach of using Euclidian distance between sample locations could be improved by deriving a measure using least cost paths which may account better for barriers to dispersal (e.g. Thomassen *et al.* 2011). Recent analyses by Morán-Ordóñez *et al.* (in press) suggest that the inclusion of landscape resistance distances informed by riverine and catchment topography may provide some extra power to the wetland model.

Additionally, as fire is a feature of Pilbara landscapes, and is likely to greatly influence patterns in compositional turnover of the terrestrial fauna (and the wetland biota to a degree), the inclusion of variables that estimate either directly or indirectly both the spatial and temporal patterns of fire, and fire intensity, is likely to improve the explanatory capacity of the GDM models. Patterns of vegetation coverage may provide some indication of the fire regime. As most of the Pilbara is under pastoral lease, a further disturbance factor likely to be unaccounted for in our models is grazing by domestic and feral herbivores. While we endeavoured to select survey sites in habitat patches that were the least disturbed, some level of disturbance was unavoidable at both terrestrial and wetland sites. Differential grazing pressure may be detected in indices of vegetation cover, and may partly explain the importance of woody vegetation cover as a predictor of compositional patterns in the terrestrial fauna.

Adequacy of the reserve system and gaps in the survey

Only about 8.5% of the Pilbara region is currently within conservation reserves, and most of this area falls within just five reserves: Karijini National Park, Millstream-Chichester National Park, Mungaroona Nature Reserve, Meentheena Conservation Park and Cane River Conservation Park (McKenzie *et al.* 2009) (in 2011, including former pastoral leases held by Parks and Wildlife). Consequently, even though these reserves are reasonably large (combined area is 1,345,000 ha), they are unlikely to capture the biodiversity of the Pilbara adequately. This lack of adequacy is supported by the spatial distribution of the representation of terrestrial fauna and wetland biodiversity in existing conservation reserves in the Pilbara (Figures 10a and 11a).

Figure 10a indicates that the terrestrial fauna diversity of the rugged Hamersley and south-western Chichester subregions is reasonably well represented in the reserve system, with the Pilbara's two large national parks (i.e. Karijini and Millstream-Chichester) and Mungaroona Nature Reserve, located in this area. While there is a conservation park (Meentheena) in the east of the Chichester subregion, there is a clear gap in the northern section of the subregion, extending from the coast to 100 km inland. With the exception of very small sections of Cane River Nature Reserve (in the far west of the Pilbara) and Millstream-Chichester National Park, there is currently no land reserved for conservation within the Roebourne subregion along the Pilbara coast. There are also no conservation reserves within the Fortescue River catchment of the Fortescue Plains subregion. While these subregions are small in area in comparison to the Chichester and Hamersley subregions (which combined cover about 80% of the region), the pattern of terrestrial fauna diversity in these areas appears to be poorly represented. With the addition of the proposed conservation reserves (Figure 10b), however, the proportional representation of terrestrial fauna increases considerably, even though the actual additional area of land proposed for reservation is small (ca. 4×10^5 ha). The improvement is the result of one new reserve within the Fortescue subregion, and another in the Roebourne subregion, although the gap in the northern Pilbara remains.

A similar pattern to that of the terrestrial fauna was observed for representation of wetland biodiversity within the conservation reserve system

of the Pilbara (Figure 11a). Again, based on the GDM predictions, the middle to upper Fortescue River valley (Fortescue subregion) and the coastal regions of the Pilbara (Roebourne subregion) are landscape types poorly represented within the 2011 conservation estate. Fortescue Marsh represents a unique landscape type not currently represented in the reserve system at all, as indicated by its red-brown colour in Figure 11a. The eastern parts of the Hamersley subregion are better represented than the western parts and the Chichester subregion, while the northern and central parts of the latter are the least represented of the upland areas. Including the proposed 2015 additions (Figure 11b) substantially improves the representation of wetland biodiversity, probably even more so than for the terrestrial fauna. Fortescue Marsh becomes very well represented and there is an improvement in representation for the Hamersley and Chichester subregions. Remaining gaps include wetlands within the Fortescue subregion (other than Fortescue Marsh), as well as wetlands in the Roebourne subregion and the northern slopes of the Chichester Range.

The predictive accuracy of patterns in compositional turnover depends on the coverage of sample points across the region of interest, and how well they represent the variation in environmental gradients. Here, for the terrestrial fauna, we observed good coverage of a large proportion of the study area, with the southern margin and south-eastern portions of the Pilbara the most poorly represented by the sample points. The inland extremity of the Fortescue subregion appears to be the largest gap in the biological survey, an area that coincides with a gap in the reserve system identified above. Similarly, gaps in survey coverage evident in the northern section of the Chichester subregion also reduce confidence in our observation that biodiversity in this area is poorly represented by the current reserve system. These areas are likely to benefit from further survey to verify the predictions. A comprehensive assessment of biological survey gaps using methods more specific to this evaluation (e.g. Funk *et al.* 2005) can be used to guide the placement of future surveys.

Given that most wetland habitats are within river channels and their associated floodplains, most catchments were well represented by the wetland biological survey locations. Exceptions might be the Ashburton River and its tributaries draining the southern slopes of the eastern Hamersley Range. Ephemeral, low-order streams were difficult to include in the survey due to their short, unpredictable hydroperiods. Such streams dominate in the highest upland areas shown as poorly represented by the sampled locations.

Limitations and future work

The Pilbara Biodiversity Survey compositional turnover models aimed to account for natural variation in biodiversity; however, some of the observed variation could be due to spatially varying disturbance regimes due to land use and management practices. In the absence of spatial data directly representing disturbance regimes (e.g. fire and herbivore grazing) we could not assess these effects in the model, nor use information about the condition of those areas in assessing the representativeness of the reserve system. To use this information effectively in biodiversity models and conservation assessments, further work is needed to distinguish natural cycles of disturbance from disturbances driven by human land use.

This analysis addresses the pattern of biological representativeness but does not address adequacy. For example, connectivity, meta-community viability and habitat condition are all important considerations in making decisions about which areas among available candidates are the best options for biodiversity conservation. These considerations, along with trade-offs related to other socioeconomic benefits, require more detailed and consultative assessments. The result of this analysis provides initial guidance as to where further biological survey may benefit decisions about the conservation of biodiversity in a network of reserves for the Pilbara. Analyses to specifically address biological survey gaps and systematic identification of new areas for survey can be developed using the approach described by Funk *et al.* (2005) combined with the outputs of GDM as proposed by Ferrier (2002), and demonstrated in Bell *et al.* (2014).

Consideration of other factors that are known to influence patterns in community composition, such as interspecific competition and historical biogeography, is likely to increase the proportion of variation explained by the GDM models. As discussed above, the inclusion of spatial environmental variables that better represent the compositional turnover of the groups considered here, such as finer resolution climatic data, and more proximal variables describing soil and hydrology processes, is also likely to increase the variation explained by the models. Moreover, our approach of aggregating the data, which includes a diverse range of taxonomic groups, may downplay the importance of variables influencing turnover specific to a taxonomic group.

A GDM model of the compositional turnover of vascular plants in the Pilbara will complement that of the terrestrial fauna and wetland biodiversity models and help to improve the assessment of gaps in the reserve system.

ACKNOWLEDGEMENTS

REFERENCES

Abbott, P.F. and Tabony, R.C. (1985). The estimation of humidity parameters. *Meteorological Magazine* **114**: 49–56.

Allan, J. and Johnson, L. (1997) Catchment-scale analysis of aquatic ecosystems. *Freshwater Biology* **37**: 107–111.

Allen, R., Pereira, L., Raes, D. and Smith, M. (1998). *Crop evapotranspiration – Guidelines for computing crop water requirements*. Food and Agriculture Organization of the United Nations, Rome.

Allnutt, T.F., Ferrier, S., Manion, G., Powell, G.V.N., Ricketts, T.H., Fisher, B.L., Harper, G.J., Irwin, M.E., Kremen, C., Labat, J-N., Lees, D.C., Pearce, T.A. and Rakotondrainibe, F. (2008). A method for quantifying biodiversity loss and its application to a 50-year record of deforestation across Madagascar. *Conservation Letters* **1**: 173–181.

Armstrong, D.P. (2005). Integrating the metapopulation and habitat paradigms for understanding broad-scale declines in species. *Conservation Biology* **19**: 1402–1410.

Arponen, A., Moilanen, A. and Ferrier, S. (2008). A successful community-level strategy for conservation prioritization. *Journal of Applied Ecology* **45**: 1436–1445.

Beard, J.S. (1990). *Plant life of Western Australia*. Kangaroo Press: Kenthurst, NSW.

Beckett, K.A. (2003). Airborne geophysics applied to groundwater modelling (pp. 8–10). *In*: I.C. Roach (ed.) *Advances in Regolith*. Cooperative Research Centre for Landscape Environments and Mineral Exploration.

Beckett, K.A. (2007). Multispectral analysis of high spatial resolution 256-channel radiometrics for soil and regolith mapping. PhD Thesis, Curtin University of Technology, Perth.

Belbin, L., Marshall, C. and Faith, D. (1983). Representing relationships by automatic assignment of color. *Australian Computer Journal* **15**: 160–163.

Bell, K.L., Heard, T.A., Manion, G., Ferrier, S. and van Klinken, R.D. (2013). The role of geography and environment in species turnover: phytophagous arthropods on a Neotropical legume. *Journal of Biogeography* **40**: 1755–1766.

Bell, K.L., Heard, T.A., Manion, G., Ferrier, S. and van Klinken, R.D. (2014). Characterising the phytophagous arthropod fauna of a single host plant species: assessing survey completeness at continental and local scales. *Biodiversity and Conservation* **23**: 2985–3003.

Berry, S., Mackey, B. and Brown, T. (2007). Potential applications of remotely sensed vegetation greenness to habitat analysis and the conservation of dispersive fauna. *Pacific Conservation Biology* **13**: 120–127.

Bestelmeyer, B.T and Wiens, J.A. (2001). Ant biodiversity in semiarid landscape mosaics: the consequences of grazing vs. natural heterogeneity. *Ecological Applications* **11**: 1123–1140.

Borcard, D., Legendre, P. and Drapeau, P. (1992). Partialling out the spatial component of ecological variation. *Ecology* **73**: 1045–1055.

Briggs, I.C. (1974). Machine contouring using minimum curvature. *Geophysics* **39**: 39–48.

Burbidge, A.H., Johnstone, R.E. and Pearson, D.J. (2010). Birds in a vast arid upland: avian biogeographical patterns in the Pilbara region of Western Australia. *Records of the Western Australian Museum*, Supplement **78**: 247–270.

Bureau of Meteorology (2010). *Australian hydrological geospatial fabric (geofabric) product guide, version 1.0*. Australian Government, Canberra.

Claridge, J., Williams, K.J. and Storey, R.J.L. (2000). *Creation of the south-east Queensland depth index rescaled using CTI*. Enhanced Resource Assessment 2000-05. A JVAP project QDN3A Technical Report. Queensland Department of Natural Resources, Brisbane.

DCCEE (2009). *Metadata: forest extent and change (v6). area-corrected, aggregated products*. Australian Government Department of Climate Change and Energy Efficiency (DCCEE), Canberra.

De Vries, R. (2009). *Australia's substrate fertility – version 0.8*. Department of the Environment, Water, Heritage and the Arts, Canberra.

Doughty, P., Rolfe, J.K., Burbidge, A.H., Pearson, D.J. and Kendrick, P.G. (2011). Herpetological assemblages of the Pilbara biogeographic region, Western Australia: ecological associations, biogeographic patterns and conservation. *Records of the Western Australian Museum*, Supplement **78**: 315–341.

Durrant, B.J., Harvey, M.S., Framenau, V.W., Ott, R. and

Waldock, J.M. (2010). Patterns in the composition of ground-dwelling spider communities in the Pilbara bioregion, Western Australia. *Records of the Western Australian Museum*, Supplement **78**: 185–204.

ESRI (2011). ArcGIS Desktop: Release 10. Environmental Systems Research Institute. Redlands, California, U.S.A.

Ferrier, S. (2002). Mapping spatial pattern in biodiversity for regional conservation planning: where to from here? *Systematic Biology* **51**: 331–363.

Ferrier, S., Drielsma, M., Manion, G., and Watson, G. (2002). Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. II. Community-level modelling. *Biodiversity and Conservation* **11**: 2309–2338.

Ferrier, S., Harwood, T. and Williams, K.J. (2010). *Using Generalised Dissimilarity Modelling to assess potential impacts of climate change on biodiversity composition in Australia, and on the representativeness of the National Reserve System.* A report prepared for the Department of Sustainability, Environment, Water, Population and Communities, Canberra. CSIRO Climate Adaptation Flagship, Canberra.

Ferrier, S., Manion, G., Elith, J. and Richardson, K. (2007). Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions* **13**: 252–264.

Ferrier, S., Powell, G.V.N., Richardson, K.S., Manion, G., Overton, J.M., Allnutt, T.F., Cameron, S.E., Mantle, K., Burgess, N.D., Faith, D.P., Lamoreux, J.F., Kier, G., Hijmans, R.J., Funk, V.A., Cassis, G.A., Fisher, B.L., Flemons, P., Lees, D., Lovett, J.C. and Van Rompaey, R.S.A.R. (2004). Mapping more of terrestrial biodiversity for global conservation assessment. *Bioscience* **54**: 1101–1109.

Fitzpatrick, M.C., Sanders, N.J., Ferrier, S., Longino J.T., Weiser, M.D. and Dunn, R. (2011). Forecasting the future of biodiversity: a test of single- and multi-species models for ants in North America. *Ecography* **34**: 836–847.

Funk, V.A., Richardson, K.S. and Ferrier, S. (2005). Survey-gap analysis in expeditionary research: where do we go from here? *Biological Journal of the Linnaean Society* **85**: 549–867.

Furby, S. (2002). *Land cover change: specification for remote sensing analysis.* Australian Greenhouse Office, Canberra.

Gallant, J. (2011). *Building the national one second digital elevation model of Australia.* Water Information Research and Development Alliance (WIRADA). Science Symposium 1–5 August 2011. CSIRO Water for a Healthy Country, Melbourne.

Gallant, J. and Austin, J. (2012). *Metadata: topographic position index derived from 1 second DEM-S.* CSIRO Land and Water, Canberra.

Gallant, J., Austin, J. and Dowling, T. (2011a). *Metadata: 300 m elevation range derived from 1 second DEM-S.* CSIRO Land and Water, Canberra.

Gallant, J., Austin, J. and Dowling, T. (2011b). *Metadata: contributing area – multiple flow direction (partial) derived from 1 second DEM-H.* CSIRO Land and Water, Canberra.

Gallant, J., Austin, J. and Dowling, T. (2011c). *Metadata: percent slope derived from 1 second DEM-S.* CSIRO Land and Water, Canberra.

Gallant, J., Austin, J. and Dowling, T. (2011d). *Metadata: plan curvature derived from 1 second DEM-S.* CSIRO Land and Water, Canberra.

Gallant, J., Austin, J. and Dowling, T. (2011e). *Metadata: profile curvature derived from 1 second DEM-S.* CSIRO Land and Water, Canberra.

Gallant, J., Austin, J. and Dowling, T. (2011f). *Metadata: slope relief derived from 1 second DEM-S.* CSIRO Land and Water, Canberra.

Gallant, J., Austin, J. and Dowling, T. (2011g). *Metadata: Topographic Wetness Index (TWI) derived from 1 second DEM-S.* CSIRO Land and Water, Canberra.

Gallant, J. and Read, A. (2009). Enhancing the SRTM data for Australia (pp. 149–154). *In*: R. Purves, S. Gruber, R. Straumann and T. Hengl (eds), *Proceedings of Geomorphometry, 31 August – 2 September 2009.* University of Zurich, Zurich.

Gallant, J.C. and Dowling, T.I. (2003). A multiresolution index of valley bottom flatness for mapping depositional areas. *Water Resources Research* **39**: 13 pp.

George, A.S., McKenzie, N.L. and Doughty, P. (eds) (2011). A biodiversity survey of the Pilbara region of Western Australia. *Records of the Western Australian Museum*, Supplement **78**: 1–311.

Geoscience Australia (2010). *The gadiometric map of Australia dataset.* Geoscience Australia, Australian Government, Canberra.

Geoscience Australia, CSIRO Land and Water (2010). *1 second SRTM derived digital elevation models user guide, version 1.0.* Geoscience Australia, Australian Government, Canberra.

Geoscience Australia, CSIRO Land and Water (2011). *1 second SRTM derived digital elevation models user guide: 1 second DSM, DEM, DEM-S and DEM-H, 3 second DSM, DEM and DEM-S, Version 1.0.4.* Geoscience Australia, Australian Government, Canberra.

Gibson, L.A. and McKenzie, N.L. (2009). Environmental associations of small ground-dwelling mammals in the Pilbara region, Western Australia. *Records of the Western Australian Museum*, Supplement **78**: 91–122.

Gollan, J.R., Ashcroft, M.B., Casis, G., Donnelly, A.P. and Lassau, S.A. (2009). Testing common habitat-based surrogates of invertebrate diversity in a semi-arid rangeland. *Biodiversity and Conservation* **18**: 1147–1159.

Gray, J., Bishop, T., Smith, P., Robinson, N. and Brough, D. (2012). *A pragmatic quantitative relationship for soil organic carbon distribution in eastern Australia.* The 5th Global workshop on Digital Soil Mapping 2012 – Digital soil assessments and beyond, 10–13 April 2012, Sydney, Australia.

Guerschman, J.P., Van Dijk, A.I.J.M., Mattersdorf, G., Beringer, J., Hutley, L.B., Leuning, R., Pipunic, R.C. and Sherman, B.S. (2009). Scaling of potential evapotranspiration with MODIS data reproduces flux observations and catchment water balance observations across Australia. *Journal of Hydrology* **369**: 107–119.

Guthrie, N.A., Weir, T. and Will, K. (2010). Localised

L.A. Gibson, K.J. Williams, A.M. Pinder, T.D. Harwood, N.L. McKenzie, S. Ferrier,
M.N. Lyons, A.H. Burbidge and G. Manion

534

and regional patterns in ground-dwelling beetle assemblages in a semi-tropical arid zone environment. *Records of the Western Australian Museum*, Supplement **78**: 169–184.

Heterick, B.E., Durrant, B. and Gunawardene, N.R. (2010). The ant fauna of the Pilbara Bioregion, Western Australia. *Records of the Western Australian Museum*, Supplement **78**: 157–167.

Hutchinson, M., Stein, J., Stein, J., Anderson, H. and Tickle, P. (2008). *GEODATA 9 Second DEM and D8 Digital Elevation Model Version 3 and Flow Direction Grid Gridded Elevation and Drainage Data Source Scale 1:250 000, User Guide Third Edition*, 3rd edition. Fenner School of Environment and Society, ANU and Geoscience Australia, Australian Government, Canberra.

IUCN (1994). *Guidelines for Protected Areas Management Categories.* IUCN (The World Conservation Union), Cambridge, UK and Gland, Switzerland.

Jacquier, D. (2011a). *Metadata: ASRIS 0-1m Plant Available Water Capacity (250m raster)*. CSIRO Land and Water, Canberra.

Jacquier, D. (2011b). *Metadata: ASRIS 0-30cm Bulk Density (250m raster)*. CSIRO Land and Water, Canberra.

Jacquier, D. (2011c). *Metadata: ASRIS 0-30cm Clay Content (250m raster)*. CSIRO Land and Water, Canberra.

Jacquier, D. (2011d). *Metadata: ASRIS Australian Soil Classification – Dominant Soil Order (250m raster)*. CSIRO Land and Water, Canberra.

Jones, M.M., Ferrier, S., Condit, R., Manion, G., Aguilar, S., and Pérez, R. (2013). Strong congruence in tree and fern community turnover in response to soils and climate in central Panama. *Journal of Ecology* **101**: 506–516.

Laurie, J., Mantle, D. and Nicoll, R.S. (2008). Customising the geological timescale, pp. 3. *In: AusGeo News.* Geoscience Australia, Australian Government, Canberra.

Leathwick, J.R., Snelder, T., Chadderton, W.L., Elith, J., Julian, K. and Ferrier, S. (2011). Use of generalised dissimilarity modelling to improve the biological discrimination of river and stream classifications. *Freshwater Biology* **56**: 21–38.

Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F., Kolt, R.D., Shurin, J.B., Law, R., Tilman, D., Loreau, M. and Gonzalez, A. (2004). The metacommunity concept: a framework for multi-scale community ecology. *Ecological Letters* **7**: 601–613.

Leibold, M.A., Economo, E.P. and Peres-Neto, P. (2010). Metacommunity phylogenetics: separating the roles of environmental filters and historical biogeography. *Ecological Letters* **13**: 1290–1299.

Leighton, K. A. (2004). Climate (pp 19–38). *In:* Van Vreeswyk, A. M. E., Payne, A. L., Leighton, K. A. and Hennig, P. (eds), *An inventory and condition survey of the Pilbara region, Western Australia.* Technical Bulletin No. 92. Western Australian Department of Agriculture, Perth, Australia.

Liu, S.F., Raymond, O.L., Stewart, A.J., Sweet, I.P., Duggan, M.B., Charlick, C., Phillips, D. and Retter, A.J. (2006). *Surface geology of Australia 1:1,000,000 scale, Northern Territory [Digital Dataset].* The Commonwealth of Australia, Geoscience Australia (http://www.ga.gov.au), Canberra.

Lyons, M.N. (2015). The riparian flora and plant communities of the Pilbara region of Western Australia. *Records of the Western Australian Museum*, Supplement **78**: 485–513.

Mackey, B.G., Berry, S., Hugh, S., Ferrier, S., Harwood, T. and Williams, K. (2012). Ecosystem greenspots: identifying potential drought, fire and climate change micro-refuges. *Ecological Applications* **22**: 1852–1864.

Manion, G. (2014). *.NET Generalised Dissimilarity Modeller (GDM) version 3.10 (development, 22/09/2014) software.* NSW Office of Environment and Heritage, Armidale.

Manion, G., Lisk, M., Ferrier, S., Nieto-Lugilde, D., Fitzpatrick, M.C. (2015). *Package 'gdm' – Functions for Generalized Dissimilarity Modelling, version 1.0.* (http://cran.r-project.org/web/packages/gdm/index.html)

McBratney, A.B., Santos, M.L.M. and Minasny, B. (2003). On digital soil mapping. *Geoderma* **117**: 3–52.

McKenzie, N.J. and Hook, J. (1992). *Interpretations of the Atlas of Australian Soils*. CSIRO Division of Soils, Canberra.

McKenzie, N.J., Jacquier, D.W., Ashton, L.J. and Cresswell, H.P. (2000a). *Estimation of Soil Properties Using the Atlas of Australian Soils*. CSIRO Land and Water, Canberra.

McKenzie, N.J., Jacquier, D.W., Maschmedt, D.J., Griffin, E.A. and Brough, D.M. (2005). *The Australian Soil Resource Information System Technical specifications.* National Committee on Soil and Terrain Information/ Australian Collaborative Land Evaluation Program (www.asris.csiro.au), Canberra.

McKenzie, N.L., Keighery, G.J., Gibson, N. and Rolfe, J.K. (2000b). Patterns in the biodiversity of terrestrial environments in the southern Carnarvon Basin, Western Australia. *Records of the Western Australian Museum*, Supplement **61**: 511–546.

McKenzie, N.L., Gibson, N., Keighery, G.J. and Rolfe, J.K. (2004). Patterns in the biodiversity of terrestrial environments in the Western Australian wheatbelt. *Records of the Western Australian Museum*, Supplement **67**: 293–335.

McKenzie, N.L., van Leeuwen S. and Pinder, A.M. (2009). Introduction to the Pilbara Biodiversity Survey, 2002–2007. *Records of the Western Australian Museum*, Supplement **78**: 3–89.

McKenzie, N.L. and Bullen, R.B. (2012). An acoustic survey of zoophagic bats in the Kimberley, Western Australia, including data on the echolocation ecology, organisation and habitat relationships of regional communities. *Records of the Western Australian Museum, Supplement* **81**: 67–108.

Micallef, L. and Rodgers, P. (2014). *eulerAPE/*: Drawing area-proportional 3-Venn diagrams using ellipses. *PLoS ONE* **9**(7), e101717.

Minty, B., Franklin, R., Milligan, P., Richardson, L.M. and Wilford, J. (2009). The radiometric map of Australia. *Exploration Geophysics* **40**: 325–333.

Mokany, K., Jones, M.M. and Harwood, T.D. (2013).

Scaling pairwise β-diversity and α-diversity with area. *Journal of Biogeography* **40**: 2299–2309.

Molnar, P., Burlando, P. and Ruf, W. (2002). Integrated catchment assessment of riverine landscape dynamics. *Aquatic Sciences* **64**, 129–140.

Morán-Ordóñez, A., Pavlova, A., Pinder, A., Sim, L., Sunnucks, P., Thompson, R. and Davis, J. (in press). Aquatic communities in arid landscapes: local conditions, dispersal-traits and landscape configuration determine local biodiversity. *Diversity and Distributions*. DOI: 10.1111/ddi.12342.

Overton, J.M., Barker, G.M. and Price, R. (2009). Estimating and conserving patterns of invertebrate diversity: a test case of New Zealand land snails. *Diversity and Distributions* **15**: 731–741.

Percival, P.J. (2010). *Index of airborne geophysical surveys (eleventh edition).* Geoscience Australia, Record 2010/13, Canberra.

Pinder, A., Halse, S.A., McRae, J.M. and Shiel, R.J. (2005). Occurrence of aquatic invertebrates of the wheatbelt region of Western Australia in relation to salinity. *Hydrobiologia* **543**: 1–24.

Pinder, A.M., Halse, S.A., Shiel, R.J. and McRae, J.M. (2010). An arid zone awash with diversity: patterns in the distribution of aquatic invertebrates in the Pilbara region of Western Australia. *Records of the Western Australian Museum*, Supplement **78**: 205–246.

Prober, S.M., Hilbert, D.W., Ferrier, S., Dunlop, M. and Gobbett, D. (2012). Combining community-level spatial modelling and expert knowledge to inform climate adaptation in temperate grassy eucalypt woodlands and related grasslands. *Biodiversity and Conservation* **21**: 1627–1650.

Raymond, O.L., Liu, S.F. and Kilgour, P. (2007a). *Surface geology of Australia 1:1,000,000 scale, Tasmania – 3rd edition [Digital Dataset].* Geoscience Australia (http://www.ga.gov.au), Canberra.

Raymond, O.L., Liu, S.F., Kilgour, P., Retter, A.J. and Connolly, D.P. (2007b). *Surface geology of Australia 1:1,000,000 scale, Victoria – 3rd edition [Digital Dataset].* Geoscience Australia (http://www.ga.gov.au), Canberra.

Raymond, O.L., Liu, S.F., Kilgour, P., Retter, A.J., Stewart, A.J. and Stewart, G. (2007c). *Surface geology of Australia 1:1,000,000 scale, New South Wales – 2nd edition [digital dataset].* Geoscience Australia (http://www.ga.gov.au), Canberra.

Sneath, P.H.A. and Sokal, R.R. (1973). *Numerical taxonomy: the principles and practice of numerical classification.* Freeman: San Francisco, U.S.A.

Sørensen, T.A. (1948). A method of establishing groups of equal amplitude in plant sociology based on similarity of species content, and its application to analyses of the vegetation on Danish commons. *Kongelige Danske Videnskabernes Selskabs Biologiske Skrifter* **5**: 1–34.

Speight, J.G. (2009). Landform. *In*: National Committee on Soil and Terrains (ed.), *Australian soil and land survey field handbook* (3rd edn*).* CSIRO Publishing, Melbourne.

Stewart, A.J., Sweet, I.P., Needham, R.S., Raymond, O.L., Whitaker, A.J., Liu, S.F., Phillips, D., Retter, A.J., Connolly, D.P. and Stewart, G. (2008). *Surface geology of Australia 1:1,000,000 scale, Western Australia [Digital Dataset].* The Commonwealth of Australia, Geoscience Australia (http://www.ga.gov.au), Canberra.

Thomassen, H.A., Buermann, W., Milá, B., Graham, C.H., Cameron, S.E., Schneider, C.J., Pollinger, J.P., Saatchi, S., Wayne, R.K. and Smith, T.B. (2010). Modelling environmentally associated morphological and genetic variation in a rainforest bird, and its application to conservation prioritization. *Evolutionary Applications* **3**: 1–16.

Thomassen, H.A., Fuller, T., Buermann, W., Mila, B., Kieswetter, C.M., Cameron, S.E., Mason, E., Schweizer, R., Schlunegger, J. and Chan, J. (2011). Mapping evolutionary process: a multi-taxa approach to conservation prioritization. *Evolutionary Applications* **4**: 397–413.

Thomaz, S., Bini, L. and Bozelli, R.L. (2007). Floods increase similarity among aquatic habitats in river-floodplain systems. *Hydrobiologia* **579**: 1–13.

Urban, M.C. (2004). Disturbance heterogeneity determines freshwater metacommunity structure. *Ecology* **85**: 2971–2978.

Viscarra Rossel, R.A. (2011). Fine-resolution multiscale mapping of clay minerals in Australian soils measured with near infrared spectra. *Journal of Geophysical Research* **116**: F04023.

Viscarra Rossel, R.A., Bui, E.N., de Caritat, P. and McKenzie, N.J. (2010a). Mapping iron oxides and the color of Australian soil using visible-near-infrared reflectance spectra. *Journal of Geophysical Research* **115**: F04031.

Viscarra Rossel, R.A., Chappell, A., Caritat, P. and McKenzie, N.J. (2010b). Mapping the information content of Australian visible-near infrared soil spectra. *19th World Congress of Soil Science, Soil Solutions for a Changing World. International Union of Soil Scientists*: 1–6 August 2010, Brisbane, Australia.

Viscarra Rossel, R.A. and Chen, C. (2011). Digitally mapping the information content of visible-near infrared spectra of surficial Australian soils. *Remote Sensing of Environment* **115**: 1443–1455.

Volschenk, E.S., Burbidge, A.H., Durrant, B.J. and Harvey, M.S. (2010). Spatial distribution patterns of scorpions (Scorpiones) in the arid Pilbara region of Western Australia. *Records of the Western Australian Museum*, Supplement **78**: 271–284.

Western, A. and McKenzie, N. (2004). *Soil hydrological properties of Australia Version 1.0.1.* CRC for Catchment Hydrology, Melbourne.

Whitaker, A.J., Champion, D.C., Sweet, I.P., Kilgour, P. and Connolly, D.P. (2007). *Surface geology of Australia 1:1,000,000 scale, Queensland – 2nd edition [digital dataset].* The Commonwealth of Australia, Geoscience Australia (http://www.ga.gov.au), Canberra.

Whitaker, A.J., Glanville, H.D., English, P.M., Stewart, A.J., Retter, A.J., Connolly, D.P., Stewart, G.A.

and Fisher, C.L. (2008). *Surface geology of Australia 1:1,000,000 scale, South Australia – 1st edition [digital dataset].* The Commonwealth of Australia, Geoscience Australia (http://www.ga.gov.au), Canberra.

Wilford, J. (2012). A weathering intensity index for the Australian continent using airborne gamma-ray spectrometry and digital terrain analysis. *Geoderma* **183**: 124–142.

Williams, K.J., Belbin, L., Austin, M.P., Stein, J. and Ferrier, S. (2012). Which environmental variables should I use in my biodiversity model? *International Journal of Geographic Information Sciences* **26**: 2009–2047.

Williams, K.J., Ferrier, S., Rosauer, D., Yeates, D., Manion, G., Harwood, T., Stein, J., Faith, D.P., Laity, T. and Whalen, A. (2010a). *Harnessing continent-wide biodiversity datasets for prioritising national conservation investment.* A report prepared for the Department of Sustainability, Environment, Water, Population and Communities, Australian Government, Canberra, by CSIRO Ecosystem Sciences, Canberra.

Williams, K.J., Ferrier. S., Rosauer, D., Yeates, D., Manion, G., Harwood, T., Stein, J., Faith, D.P., Laity, T. and Whalen, A. (2010b). *Harnessing continent-wide biodiversity datasets for prioritising national conservation investment: appendices.* A report prepared for the Department of Sustainability, Environment, Water, Population and Communities, Australian Government, Canberra, by CSIRO Ecosystem Sciences, Canberra.

Wilson, J. and Gallant, J. (2000). Digital terrain analysis (pp. 1–27). *In*: J.P. Wilson and J.C. Gallant (eds), *Terrain Analysis: Principles and Applications*. John Wiley, New York.

Woinarski, J.C.Z., Braithwaite, R.W., Menkhorst, K.A., Griffin, S., Fisher, A. and Preece, N. (1992). Gradient analysis of the distribution of mammals in Stage III of Kakadu National Park, with a review of the distribution patterns of mammals across north-western Australia. *Wildlife Research* **19**: 233–262.

Xu, T. and Hutchinson, M. (2011). *ANUCLIM version 6.1 user guide.* The Australian National University, Fenner School of Environment and Society, Canberra.

APPENDIX 1

Environmental data compiled for GDM analysis of Pilbara Biodiversity Survey data.

For the Pilbara region analysis, we targeted the collation of spatial environmental data describing facets of the climate, regolith, landform, hydrology and biological habitats, or their proxies (see Appendix 1A). Ordinal or continuous variables with consistent spatial dimensions are required for GDM. Minor differences in spatial extent and data/ no-data areas were filled using focal-mean or focal-majority algorithms for continuous and ordinal variables respectively, using ArcGIS software (ESRI 2011).

Climate

Long-term (30 years, 1975–2005) monthly variation in climate was derived from ANUCLIM v6.1 (Xu and Hutchinson 2011) using version 3 of the 9 second (9sec) digital elevation model (DEM) for Australia (Hutchinson *et al.* 2008). Indices of annual and seasonal variation in a range of climatic variables were subsequently derived. Vapour pressure deficit was derived using monthly dew point temperature, and wet and dry bulb temperatures based on the equations outlined in Allen *et al.* (1998). The psychrometric data method was used to estimate actual vapour pressure from gridded values of atmospheric pressure based on altitude from the DEM. For relative humidity we adopted the equation used by the Australian Bureau of Meteorology (Abbott and Tabony 1985).

These climatic variables characterise general patterns of seasonal wetting and drying but lack detail about inter-annual variability. This variability, combined with the sparse distribution of weather stations throughout the Pilbara, limits the accuracy of the climatic signal and its correlation with patterns of regional variation in the biota. As such, two indirect location-based predictors were included as proxies for unexplained variation correlated with climate – Euclidean distance from coast and elevation. The elevation data were sourced from the 3 second (3sec) DEM for Australia which is a derivative of the 1 second (1sec) DEM (Gallant and Read 2009; Gallant 2011; Geoscience Australia and CSIRO Land and Water 2011). The elevation mean and range within each 9sec grid were compiled as candidates for explaining regional to local elevation heterogeneity.

We also included a remotely sensed variable of evapotranspiration derived from MODIS-derived potential evapotranspiration (Guerschman *et al.* 2009) to characterise variation due to the interaction among climate, landform and vegetation patterns.

Regolith

Because spatial data describing soils is limited in resolution, a wide range of candidate variables that are potentially correlated with patterns in soil variability at the site level were compiled. These variables fall into four classes: 1) attributes derived from soil maps; 2) models of soil properties; 3) attributes derived from geology maps; and 4) geophysics variables such as radiometric data.

Two sources of attributes from soil maps were available; one set was based on an interpretation of the *Atlas of Australian Soils* (McKenzie and Hook 1992; McKenzie *et al.* 2000a; Western and McKenzie 2004), and another set based on a composite of best available soil maps (McKenzie *et al.* 2005; Jacquier 2011a, b, c, d), available through the Australian Soil Resource Information System (http://www.asris. csiro.au/themes/NationalGrids.html). Where these attributes represent the same property, the more recent set based on the best composite of maps was used, even though map boundaries may influence the spatial appearance of the prediction.

Variation in surface geology for Australia was compiled by Geoscience Australia from 1:250,000 source data (Liu *et al.* 2006; Raymond *et al.* 2007a, b, c; Whitaker *et al.* 2007, 2008; Stewart *et al.* 2008). However, only two attributes are currently available in continuous or ordinal form – inherent rock fertility (De Vries 2009) and geological age based on the timescales database (Laurie *et al.* 2008) compiled by Williams *et al.* (2010a, b) at 0.01 resolution.

Digital soil mapping is a rapidly developing field of science (McBratney *et al.* 2003) and a number of modelled soil properties are becoming available nationally with potential application in ecology (e.g. Viscarra Rossel *et al.* 2010a, b; Viscarra Rossel and Chen 2011; Viscarra Rossel 2011; Gray *et al.* 2012; Wilford 2012). We explored the potential predictive utility of several of these modelled soil properties for the first time – weathering intensity index (Wilford 2012); three principal components of soil colours (Viscarra Rossel *et al.* 2010b) and three clay minerals (Viscarra Rossel 2011) at two soil depth profiles – 0 to 20 cm and 60 to 80 cm.

Radiometric or gamma-ray spectrometry data is another potential source of information about the structure and composition of the top 30 to 40 cm of the land surface. All rocks and soils contain radioactive isotopes, and almost all the gamma-rays detected near the Earth's surface are the result of the natural radioactive decay of potassium, uranium and thorium. Changes in lithology, or soil type, are often accompanied by changes in

538

L.A. Gibson, K.J. Williams, A.M. Pinder, T.D. Harwood, N.L. McKenzie, S. Ferrier,
M.N. Lyons, A.H. Burbidge and G. Manion

the concentrations of radio-elements and may also indicate mineral deposits. Potassium, uranium and thorium behave quite differently from one another during weathering and pedogenesis and in combination with digital terrain data can be used to define regolith properties such as weathering intensity (Wilford 2012). Quantitative soil characterisation based on radiometric data is an area of continuing research (Beckett 2003, 2007). We obtained the 2010 edition of the radiometric map of Australia dataset (Geoscience Australia 2010). This dataset comprises grids of potassium (K), uranium (U) and thorium (Th) element concentrations, and derivatives of these grids, that were derived by seamlessly merging over 550 airborne gamma-ray spectrometric surveys in the national radioelement database (Percival 2010) using the method described by Minty *et al.* (2009). The original survey grids were levelled and then re-sampled, using minimum curvature (Briggs 1974), onto the Radiometric Map of Australia Grids with a cell size of about 100 m (0.001 degrees).

Landform

Variation in topography and landform can be captured by indices derived from digital terrain models (Wilson and Gallant 2000). Gallant and co-workers have been generating terrain indices based on the 1sec DEM for Australia (Geoscience Australia and CSIRO Land and Water 2010). These indices vary in complexity from simple, such as slope (Gallant *et al.* 2011c), relief (Gallant *et al.* 2011f) and elevation diversity (Gallant *et al.* 2011a), to more complex algorithms such as topographic wetness indices (Gallant *et al.* 2011g) based on contributing area (Gallant *et al.* 2011b) and landform shape (Gallant *et al.* 2011d; Gallant *et al.* 2011e). Classes of topographic position commonly observed in the field (Speight 2009) also can be interpreted from a

DEM (Gallant and Austin 2012) or derived using multi-resolution methods (Gallant and Dowling 2003). We compiled a few of the available indices describing landform diversity from existing datasets derived from different resolution DEMs, all of which are consistent with our 9sec analysis resolution.

Vegetation

Indices of vegetation greenness or fPAR (fraction of photosynthetically active radiation intercepted by the sunlit canopy) based on time series of normalised-difference vegetation index (NDVI) values from the NASA MODIS 16-day L3 Global 250 m (MOD13Q1) satellite imagery were available for use in this study (Berry *et al.* 2007; Mackey *et al.* 2012).

Distance to woody vegetation and the extent of woody vegetation detected in a 250 m grid cell were derived from the forest extent and change area corrected aggregate products (V6) developed by the Department of Climate Change and Energy Efficiency (DCCEE 2009), from the 2006 satellite imagery based on the method of Furby (2002).

Hydrologic

Additional spatial environmental indices related to landform were developed to describe the wetland environments of the Pilbara region. We approached this by first developing an ecological rationale for why particular attributes of the environment may be important based on scientific understanding of aquatic systems in the Pilbara region (Appendix 1B). Indices subsequently derived from available data sources are described in Appendix 1A, additional to other variables (described above).

APPENDIX 1A

Environmental indices compiled and selected as candidates for the Pilbara analysis region.
* Additional indices customised for wetland biodiversity modelling.

Code	Group	Subgroup	Description
COASTdist	Climate	DIST	Euclidean distance to coast in degrees based on 9sec grid
ELEVav	Climate	DIST	Mean elevation within 9sec grid based on 3sec DEM
EVAPmaxM	Climate	ENERGY	Maximum monthly evaporation (mm)
EVAPminM	Climate	ENERGY	Minimum monthly evaporation (mm)
EVAPOTRANS	Climate	ENERGY	MODIS-derived potential evapotranspiration annual mean (10 years)
MaxTEMPcldM	Climate	ENERGY	Maximum temperature coolest (minimum) month (°C)
RngTEMPmaxM	Climate	ENERGY	Maximum monthly diurnal temperature range (°C)
RngTEMPminM	Climate	ENERGY	Minimum monthly diurnal temperature range (°C)
TopoRADmaxM	Climate	ENERGY	Topographically-adjusted solar radiation December (maximum)
TopoRADminM	Climate	ENERGY	Topographically-adjusted solar radiation in June (minimum)
MOISTmaxM	Climate	WATER	Highest Period Moisture Index
MOISTminM	Climate	WATER	Lowest Period Moisture Index
PRECIPdryM	Climate	WATER	Precipitation of the driest (minimum) month (mm)
PRECIPwetM	Climate	WATER	Precipitation of the wettest (maximum) month (mm)
RELHUMIDmaxM	Climate	WATER	Maximum monthly relative humidity (%)
RELHUMIDminM	Climate	WATER	Minimum monthly relative humidity (%)
VAPDEFmaxM	Climate	WATER	Maximum monthly vapour pressure deficit (KPa)
VAPDEFminM	Climate	WATER	Minimum monthly vapour pressure deficit (KPa)
WATERDEFmaxM	Climate	WATER	Maximum monthly precipitation deficit (RAIN minus EVAP) (mm)
WATERDEFminM	Climate	WATER	Minimum monthly precipitation deficit (RAIN minus EVAP) (mm)
AgeROCKav	Regolith	GEOLMAPPING	Mean geological age in millions of years
ROCKfertility	Regolith	GEOLMAPPING	An index of inherent rock fertility
SALINITY	Regolith	HYDROMAPPING	Median water table salinity
ROCKweathering	Regolith	MODELLED	Weathering intensity index
SOILspectra1	Regolith	MODELLED	Spectra of surficial soils – Principal component 1
SOILspectra2	Regolith	MODELLED	Spectra of surficial soils – Principal component 2
SOILspectra3	Regolith	MODELLED	Spectra of surficial soils – Principal component 3
SubSOILkaolin	Regolith	MODELLED	Kaolinite clay minerals in subsurface soil (60–80 cm)
SubSOILsmectite	Regolith	MODELLED	Smectite clay minerals in subsurface soil (60–80 cm)
TopoSOILdepth	Regolith	MODELLED	Soil depth in metres scaled from a topographic wetness index (Claridge *et al.* 2000)
TopoSOILwater	Regolith	MODELLED	Soil water holding capacity scaled from a topographic wetness index (Claridge *et al.* 2000)
TopSOILillite	Regolith	MODELLED	Illite clay minerals in surface soil (0–20 cm)
TopSOILkaolin	Regolith	MODELLED	Kaolinite clay minerals in surface soil (0–20 cm)
TopSOILsmectite	Regolith	MODELLED	Smectite clay minerals in surface soil (0–20 cm)
POTASSIUMpct	Regolith	RADIOMETRICS	Potassium element concentrations (% K)
rSOILRadiometry	Regolith	RADIOMETRICS	Terrestrial dose rate from K, U and Th grids (nG/h)
THORIUMKratio	Regolith	RADIOMETRICS	Ratio of thorium over potassium
THORIUMppm	Regolith	RADIOMETRICS	Thorium element concentrations (ppm eTh)
tSOILRadiometry	Regolith	RADIOMETRICS	Total dose rate (terrestrial + cosmic radiation) (nG/h)
URANIUM2THratio	Regolith	RADIOMETRICS	Ratio of U2 over thorium
URANIUMKratio	Regolith	RADIOMETRICS	Ratio of uranium over potassium

Code	Group	Subgroup	Description
URANIUMppm	Regolith	RADIOMETRICS	Uranium element concentrations (ppm eU)
URANIUMTHratio	Regolith	RADIOMETRICS	Ratio of uranium over thorium
CALCRETE	Regolith	SOILMAPPING	Presence (1) or absence (0) of calcrete in or below the soil profile
SOILclay	Regolith	SOILMAPPING	Soil clay content to 30 cm
SOILcoarse	Regolith	SOILMAPPING	Percent of coarse fragments throughout the soil profile
SOILdensity	Regolith	SOILMAPPING	Bulk soil density to 30 cm
SOILdepth	Regolith	SOILMAPPING	Solum soil depth (m)
SOILhydraulic	Regolith	SOILMAPPING	Average soil horizon saturated hydraulic conductivity
SOILnutrient	Regolith	SOILMAPPING	Gross soil nutrient status
SOILstructure	Regolith	SOILMAPPING	Hydrological scores for grades of soil pedality (ordered by infiltration rate)
SOILwater	Regolith	SOILMAPPING	Plant-available soil water content (mm/100 cm)
ELEVrange1s	Landform	COMPLEX	Mean elevation range (m) within 300 m window derived from 1sec DEM
RIDGEflatness	Landform	COMPLEX	Ridgetop flatness index
TOPOav	Landform	COMPLEX	Zonal majority yellowbook topographic position (slope relief interpreted from Table 5 second edition) extracted from 1sec DEM
TOPOposition	Landform	COMPLEX	Topographic position: upper, middle and lower parts of the landscape
TOPOrange	Landform	COMPLEX	Zonal range yellowbook topographic position (slope relief interpreted from Table 5 second edition) extracted from 1sec DEM
VALLEYflatness	Landform	COMPLEX	Valley bottom flatness class (deposited material) from 3sec DEM
SLOPEav	Landform	SLOPE	Focal mean slope angle in percent from 3sec DEM
SLOPEcurve	Landform	SLOPE	Rate of change of gradient down a slope: flow acceleration, erosion/deposition rates
SLOPErange	Landform	SLOPE	Focal range of percent slope based on 1sec DEM
SLOPEshape	Landform	SLOPE	Rate of change of slope: accumulation or dispersion of water
TOPOflatarea	Landform	SLOPE	A mask for relatively flat areas where DEM cannot discriminate topographic position
avTOPOwetness	Landform	TOPOMOISTURE	Focal mean of topographic wetness index based on 1sec DEM
HILLshade	Landform	TOPOMOISTURE	Topography hillshade data set from 3sec DEM-S
rngTOPOwetness	Landform	TOPOMOISTURE	Focal range of topographic wetness index based on 1sec DEM
TOPOwetness	Landform	TOPOMOISTURE	Compound topographic index: relative wetness based on catchment contributing area
SCLEROint	Biotic	FPAR	Sclerophyll fPAR interception (MODIS 9sec grids)
TURGORint	Biotic	FPAR	Turgour fPAR interception (MODIS 9sec grids)
WOODYVEGcov	Biotic	VEG	Extent of woody vegetation cover within 250 m grids (ha) (LANDSAT 2006)
WOODYVEGdist	Biotic	VEG	Euclidean distance to woody vegetation of at least 1 ha extent (LANDSAT 2006)
*CLIFFdist	Hydrologic	DISTANCE	Distance to cliff lines
*CREEKdist	Hydrologic	DISTANCE	Distance to minor drainage lines
*PERENdist	Hydrologic	DISTANCE	Distance to perennial pool or water body
*RIVERdist	Hydrologic	DISTANCE	Distance to major drainage lines
*RIVERsize	Hydrologic	DISTANCE	Highest catchment stream order
*WATERHOLEdist	Hydrologic	DISTANCE	Distance to water points

APPENDIX 1B

General rationale used to target the development of environmental indices relevant to aquatic systems in the Pilbara region.

Key environmental features influencing the character of aquatic systems	Description/application
Water feature type (e.g. riverine/non-riverine, water body)	The distinction between riverine and non-riverine systems is critical to the composition of the biota. This information is available from river and water body mapping.
Water body permanence	Water body permanence in rivers and non-rivers has a strong influence on fauna and at the ephemeral/seasonal end of the temporal scale on plants. This requires sequential remote sensing data to verify the proportion of winters (dry season) with water present.
Water body turbidity	Turbidity has a strong influence on biotic composition. In the absence of a specific layer, a surrogate could be derived from regolith and terrain attributes in the vicinity of the wetland/river (low rockiness, high clay content, flatness).
Regolith and terrain features associated with water and riparian zone	Regolith and landform influence the physical and chemical attributes of the aquatic and riparian habitats. Measures such as catchment ruggedness, weathering index, clay content and salinity will be relevant. Such measures are important for the riparian plants but also may be surrogates for turbid *vs.* clear pools. Massive rocks reflect gorge sites with steep valley slopes. Topographic shadowing effects will help identify gorges, as well as indices that discriminate a broad flat or a valley, derived from slope measures.
Catchment hierarchy (stream order)	For the aquatic plants there was little compositional difference except between the extremities of the scale. Basically, creeks/springs are different from everything else. For the invertebrates there were clear differences between first and second order streams (i.e. headwaters) and lower order (downstream) sites. Elevation may also be a useful indicator of springs/head water creeks.
Slope along valley floor (Thalweg)	The gradient of the valley floor influences the formation of waterholes and pools. This indicator is derived from altitudinal differences between points along the valley floor, i.e. what is altitudinal difference between a point along the valley floor and another point 1 km downstream.
Slope perpendicular to valley bottom	The steepness of the sides of the river valleys influences the character of waterholes and pools. This could be derived as the slope from the valley floor to the nearest maximum altitude perpendicular to the river channel, or the average slope over a fixed distance (e.g. 200 m), perpendicular to the river channel.
Channel width	An indicator of the volume of habitat available.
River channel shape	Deeper pools tend to be at bends in the river where flows have eroded the bed against cliffs. Straighter sections of rivers tend to have more seasonal pools.
River discharge category	Coastal or internal, size of contributing catchment.
Vegetation density in riparian zone	The presence of dense vegetation within/downstream of first order creeks may indicate presence of springs. Difficult to separate springs from creeks using other variables. At other positions in the landscape it will be indicating the gallery forest of riparian zones of big river pools with its role in shading and carbon inputs.
Climate	Climate broadly influences the region through annual and seasonal rainfall indices and temperature during the wet and dry seasons.

542

L.A. Gibson, K.J. Williams, A.M. Pinder, T.D. Harwood, N.L. McKenzie, S. Ferrier,
M.N. Lyons, A.H. Burbidge and G. Manion

APPENDIX 2

Partial importance of the 34 predictors included in the GDM model of terrestrial fauna determined by successively removing and then replacing each predictor in the model. Relative contribution to the model is shown for the partial % deviance explained and for compositional turnover is the partial sum of coefficients for predictor variable splines. Predictors are sorted by group, then by the partial % deviance explained from largest to smallest.

Predictor	Group	Model intercept	Partial % deviance explained	Partial sum of coefficients
SOILclay	Regolith	0.298951	2.784955	0.347575
SOILspectra1	Regolith	0.301226	1.730941	0.663065
AgeROCKav	Regolith	0.297129	0.825033	-0.069820
SOILdensity	Regolith	0.293673	0.413231	0.154999
SOILwater	Regolith	0.286182	0.307540	-0.141780
ROCKfertility	Regolith	0.289180	0.291539	0.046018
SOILhydraulic	Regolith	0.293222	0.263604	0.152569
TopSOILillite	Regolith	0.296354	0.248420	0.071942
ROCKweathering	Regolith	0.296185	0.221927	0.105939
URANIUMTHratio	Regolith	0.292746	0.166037	0.057519
SOILspectra2	Regolith	0.295651	0.153569	0.046256
SOILspectra3	Regolith	0.295097	0.127903	0.096127
rSOILRadiometry	Regolith	0.291035	0.108918	0.052842
WOODYVEGcov	Biotic	0.283506	1.492938	0.236135
TURGORint	Biotic	0.291291	0.155424	0.096720
SCLEROint	Biotic	0.289195	0.054238	0.096521
Geographic Distance	GEOG	0.383649	0.514363	0.052627
PRECIPwetM	Climate	0.278708	0.481020	0.068095
VAPDEFmaxM	Climate	0.286097	0.283565	0.015435
RngTEMPminM	Climate	0.286173	0.236309	0.122520
PRECIPdryM	Climate	0.284814	0.219142	0.030538
TopoRADminM	Climate	0.290226	0.097757	0.053116
MOISTminM	Climate	0.286179	0.085649	-0.002450
EVAPOTRANS	Climate	0.294346	0.081462	0.047981
avTOPOwetness	Landform	0.299923	0.352871	0.163034
ELEVrange1s	Landform	0.287143	0.201984	0.011518
HILLshade	Landform	0.289405	0.194709	0.114221
SLOPEshape	Landform	0.296649	0.148605	0.107851
TOPOflatarea	Landform	0.287174	0.138988	-0.009520
RIDGEflatness	Landform	0.295549	0.118471	0.019753
SLOPErange	Landform	0.287166	0.077396	-0.004800
PERENdist	Hydrologic	0.294549	0.337370	0.084439
CREEKdist	Hydrologic	0.296091	0.297972	0.060220
WATERHOLEdist	Hydrologic	0.295536	0.137967	0.017756

APPENDIX 3

Partial importance of the 31 predictors included in the GDM model of wetland biodiversity determined by successively removing and then replacing each predictor in the model. Relative contribution to the model is shown for the partial % deviance explained and for compositional turnover is the partial sum of coefficients for predictor variable splines. Predictors are sorted by group, then by the partial % deviance explained from largest to smallest.

Predictor	Group	Model intercept	Partial % deviance explained	Partial sum of coefficients
SOILspectra1	Regolith	0.440868	3.726465	0.284270
SOILspectra3	Regolith	0.397573	1.915270	0.273269
SOILcoarse	Regolith	0.349937	1.604874	0.441083
SOILclay	Regolith	0.390150	0.854646	0.257399
CALCRETE	Regolith	0.431222	0.803488	0.285845
THORIUMppm	Regolith	0.392662	0.606473	-0.052733
SOILwater	Regolith	0.386290	0.475515	0.098623
POTASSIUMpct	Regolith	0.398290	0.181537	0.084127
SALINITY	Regolith	0.396262	0.146821	0.129120
TopoSOILdepth	Regolith	0.394146	0.142990	0.111956
tSOILRadiometry	Regolith	0.384640	0.116939	0.111024
SOILhydraulic	Regolith	0.397950	0.114711	0.042447
TopoSOILwater	Regolith	0.383324	0.066574	0.050283
SOILdensity	Regolith	0.389579	0.063597	0.025561
VALLEYflatness	Landform	0.387216	2.317408	0.265478
RIDGEflatness	Landform	0.406098	0.788395	0.097823
ELEVrange1s	Landform	0.380524	0.415711	0.009781
TOPOrange	Landform	0.39888	0.227913	0.033557
TOPOav	Landform	0.382598	0.192719	-0.020880
TOPOposition	Landform	0.396753	0.069683	0.038118
RIVERdist	Hydrologic	0.402223	0.400610	0.058104
WATERHOLEdist	Hydrologic	0.400425	0.223623	0.005736
EVAPOTRANS	Climate	0.406562	1.234791	0.408109
MOISTmaxM	Climate	0.426949	0.693327	0.257617
TopoRADmaxM	Climate	0.410757	0.567435	0.223142
ELEVav	Climate	0.390900	0.486699	0.146071
MOISTminM	Climate	0.388643	0.452929	0.130371
RngTEMPminM	Climate	0.386065	0.382681	0.072267
RngTEMPmaxM	Climate	0.388513	0.120242	0.045086
RELHUMIDminM	Climate	0.386680	0.082719	0.033791
WOODYVEGdist	Biotic	0.402493	0.180799	0.083225

L.A. Gibson, K.J. Williams, A.M. Pinder, T.D. Harwood, N.L. McKenzie, S. Ferrier,
M.N. Lyons, A.H. Burbidge and G. Manion

APPENDIX 4A

Terrestrial biodiversity variable partitioning based on using 77 candidate variables. Partitioning groups: biotic and climate (BC), hydrologic and landform (HL), regolith – rock and soil (RS).

Partitions*	Intercept	%Deviance explained	Sum of coefficients	Predictors	% Deviance partitioning	% Unexplained deviance
BC+HL+RS	0.255	46.5	5.92	66 of 77	9.86	
BC	0.648	18.2	2.48	21	4.03	
HL	0.443	23.9	2.04	20	4.57	
RS	0.370	37.1	4.22	31	15.12	
BC+HL	0.419	31.4	3.54	38	0.83	
BC+RS	0.355	42.0	5.22	45	3.52	
HL+RS	0.248	42.5	5.10	49	8.61	
TOTAL					**46.6**	**53. 5**

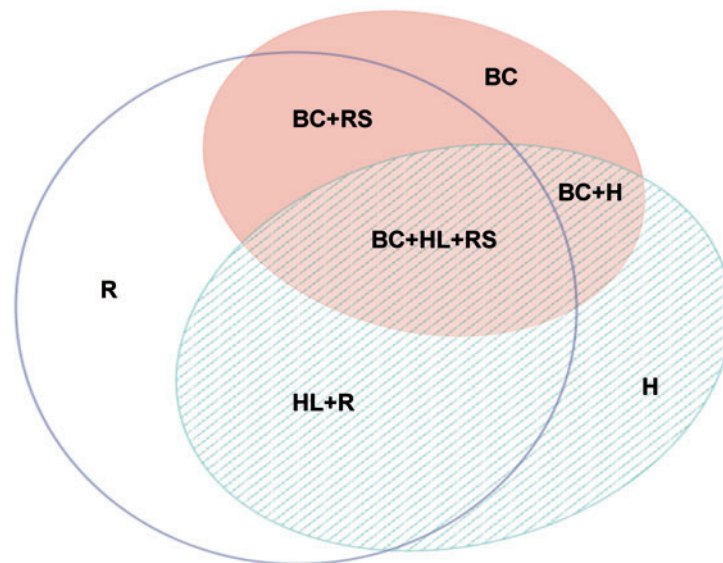* Geographic distance was included as a predictor in all cases
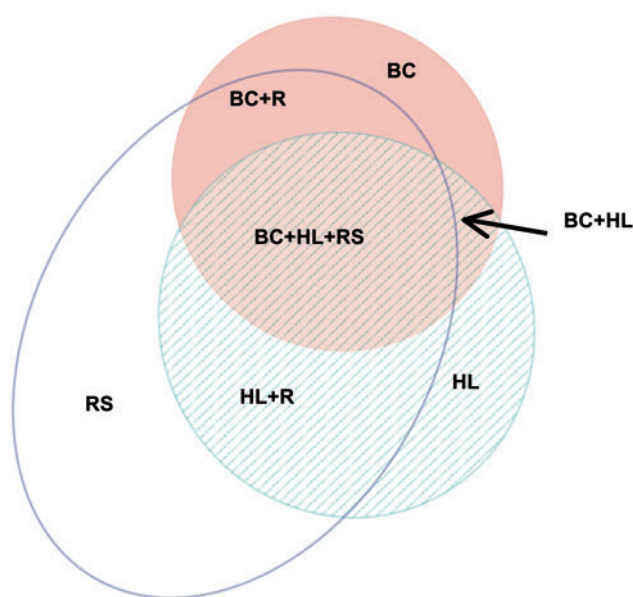
Shared and independent contributions of the groups to explaining the deviance – developed using *eulerAPE* software by Micallef and Rodgers (2014).

APPENDIX 4B

Wetland biodiversity variable partitioning based on using 77 candidate variables. Partitioning groups: biotic and climate (BC), hydrologic and landform (HL), regolith – rock and soil (RS).

Partitions*	Intercept	%Deviance explained	Sum of coefficients	Predictors	% Deviance partitioning	% Unexplained deviance
BC+HL+RS	0.303	59.3	8.44	48 of 77	7.64	
BC	0.891	18.6	2.82	15	3.86	
HL	0.764	38.1	3.22	16	7.07	
RS	0.594	46.5	6.78	23	12.05	
BC+HL	0.528	47.2	5.11	30	1.82	
BC+RS	0.441	52.2	7.96	35	5.23	
HL+RS	0.425	55.4	7.54	37	21.62	
TOTAL					**59.3**	**40.7**

Shared and independent contributions of the groups to explaining the deviance – developed using *eulerAPE* software by Micallef and Rodgers (2014).