ECOGRAPHY 29: 733–744, 2006

Predicting patterns of plant species richness in megadiverse South Africa

Wilfried Thuiller, Guy F. Midgley, Mathieu Rouget and Richard M. Cowling

Thuiller, W., Midgley, G. F., Rouget, M. and Cowling, R. M. 2006. Predicting patterns of plant species richness in megadiverse South Africa. – Ecography 29: 733–744.

Using new tools (boosted regression trees) in predictive biogeography, with extensive spatial distribution data for >19 000 species, we developed predictive models for South African plant species richness patterns. Further, biome level analysis explored possible functional determinants of country-wide regional species richness. Finally, to test model reliability independently, we predicted potential alien invasive plant species richness with an independent dataset. Amongst the different hypotheses generally invoked to explain species diversity (energy, favorableness, topographic heterogeneity, irregularity and seasonality), results revealed topographic heterogeneity as the most powerful single explanatory variable for indigenous South African plant species richness. Some biome-specific responses were observed, i.e. two of the five analyzed biomes (Fynbos and Grassland) had richness best explained by the "species-favorableness" hypothesis, but even in this case, topographic heterogeneity was also a primary predictor. This analysis, the largest conducted on an almost exhaustive species sample in a species-rich region, demonstrates the preeminence of topographic heterogeneity in shaping the spatial pattern of regional plant species richness. Model reliability was confirmed by the considerable predictive power for alien invasive species richness. It thus appears that topographic heterogeneity controls species richness in two main ways: firstly, by providing an abundance of ecological niches in contemporary space (revealed by alien invasive species richness relationships) and secondly, by facilitating the persistence of ecological niches through time. The extraordinary richness of the South African Fynbos biome, a world-renowned hotspot of biodiversity with the steepest environmental gradients in South Africa, may thus have arisen through both mechanisms. Comparisons with similar regions of the world outside South Africa are needed to confirm the generality of topographic heterogeneity and favorableness as predictors of plant richness.

W. Thuiller (wilfried.thuiller@ujf-grenoble.fr), Laboratoire d'Ecologie Alpine, UMR-CNRS 5553, Univ. Joseph Fourier, BP 53, FR-38041 Grenoble Cedex 9, France. — G. F. Midgley, Kirstenbosch Research Centre, South African National Biodiversity Inst., P/Bag x7, Claremont 7735, Cape Town, South Africa. — M. Rouget, South African National Biodiversity Inst., Private Bag x101, Pretoria 0001, South Africa. — R. M. Cowling, Dept of Botany, Nelson Mandela Metropolitan Univ., PO Box 77000, Port Elizabeth 6031, South Africa.

For much of the last century, regional-scale species diversity patterns were studied by biogeographers concerned with faunal and floristic richness in different parts of the globe (Hutchinson 1959). Resulting narrative analyses relied heavily on geographical and historical factors. Over the past three decades, ecologists have sought predictors of regional-scale patterns, using a wide range of explanatory variables, but especially measures of available energy and environmental heterogeneity (Currie 1991, Francis and Currie 1998, Whittaker

Accepted 21 June 2006

ISSN 0906-7590

1999, Whittaker et al. 2001, Chown et al. 2003, Field et al. 2005, Storch et al. 2005). A new synthesis of diversity theory is now emerging that acknowledges the role of ecology, geography and evolution in determining diversity patterns at all spatial scales (Gaston 2000, Ricklefs 2000, Hubbell 2001, Currie et al. 2004, Ricklefs et al. 2004), and ecological and historical perspectives have been fused.

In this paper we analyze plant regional richness patterns in South Africa. This country is ideal for such analyses due to its high environmental heterogeneity and extraordinarily rich flora (Goldblatt 1978, Cowling et al. 1989). Furthermore, it contains not only some of the most species-rich regions in the world (such as the southwestern part of fynbos biome), but also areas of floristic impoverishment, such as parts of the Upper Karoo and Kalahari Basin (Cowling et al. 1989, Cowling and Hilton-Taylor 1992, Cowling et al. 1997a).

Previous work on regional-scale plant diversity patterns in South Africa provides a somewhat confusing picture on the relative role of climate, available energy and heterogeneity explanatory variables, especially because individual biomes or climate regions have been analysed individually. O'Brien (1993, 1998) demonstrated a curvilinear relationship between energy and species richness of trees in one degree squares; overall, richness was best explained by a simple model based on a linear relationship with annual rainfall and a parabolic relationship with minimum monthly potential evapotranspiration (PET). The analysis excluded the vast majority of South African woody plants (low and dwarf shrubs associated with fynbos and karoo vegetation) limiting the generality of the pattern (Cowling et al. 1997b). Hoffman et al. (1994) showed a negative relationship between richness and available energy (PET) for a full complement of plant species from intensively surveyed sites of roughly equal area across South Africa, interpreted as a consequence of high water deficit impacting negatively on richness in areas of high PET (the Kalahari and Nama-karoo deserts). Both O'Brien (1993) and Hoffman et al. (1994) identified topographic heterogeneity as a relatively unimportant factor in explaining richness, but nonetheless accounting for some otherwise unexplained variance in their models.

Using 12 explanatory variables to reflect heterogeneity, favorableness, energy, seasonality and irregularity, Cowling et al. (1997b) demonstrated that heterogeneity explained plant species richness in the temperate biomes (Fynbos and Karoo), but productivity was the strongest predictor of plants species richness in the subtropical biomes (Grassland and Savanna). They argued that speciation/extinction history, and concomitantly high levels of beta and gamma diversity, was important in the former, whereas ecology was more important in the latter and richness peaked in productive environments as expected for a flora of tropical origins.

Substantial effort has focused on explaining the high richness in the Fynbos biome. While topographic and climatic heterogeneity are good predictors of species richness at the regional scale within the western, strongly winter-rainfall part of the biome (Linder 1991), other studies suggest that historical factors are pre-eminent. Using species-area analysis of regional-scale data sets, and analysis of covariance, Cowling and Lombard (2002) showed a clear geographic diversity pattern in the Fynbos biome: western sites with high winter rainfall have more than double the number of species as eastern sites, a difference which is not explicable by heterogeneity or available energy. Cowling and Lombard (2002) invoked historical factors, arguing that higher Pleistocene climatic stability in the west had resulted in elevated speciation rates and depressed extinction rates there, and hence higher regional richness than in the less climatically stable east, further supported by the higher incidence of natural rare species (both neoendemics and relictual endemics) in the west. A similar climate-stability explanation has been posited to explain the west-east gradient in regional richness in the karroid biomes (Cowling et al. 1998). There is now considerable evidence that climatically stable Pleistocene environments are associated with high richness and endemism across many taxa (Fjeldså and Lovett 1997, Midgley et al. 2001, Cowling et al. 2004, Dynesius et al. 2004).

No general explanation for patterns of regional-scale plant diversity in South Africa is accepted, and the wide range of diversity values recorded here within several phylogenetically distinct floras and six regionally extensive biomes offers great potential for exploring determinants of regional plant richness patterns. Such a study could also overcome problems of previous studies since a) we have region-wide data on $>19\,000$ plant species at a fixed spatial scale, therefore avoiding the pitfalls of variably sized sites and the attendant species-area analyses required to overcome these; and b) new tools in predictive biogeography are now available. Hence, we now have an opportunity to gain a predictive understanding of richness across an entire megadiverse country.

To assess the many deterministic explanations of regional richness (Table 1), we have used a wide array of explanatory variables reflecting contemporary environmental conditions, namely topographic heterogeneity, energy, bioclimatic suitability, seasonality and irregularity (Table 2). We do not address the role historic drivers such as climate history in this analysis, though topographic heterogeneity can be interpreted in this light through providing climatic refugia. Finally, as an independent test of model reliability, we predict invasive plant diversity in South Africa and compare this with observed patterns.

Table 1. Common hypotheses invoked to explain species richness pattern in southern Africa.

General hypothesis	Result	Southern Africa study	Scale
Available energy	Curvilinear relationship between tree species richness and potential evapotranspiration.	O'Brien 1993, 1998	Southern Africa
	Negative and linear relationship between species richness and potential evapotranspiration in arid and semi-arid South Africa.	Hoffman et al. 1994	Semi arid South Africa
Topographic heterogeneity	Positive relationship for Fynbos and the two karroid biomes.	Cowling et al. 1997b	South Africa (plots)
Favourableness	Positive relationship between woody species richness and rainfall.	O'Brien 1993, 1998	Southern Africa
	Weak relationship between plant species and rainfall in semi-arid South Africa.	Hoffman et al. 1994	Regional
Climate seasonality and irregularity	Plant species richness negatively related to rainfall reliability.	Cowling et al. 1994, 1997b	South Africa (plots)

Materials and methods

Indigenous species richness datasets

Distribution data for the South African plant species were extracted from the PRECIS database (Germishuizen and Meyer 2003), the largest plant specimen collection in Africa, comprising georeferenced data of species presences only (at quarter degree square resolution, QDS, ca 25 × 25 km at this latitude), for >27 000 plant taxa and comprising 800 000 records, digitised from specimens collected over more than a century, and held in the National Herbarium of the SANBI (for other uses of this dataset, see Thuiller et al. 2005). Despite the large number of records, some geographic areas have been under-sampled (Gibbs Russel et al. 1984). To remove these from the analysis, QDS in which <20 species were recorded were rejected, following Rouget et al. (2004), leaving 1513 pixels. Taking into account the spatial resolution used in this study (25 × 25 km), as well as the relatively complete surveys, we assume that absences in PRECIS (i.e. where a particular species has not been recorded) are reliable. We define species here to include subspecies and varieties, and merge their presence data. Alien species data were excluded from this data set. We extracted data for 19 000 indigenous species to produce this database of indigenous plant species richness in South Africa.

Table 2. Explanatory variables used to model pattern of plant species richness at the regional (QDS) scale in South African biomes. The variables are grouped according to hypotheses invoked to explain patterns of regional richness. Abbreviations used throughout the text.

Hypothesis variable	Derivation
Topographic heterogeneity	
AHI	StD altitude: standard deviation of all the grid altitude values at 200 × 200 m in a QDS pixel.
Favourableness	
MAP^1	Mean annual precipitation (mm)
MTC^1	Mean temperature of the coldest month (°C)
Energy	
PET	Potential evapotranspiration
AET/PET^3	Index of humidity: ratio actual to potential evapotranspiration
MAT^1	Mean annual temperature (°C)
NPP^4	Net primary productivity (ton ha^{-1})
Seasonality	
PPI^1	Plant productivity index: number of months per year receiving more rainfall (mm) than twice the mean annual temperature (°C)
Irregularity	
PCV^1	Coefficient of variation of annual rainfall

[1] Climatic variables have been derived from Schulze (1997) at a resolution of 1 × 1 km, and were scaled up to QDS by averaging.
[2] Potential evapotranspiration estimates were calculated using the FAO 56 Penman Monteith combination equation.
[3] The index of humidity was derived using the LPJ dynamic global vegetation model (Hickler et al. 2004) at a resolution of 10 × 10, and were scaled up to QDS by averaging.
[4] The net primary productivity was derived using the Sheffield dynamic global vegetation model (Woodward and Lomas 2004) at a resolution of 12 × 12, and were scaled up to QDS by averaging.

Invasive species richness datasets

In this part of the study, invasive alien species data only are used to test the predictive power of the models. We hypothesized that invasive species will reflect suitable contemporary environmental conditions for the indigenous flora: hence, we would expect a strong positive relationship between indigenous and alien plant diversity if historical determinants of species richness are not significant (Brown 1988, Lonsdale 1999, Stohlgren et al. 2003, Richardson et al. 2005). Using invasive species to evaluate the model predictive accuracy could be seen as somewhat problematic because the equilibrium assumption of species models are more likely violated than with indigenous species (Guisan and Thuiller 2005). However, several recent studies (Stohlgren et al. 2003, Richardson et al. 2005) have shown positive correlations between indigenous and invasive species richness, demonstrating that despite individual invasive species not achieving equilibrium, the sum of all the invasive species is a good surrogate of environmental suitability (Richardson et al. 2005).

The Southern African Plant Invaders Atlas (SAPIA) is the best source of data on the distribution of invasive alien plants in South Africa, Lesotho and Swaziland. The SAPIA database has records for 500 species, with information on their distribution, abundance, habitat preferences, and dates of introduction (Henderson 1998, 1999, 2001). Records are geo-referenced at a QDS resolution. Only alien species invading natural or semi-natural habitats are listed in SAPIA (weeds of agricultural lands and human-dominated systems are not included); we considered all species in SAPIA to be naturalized or invasive (sensu Richardson et al. 2000).

Biome distribution

To assess consistent differences in richness patterns between biomes (defined in terms of climatic and vegetation structure criteria) within South Africa (Rutherford and Westfall 1994), we developed models both for the entire database, and separately for different biomes. We used the biome map recently compiled following a rigorous regional vegetation mapping exercise (Mucina and Rutherford 2005). This map is the best available dataset of its kind at a scale appropriate for this study. Eight major biomes have been identified for South Africa, namely, Desert, Succulent Karoo, Nama-karoo, Fynbos, Albany thicket, Forest, Grassland and Savanna. The forest biome was excluded in the analysis owing to its fragmented and small extent. The remaining biome polygons were rasterised at QDS grid resolution to match the climatic and species datasets. Because of this rasterisation and the fact they have a particularly low extent, the Desert and the Albany thicket biomes were also removed. We thus retained five biomes, two semi-arid (Nama and succulent karoo), a Mediterranean type (fynbos), a grassland and a savanna biome. To avoid ecotonal areas, we considered only QDS where the dominant biome covers $>75\%$ of the QDS area.

Environmental variables

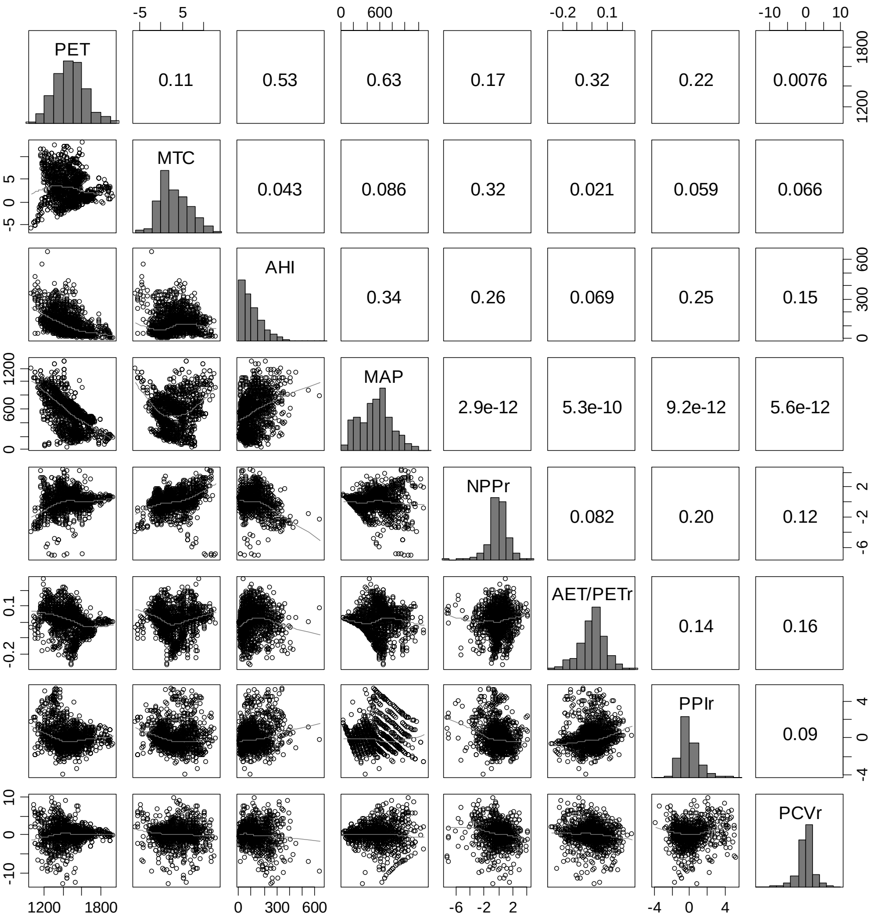
The variables which were selected to test hypotheses on regional richness (Richerson and Lum 1980, Currie 1991) are shown in Table 2, where they are grouped according to the appropriate hypothesis and the methods of derivation are described. We acknowledge that some of the variables were appropriate for more than one hypothesis. All environmental data were available at a 1 km-resolution and were rescaled to match the resolution of the species richness data.

As expected, several variables were inter-correlated with precipitation (e.g. net primary productivity (NPP), index of humidity (AET/PET), coefficient of variation of annual rainfall (PCV), plant productivity index (PPI)). To avoid problems in regression fitting caused by the strong correlations among precipitation, and net primary productivity, index of humidity and coefficient of variation of annual rainfall, we applied transformation to NPP, AET/PET, PCV and PPI (see Leathwick et al. 2006 for more details on the approach). First, we adjusted these variables for precipitation by fitting univariate non-linear regressions (generalised additive model with 4-degrees of freedom) describing the average relationship between precipitation and the other correlated variables. We then used the residuals of each of the univariate regression as a predictor (NPPr, AET/PETr, PCVr, PPIr), these indicating for any pixel the deviation from the average NPP, AET/PET, PCV or PPI expected at its precipitation level. For instance, for net primary productivity, positive values of NPPr indicate very humid and productive areas where NPP is higher than would be expected from an average precipitation level (eastern part of South Africa, along the Indian Ocean coast). Negative values indicate very unproductive sites compared to the level of precipitation found mostly in the Cape Floristic Region and in the central highlands of South Africa. After correcting for multi-collinearity, all the variables selected were poorly auto-correlated, ensuring meaningful variable selection (Fig. 1).

Models

We developed sophisticated multivariate models (boosting methods, see also Elith et al. 2006, Leathwick et al. 2006) to explain indigenous species richness and predict both indigenous and alien invasive richness. Least-square or log-likelihood models are usually performed for such purposes (Pausas and Austin 2001) using stepwise regression to rank and select the most explanatory variables. While the latter fit the single most parsimonious explanatory model for species richness given a range of ecological predictors, boosting methods fit a large number of relatively simple models whose predictions are combined to give more robust estimates of the response. The algorithm we used in this study was a boosted regression tree assuming a Poisson distribution of species richness (BRT, Ridgeway 1999, Friedman 2001) where each individual model consists of a simple classification or regression tree (see also Rouget and Richardson 2003, Thuiller et al. 2003, Thuiller 2004), i.e. a rule based classifier that consists of recursive partitions of the dimensional space defined by the predictors into groups that are as homogeneous as possible in terms of response (Hastie et al. 2001). The tree is built by repeatedly splitting the data, defined by a simple rule

Fig. 1. Scatter plot matrix for all the variables selected for the models (see Table 2 for the abbreviations). The lower diagonal represents the correlation plots between the variables and the associated smooth curve showing the main correlation trend. The diagonal represents the frequency histogram of the variables. The upper panel represents the correlation values between the variables.



based on a single explanatory variable. At each split, the data are partitioned into two exclusive groups, each of which is as homogeneous as possible (Breiman et al. 1984).

Ordinary generalised linear models have the form:

$$f(x) = \sum_{j=1}^{j} \beta_j x_j$$

where the algorithm seeks to estimate the β_j throughout various optimisation procedures (often maximum likelihood estimation). Special cases of basis expansions like generalised additive models (GAM) have also been using the same form:

$$f(x) = \sum_{j=1}^{j} \beta_j h_j(x)$$

where h(x) is a non parametric function (e.g. spline). These methods have so far fixed the h_js and then found β_j using standard techniques (e.g. ordinary least squares regression – OLS). Regression trees also have this form where the h_js are indicator functions indicating whether x falls into a particular "box" and β is just the terminal node means. Regression trees do not preselect the h_js nor J, rather they are estimated iteratively through the recursive partitioning algorithm (Breiman et al. 1984).

BRT makes each h_j take the form of a regression tree. They are fitted incrementally so that $h_1(x)$ is the single best tree, $h_2(x)$ is the best tree that predicts the residuals of $h_1(x)$, and so on (Friedman et al. 2000). By this method, the BRT uses an iterative method for developing a final model progressively adding trees to the model, while re-weighting the data to emphasise cases poorly predicted by the previous trees.

For this study, given our objective of producing a highly general model, we used a 50-fold cross-validation procedure to identify an optimal number of trees that maximises the ability of a model to make accurate predictions to independent sites while avoiding excessive model complexity. We also selected a tree size of 2, allowing inclusion of simple two-term interactions between variables.

More importantly, BRT allowed the estimation of the relative importance of each variable in the model. We used a permutation method, which randomly permutates each predictor variable independently, and computes the associated reduction in predictive performance.

To run the generalised booted model, we used the library GBM implemented by Ridgeway (<http://www.i-pensieri.com/gregr/gbm.shtm >) under the free source R-CRAN software (Ihaca and Gentleman 1996).

Finally, we also computed the response curve of the regional plant species richness in South African biomes against the variable with the highest explanatory power. To do this, we used the evaluation strip, a novel, robust method for plotting predicted responses from statistical models proposed recently by Elith et al. (2005).

Results

Model accuracy

Predictions of species richness derived from BRT models showed significant correlations with observed species richness for all of our datasets (Table 3). The highest correlation was observed for the Fynbos biome, evaluated using both the initial and the independent alien invasive dataset as test datasets. Models for the Savanna biome and for all South African biomes combined were also highly predictive using both the initial and the independent alien invasive dataset as test datasets (Table 3). The lowest correlation was observed for the Nama-karoo where the BRT rather efficiently predicted current regional species richness, but only poorly predicted alien invasive richness.

Test of hypotheses

The index of topographic heterogeneity emerged as the most influential variable in explaining and predicting plant diversity pattern country-wide and in most of the biomes (AHI, Table 3, Fig. 2). However, in the Fynbos and Grassland biomes patterns were best explained by mean annual precipitation, a variable associated with the favorableness hypothesis. Overall, species diversity patterns in South Africa were well predicted by a simple index of topographic heterogeneity (relative influence higher than 45%, Fig. 2). As spatial heterogeneity increased, the model predicted an increase in species richness until a plateau was reached (Fig. 3). The same pattern was observed for the two karroid and the Savanna biomes, where diversity was predicted to be strongly positively influenced by topographic heterogeneity and temperature of the coldest month (linear positive relationships).

Though topographic heterogeneity was the main driver of the species diversity in most biomes, other variables were also important (Fig. 2). For instance, Nama-karoo biome patterns were also strongly influenced by the index of humidity (AET/PETr: 27% of relative influence, polynomial positive relationship), a surrogate of available energy in semi-arid region like South Africa. Fynbos biome patterns were most strongly predicted by three variables associated with three different diversity hypotheses, the first being mean annual precipitation, invoked for the favorableness hypothesis, explaining 31% of the total variation. The second and third (net primary productivity (residuals) with a negative sigmoid relationship, and topographic heterogeneity with a positive sigmoid relationship), explaining 23 and 20% of the variation, respectively, were related to the energy (albeit the observed relationship was opposite that predicted by the hypothesis) and heterogeneity hypotheses, respectively. Species richness in the two karroid biomes was predicted to be strongly driven by index of topographic heterogeneity. Nama-karoo species diversity also seems explained by both the residual index of humidity (negative sigmoid relationship), and the residual precipitation coefficient of variation (bell-shaped curve centred on 0). Such relationships imply that Nama-karoo species diversity is higher where the index of humidity is lower than expected as a function of mean precipitation, and is higher when precipitation variation is not extreme.

Finally, Grassland biome diversity patterns were not driven primarily by topographic heterogeneity, but by precipitation (positive sigmoid relationship (Fig. 3)) and residual rainfall irregularity (negative sigmoid relationship) implying that species diversity here is higher for lower irregularity than predicted by mean precipitation.

Prediction of plant diversity in South Africa

The spatial pattern of plant diversity in South Africa was well summarised by the generalised boosted models

Table 3. Variables selected (sum of importance >60%) by the generalized boosted models, and the Pearson Correlation (r_p) between the predicted data and the PRECIS database (used for the models), and) the independent SAPIA invasive alien plant database (not used for the model). All the correlations are significant ($p < 0.001$). See Table 2 for explanation of abbreviations for explanatory variables.

Biome	Environmental factors	r_p (PRECIS)	r_p (SAPIA)
Fynbos	MAP; NPPr; AHI	0.798	0.534
Nama-karoo	AHI; AET/PETr; NPPr; PPIr	0.515	0.225
Succulent karoo	AHI; MTC; MAP	0.660	0.344
Grassland	MAP; PCVr; AHI	0.600	0.280
Savanna	AHI; MAP; PET	0.655	0.543
Allbiomes	AHI; MTC, MAP	0.674	0.470

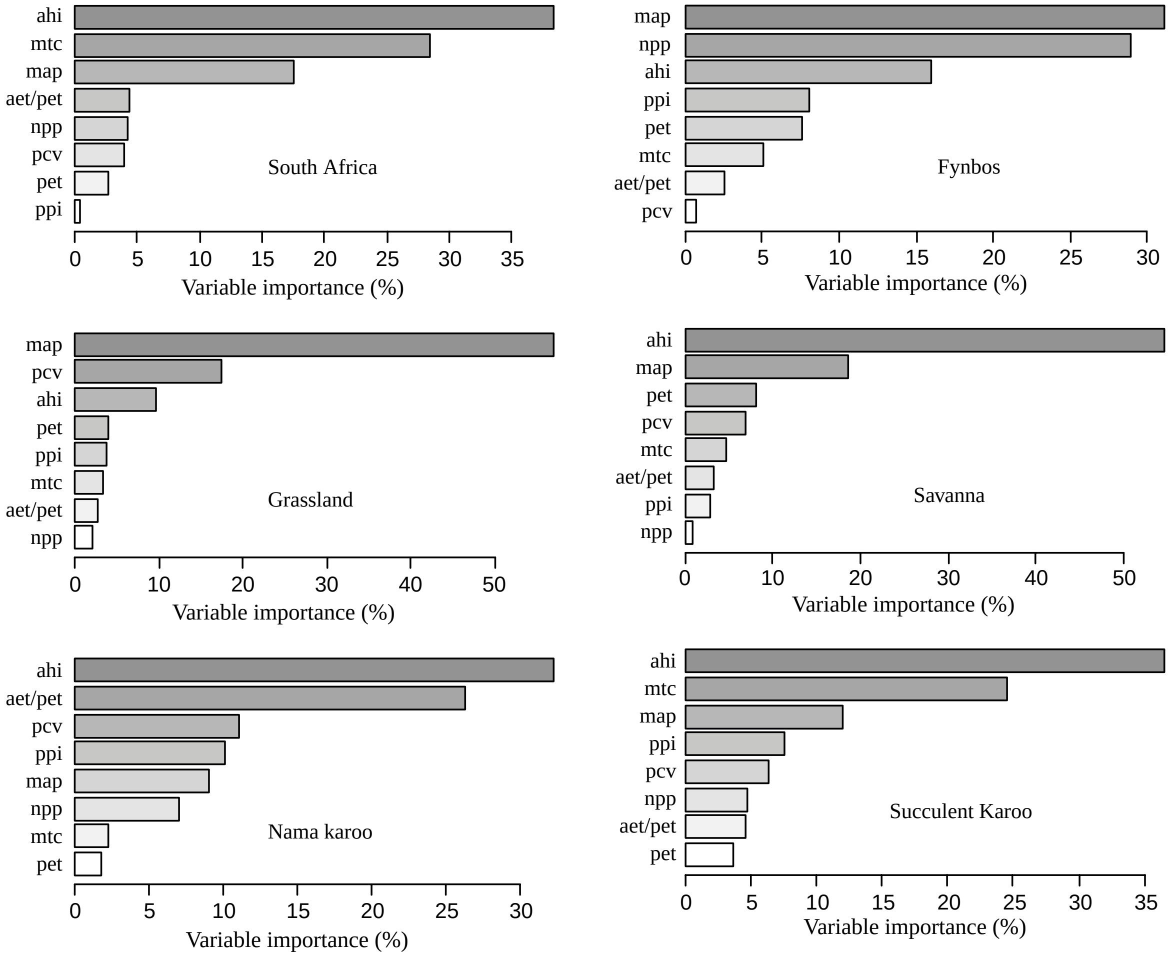


Fig. 2. Relative importance of the environmental variables selected in generalized boosted models in explaining the plant diversity patterns in South Africa, for the all regions and for the different biomes. See Table 2 for explanation of abbreviations for explanatory variables.

(Fig. 4). An important feature of the (observed) PRECIS database was the bias for high diversity to be recorded around the major towns and along the major roads. This was notable in Fig. 4A, where the Gauteng urban complex was observed as one of the most species-rich areas in South Africa, although this may actually be an artefact of over-representation in the database.

The predicted maps (BRT for all South Africa and aggregated BRTs for biomes) produced a much smoother pattern than is evident in the observed species richness map (Fig. 4B, C). The high plant observed species richness node observed in Gauteng was thus predicted to be moderately rich, while the Mpumalanga Drakensberg escarpment was predicted to have a higher richness than observed, a pattern to be expected in a poorly explored and climatically and topographically heterogeneous region of high plant richness and endemism (Matthews et al. 1993). Another interesting point is that the difference in species richness in the two karroid biomes was not clearly visible using the PRECIS database, yet the predicted patterns showed the Succulent Karoo to be markedly richer than the Nama-karoo, a pattern observed using species-area analyses (Cowling et al. 1998)

Interestingly, there were some important differences between the maps produced using the model for all South Africa and the models derived separately for each biome. Not surprisingly, the major difference occurred in the Grassland biome where the overall model and the biome-specific model did not use the same variables. This is noticeable around the Drakensberg Mountains, where the model for South Africa predicted very high species richness due to the high topographic heterogeneity of the region. Alternatively, the biome specific model did not predict high plant species richness in this area, because the measure of heterogeneity was not the most important variable for explaining patterns of richness here. Another major difference occurred for

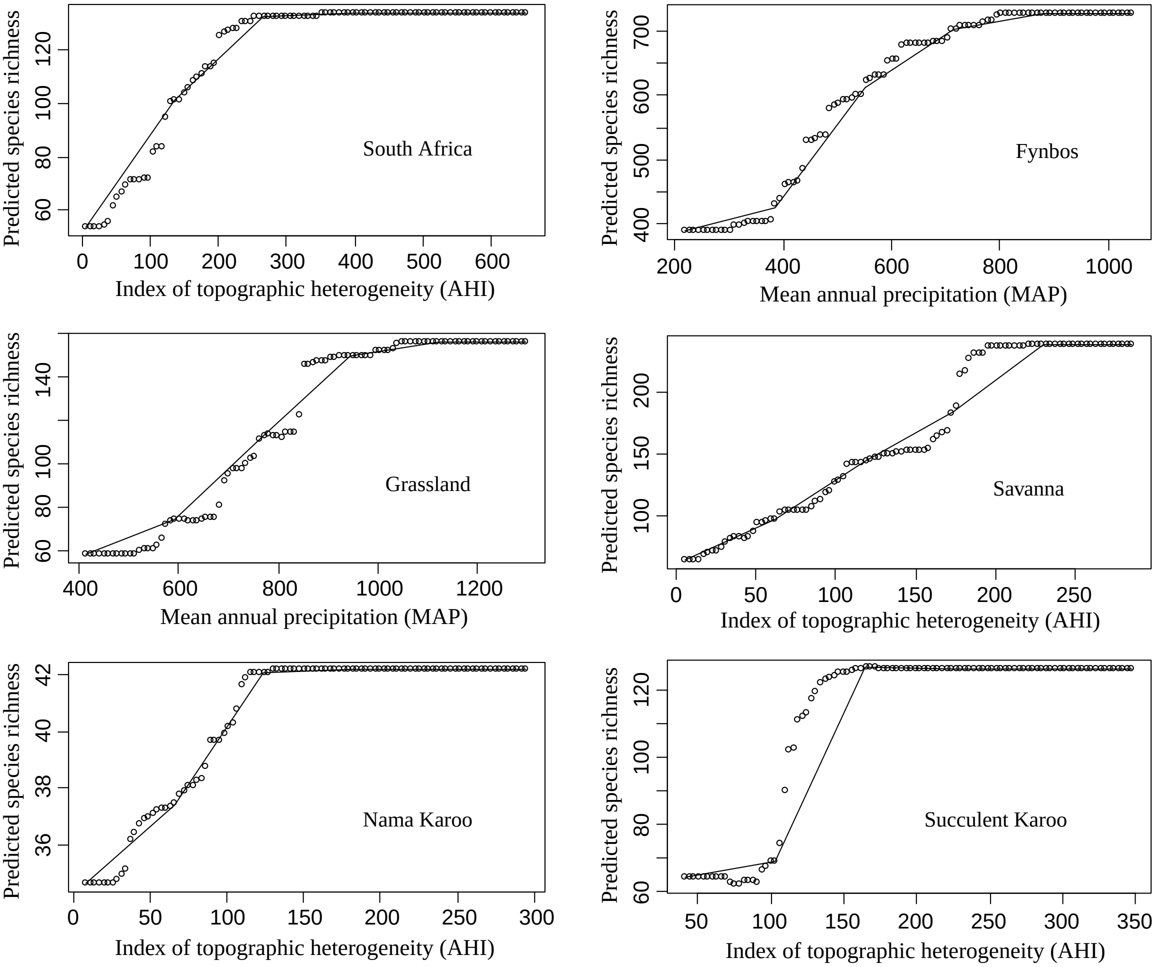


Fig. 3. Predicted species richness response to the most influential variable of the generalized boosted models for all South Africa and different biomes.

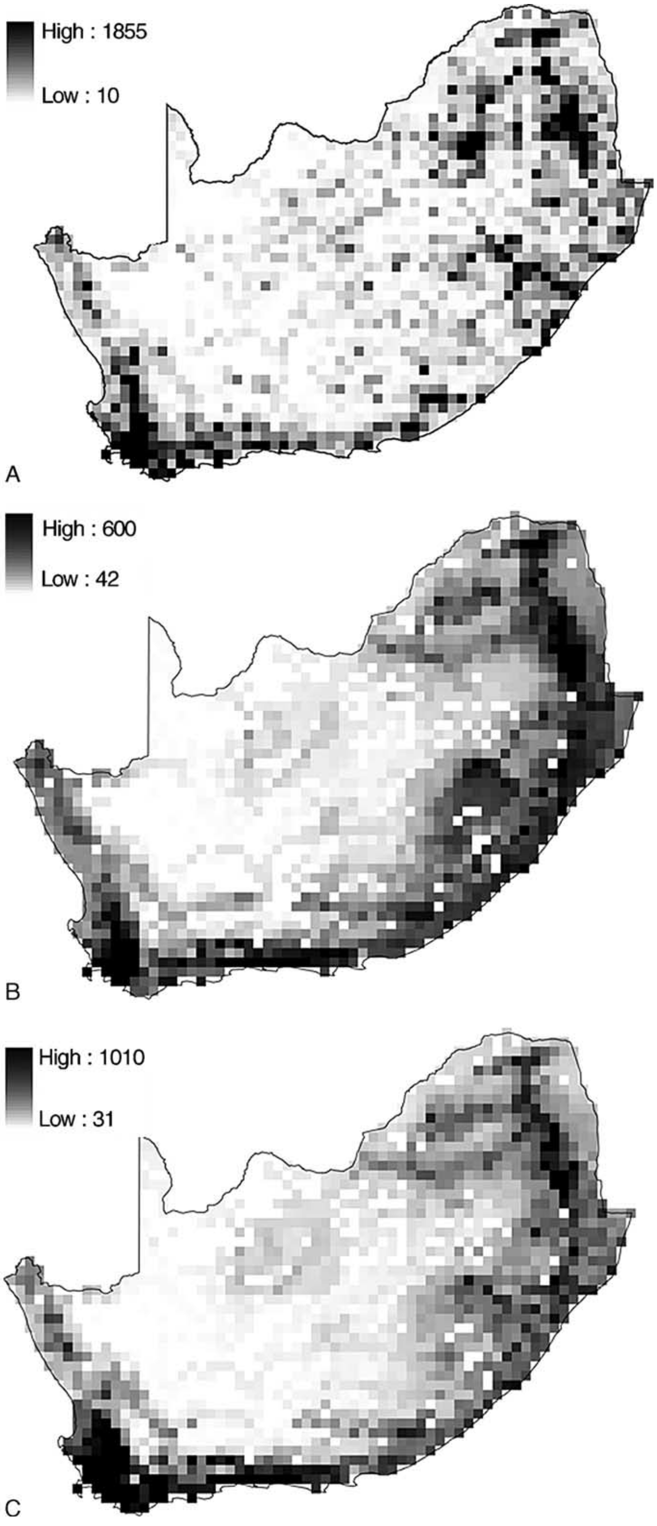
the karroid biomes. Using the model developed for South Africa, the distinction between the two biomes is not straightforward, the model predicting some areas with moderate richness in the Nama-karoo (a species-poor biome). On the other hand, the biome-specific map accurately distinguished the two karroid biomes.

Discussion

Pattern of regional richness in South Africa

The spatial pattern of species richness in South Africa demonstrates that species are neither uniformly nor randomly distributed across the landscape. Generally, areas of high richness are predicted to occur along and below the Great Escarpment, an ancient geomorphological feature associated with the break up of Gondwana, which separates the highly dissected and relatively warm coastal plain from a somewhat featureless and frost-prone interior plateau. A similar conclusion was reached by Cowling et al. (1997b), albeit with cruder data and analyses. Not surprisingly, both observed and modelled distributions show that regions of especially high richness occur in the highly dissected landscapes of the southwestern Fynbos biome, and some areas along the subtropical east coast and eastern escarpment, where there is a complex mosaic of forest, savanna and grassland communities. Regions of relatively low diversity occur mainly in the arid, central areas of the Nama-karoo biome. The Succulent karoo is predicted to be richer than the Nama-karoo, as has already been observed in the literature (Cowling et al. 1998). However, western Fynbos regions are not predicted to be richer than those in the east, possibly because rainfall reliability was modelled on an annual rather than a seasonal basis (see Cowling and Lombard 2002), but also owing to the role of historical climatic changes which we do not explore in this paper.

Fig. 4. Maps of (A) observed species richness, (B) predicted richness using generalised boosted models for all South Africa, and (C) predicted richness using generalised boosted models for the different biomes.



These results are consistent with a study by Cowling et al. (1997b) who modelled regional-scale richness using data from variously sized areas. The fact that both analyses are concordant using completely different data sets (for both the response and the explanatory variables) and analytical techniques, highlight the robustness of our findings.

Test of hypotheses

Whole region

The pattern of plant species richness over South Africa was most strongly predicted by topographic heterogeneity followed by energy/favourableness variables. Surprisingly, this does not concur with studies of regional-scale patterns in other parts of the world. Thus, studies in California (Richerson and Lum 1980) and North America (trees only; Currie and Paquin 1987, Currie 1991) show that topographic heterogeneity plays a subsidiary role relative to favorableness (mean rainfall) and energy availability. Nor does it corroborate the results of O'Brien (O'Brien 1993, 1998), as we found no evidence of a curvilinear relationship between energy variables and plant richness across the region.

Why high productivity regions should support more plant species than low-productivity ones remains unclear. In South Africa, some of the richest areas fall within the low-energy, low-rainfall parts of the Fynbos and Succulent karroid biomes. Advocates of species-energy theory argue that high levels of available energy – or productivity – favour large numbers of individuals (Wright et al. 1993); however, the next logical step – that large numbers of individuals result in species-rich assemblages – is unconvincing (Brown 1988). Perhaps, as pointed out by Ricklefs (2004), productive environments are rich in species simply because they prevailed over large parts of the world throughout most of the Cenozoic.

Spatial heterogeneity is a surrogate for niche diversity, and it follows that – all else being equal – highly heterogeneous areas should support more species than areas of lower heterogeneity (Rosenzweig 1995). This ecological component of heterogeneity is supported by the fact that heterogeneity is the best predictor of invasive alien plant richness. However, heterogeneity is also important in an historical, evolutionary context, since speciation rates are likely to be higher, and extinction rates lower, in heterogeneous environments (Poethke et al. 2003, Konvicka et al. 2006). Thus, invoking heterogeneity as a determinant of richness requires a fusion of ecological and evolutionary perspectives on diversity (Cowling et al. 1997b, Ricklefs 2004). Our results show it as an important predictor and not a factor that should be used to account for the unexplained variance in deterministic, ecological models (Whittaker et al. 2001).

Given the antiquity and habitat diversity of the South African landscape (Partridge 1997), the extremely steep environmental gradients over much of its area (Schulze 1997), the relative absence of catastrophic change associated with Plio-Pleistocene climatic cycles (Scott et al. 1997), and the high number of habitat specialist, range-restricted species (Cowling and Hilton-Taylor 1992), it is likely that relatively recent diversification of certain lineages, resulting in the development of "species flocks" (Brooks 1950), is a major determinant of regional richness in the sub-continent. The high beta diversity of specks-rich regions such as the Fynbos biome (Cowling and Holmes 1992) and the Succulent karoo (Cowling et al. 1989), as well as the importance of spatial heterogeneity as a predictor of regional richness, support this contention. Furthermore, the relationship between available energy and Fynbos species richness is negative, showing that the low-energy sites in the Fynbos biome support more species than the high-energy sites. In short, our results are consistent with the hypothesis the massive richness of the escarpment and coastal foreland floras is a consequence of ongoing diversification and low extinction in an environment that became increasingly but gradually dissected and has remained relatively stable climatically since the early Miocene (Cowling and Proches 2005). Available phylogenetic data from the Fynbos and Succulent karroid biomes support this hypothesis (Linder 2003, Klak et al. 2004).

However, this historical explanation for the role of topographic heterogeneity does not account for the explanatory power of the model for alien invasive species richness. Because of this, we suggest topographic heterogeneity contributes to richness in two ways: firstly by providing an abundance of niches in space, and secondly providing an abundance of relatively stable niches in time. Our modelling approach would not be able to dissect the relative importance of these two mechanisms.

Biome-level analyses

Our results also show that different suites of variables emerged to predict species richness for different biomes. No single explanation for such variance is obvious, but it likely reflects different ecological and phylogenetic circumstances across biomes. Nevertheless, for several biomes, plant species richness is determined largely by environmental spatial heterogeneity and not by available energy. This is particularly true of the Fynbos and karroid biomes, where explosive diversification within certain lineages has resulted in the evolution of numerous habitat specialists and ecological equivalents, resulting in very high beta and gamma diversity (Cowling et al. 1989). Rainfall irregularity, as suggested by Cowling et al. (1997b), does not seem to be the major driver of plant species richness in the Fynbos biome.

probably because rainfall irregularity was modeled on an annual rather than a seasonal basis in our study.

In the tropical-derived Savanna and Grassland floras, total productivity emerged as an important predictor of richness, consistent with Cowling et al. (1997b). However, rather than providing support for the species-energy or species-favorableness theory, we concur with Cowling et al. (1997b) that this pattern probably reflects the environmental conditions (warm and moist) where these tropical lineages originated. Low productivity savanna and grassland environments are relatively young (Scott et al. 1997), and few tropical lineages have diversified there. The role of fire remains to be explored in this framework.

Conclusions

Owing to varied climates, topographies and floras of different phylogenetic affinity, South Africa provides an excellent opportunity for modelling plant richness at the regional scale. This study shows the overwhelming importance of topographic heterogeneity in driving richness patterns, a likely consequence of a flora that has diversified explosively in concert with increasing aridity, gradual topographical rejuvenation, and relative climatic stability (on the coastal forelands) since the early Miocene. However, energy does play an important role, especially in the subtropical Savanna and Grassland biomes, probably related to conditions suitable both now and in the past for a tropical-derived flora. Our analysis also shows that the species richness pattern in South Africa appears to be biome-specific with phylogenetic history or clade-specific tolerances of favourable ecological conditions explaining certain relationships. Overall, it seems plausible that topographic heterogeneity contributes to species richness in two ways, firstly by providing niches in space (as shown by richness patterns of the recently introduced alien flora), and secondly, by providing relatively stable niches over time that may promote speciation rates and retard extinction rates (as shown by richness of endemic species flocks).

Acknowledgements – We thank L. Henderson for supplying the SAPIA data and the South African National Biodiversity Inst. for the use of data from the National Herbarium (PRE) Computerized Information System (PRECIS). WT acknowledges support from the GDRI "France South Africa – Dynamics of biodiversity in Southern African ecosystems and sustainable use in the context of global change: processes and mechanisms involved". We finally thank Jane Elith, and particularly Jens-Christian Svenning for their insightful comments and inputs on the manuscript.

References

Breiman, L. et al. 1984. Classification and regression trees. – Chapman and Hall.

Brooks, J. L. 1950. Speciation in ancient lakes. – Quart. Rev. Biol. 25: 30–176.

Brown, J. H. 1988. Species diversity. – In: Myers, A. A. and Giller, P. S. (eds), Analytical biogeography. An integrated approach to the study of animal and plant distributions. Chapman and Hall, pp. 57–89.

Chown, S. L. et al. 2003. Energy, species richness, and human population size: conservation implications at a national scale. – Ecol. Appl. 13: 1233–1241.

Cowling, R. M. and Hilton-Taylor, C. 1992. Patterns of plant diversity and endemism is southern Africa: an overview. – Strelitzia 1: 31–52.

Cowling, R. M. and Holmes, P. M. 1992. Endemism and speciation in a low-land flora from the Cape Floristic Region. – Biol. J. Linn. Soc. 47: 367–383.

Cowling, R. M. and Lombard, A. T. 2002. Heterogeneity, speciation/extinction history and climate: explaining regional plant diversity patterns in the Cape Floristic Region. – Div. Distrib. 8: 163–179.

Cowling, R. M. and Proches, S. 2005. Patterns and evolution of plant diversity in the Cape Floristic Region. – Biol. Skrifter 55: 273–288.

Cowling, R. M. et al. 1989. Patterns of plant species diversity in southern Africa. – In: Huntley, B. J. (ed.), Biotic diversity in southern Africa. Concepts and conservation. Oxford Univ. Press, pp. 19–50.

Cowling, R. M. et al. 1994. Plant functional diversity, species diversity and climate in arid and semi-arid southern Africa. – J. Arid Environ. 27: 141–158.

Cowling, R. M. et al. 1997a. Vegetation of southern Africa. – Cambridge Univ. Press.

Cowling, R. M. et al. 1997b. Species diversity at the regional scale. – In: Cowling, R. M. et al. (eds), Vegetation of southern Africa. Cambridge Univ. Press, pp. 447–467.

Cowling, R. M. et al. 1998. Extraordinarily high regional-scale plant diversity in southern African arid lands: subcontinental and global comparisons. – Div. Distrib. 4: 27–36.

Cowling, R. M. et al. 2004. Climate stability in mediterranean-type ecosystems: implications for the evolution and conservation of biodiversity. – In: Arianoutsou, M. (ed.), Proc. 10' MEDECOS – International Conference on Ecology, Conservation and Management of Mediterranean Climate Ecosystems. Millpress, pp. 1–11.

Currie, D. J. 1991. Energy and large-scale patterns of animal- and plant-species richness. – Am. Nat. 137: 27–49.

Currie, D. J. and Paquin, V. 1987. Large-scale biogeographical patterns of species richness of trees. – Nature 329: 326–327.

Currie, D. J. et al. 2004. Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. – Ecol. Lett. 7: 1121–1134.

Dynesius, M. et al. 2004. Intercontinental similarities in riparian-plant diversity and sensitivity to river regulation. – Ecol. Appl. 14: 173–191.

Elith, J. et al. 2005. The evaluation strip: a new and robust method for plotting predicted responses from species distribution models. – Ecol. Modell. 186: 280–289.

Elith, J. et al. 2006. Novel methods improve predictions of species' distributions from occurrence data. – Ecography 29: 129–151.

Field, R. et al. 2005. Global models for predicting woody plant richness from climate: development and evaluation. – Ecology 86: 2263–2277.

Fjeldsa, J. and Lovett, J. C. 1997. Biodiversity and environmental stability. – Biodiv. Conserv. 6: 315–323.

Francis, A. P. and Currie, D. J. 1998. Global patterns of tree species richness in moist forests: another look. – Oikos 81: 598–602.

Friedman, J. H. 2001. Greedy function approximation: a gradient boosting machine. – Ann. Stat. 29: 1189–1232.

Friedman, J. H. et al. 2000. Additive logistic regression: a statistical view of boosting. – Ann. Stat. 28: 337–374.

Gaston, K. J. 2000. Global patterns of biodiversity. – Nature 405: 220–227.

Germishuizen, G. and Meyer, N. L. 2003. Plants of southern Africa: an annotated checklist. – Strelitzia 14: 1–1231.

Gibbs Russel, G. E. et al. 1984. Intensity of plant collecting in southern Africa. – Bothalia 15: 131–138.

Goldblatt, P. 1978. An analysis of the flora of southern Africa: its characteristics, relationships and origins. – Ann. Missouri Bot. Gard. 65: 369–436.

Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. – Ecol. Lett. 8: 993–1009.

Hastie, T. J. et al. 2001. The elements of statistical learning: data mining, inference, and prediction. – Springer.

Henderson, L. 1998. Southern African plant invaders atlas (SAPIA). – Appl. Plant Sci. 12: 31–32.

Henderson, L. 1999. The southern African plant invaders atlas (SAPIA) and its contribution to biological weed control. – Afr. Entomol. 1: 159–163.

Henderson, L. 2001. Alien weed and invasive plants. – Agricultural Research Council of South Africa.

Hickler, T. et al. 2004. Using a generalized vegetation model to simulate vegetation dynamics in northeastern USA. – Ecology 85: 519–530.

Hoffman, M. T. et al. 1994. Plant richness is negatively related to energy availability in semi-arid southern Africa. – Biodiv. Lett. 2: 35–38.

Hubbell, S. P. 2001. A unified neutral theory of biodiversity and biogeography. – Princeton Univ. Press.

Hutchinson, G. E. 1959. Homage to Santa Rosalia, or why are there so many kinds of animals? – Am. Nat. 93: 145–159.

Ihaca, R. and Gentleman, R. 1996. R: a language for data analysis and graphics. – J. Comp. Graph. Stat. 5: 299–314.

Klak, C. et al. 2004. Unmatched tempo of evolution in southern African semi-desert ice plants. – Nature 427: 63–65.

Konvicka, M. et al. 2006. Butterfly extinctions in European states: do socioeconomic conditions matter more than physical geography? – Global Ecol. Biogeogr. 15: 82–92.

Leathwick, J. R. et al. 2006. Variation in demersal fish species richness in the oceans surroundings New Zealand: an analysis using boosted regression trees. – Marl. Ecol. Prog. Ser., in press.

Linder, H. P. 1991. Environmental correlates of patterns of species richness in the south-western Cape Province of South Africa. – J. Biogeogr. 18: 509–518.

Linder, H. P. 2003. The radiation of the Cape flora. – Biol. Rev. 78: 597–638.

Lonsdale, W. M. 1999. Global patterns of plant invasions and the concept of invasibility. – Ecology 80: 1522–1536.

Matthews, W. S. et al. 1993. Endemic flora of the north-eastern Transvaal Escarpment South Africa. – Biol. Conserv. 63: 83–94.

Midgley, G. F. et al. 2001. Have Pleistocene climatic cycles influenced species richness patterns in the Greater Cape Mediterranean Region? – J. Mediterranean Ecol. 2: 137–144.

Mucina, L. and Rutherford, M. C. 2005. Vegetation map of South Africa, Lesotho and Swaziland. – National Botanical Inst.

O'Brien, E. M. 1993. Climatic gradients in woody plant species richness: toward an explanation based on an analysis of Southern Africa's woody flora. – J. Biogeogr. 20: 181–198.

O'Brien, E. M. 1998. Water-energy dynamics, climate, and prediction of woody plant species richness: an interim general model. – J. Biogeogr. 25: 379–398.

Partridge, T. C. 1997. Evolution of landscapes. – In: Cowling, R. M. et al. (eds), Vegetation of southern Africa. Cambridge Univ. Press, pp. 5–20.

Pausas, J. G. and Austin, M. P. 2001. Patterns of plant species richness in relation to different environments: an appraisal. – J. Veg. Sci. 12: 153–166.

Poethke, H. et al. 2003. Local extinction and the evolution of dispersal rates: causes and correlations. – Am. Nat. 161: 631–640.

Richardson, D. M. et al. 2000. Naturalization and invasion of alien plants: concepts and definitions. – Div. Distrib. 6: 93–107.

Richardson, D. M. et al. 2005. Species richness of alien plants in South Africa: environmental correlates and the relationship with native plant species richness. – Ecoscience 12: 391–402.

Richerson, P. J. and Lum, K. 1980. Patterns of plant species diversity in California: relation to weather and topography. – Am. Nat. 116: 504–536.

Ricklefs, R. E. 2000. The relationship between local and regional species richness in birds of the Caribbean basin. – J. Anim. Ecol. 69: 1111–1116.

Ricklefs, R. E. 2004. A comprehensive framework for global patterns in biodiversity. – Ecol. Lett. 7: 1–15.

Ricklefs, R. E. et al. 2004. The region effect on mesoscale plant species richness between eastern Asia and eastern North America. – Ecography 27: 129–136.

Ridgeway, G. 1999. The state of boosting. – Comput. Sci. Stat. 31: 172–181.

Rosenzweig, M. L. 1995. Species diversity in space and time. – Cambridge Univ. Press.

Rouget, M. and Richardson, D. M. 2003. Inferring process from pattern in plant invasions: a semimechanistic model incorporating propagule pressure and environmental factors. – Am. Nat. 162: 713–724.

Rouget, M. et al. 2004. Mapping the potential ranges of major plant invaders in South Africa, Lesotho and Swaziland using climatic suitability. – Div. Distrib. 10: 475–484.

Rutherford, M. C. and Westfall, R. H. 1994. Biomes of southern Africa f an objective categorization. – Mem. Bot. Surv. South Africs 63: 1–94.

Schulze, R. E. 1997. South African atlas of Agrohydrology and Climatology. – Water Research Commission.

Scott, L. et al. 1997. Vegetation history. – In: Cowling, R. M. et al. (eds), Vegetation of southern Africa. Cambridge Univ. Press, pp. 62–84.

Stohlgren, T. J. et al. 2003. The rich get richer: patterns of plant invasions in the United States. – Front. Ecol. Environ. 1: 11–14.

Storch, D. et al. 2005. The species-area-energy relationship. – Ecol. Lett. 8: 487–492.

Thuiller, W. 2004. Patterns and uncertainties of species' range shifts under climate change. – Global Change Biol. 10: 2020–2027.

Thuiller, W. et al. 2003. Large-scale environmental correlates of forest tree distributions in Catalonia (NE Spain). – Global Ecol. Biogeogr. 12: 313–325.

Thuiller, W. et al. 2005. Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. – Global Change Biol. 11: 2234–2250.

Whittaker, R. J. 1999. Scaling, energetics and diversity. – Nature 401: 865–866.

Whittaker, R. J. et al. 2001. Scale and species richness: towards a general, hierarchical theory of species diversity. – J. Biogeogr. 28: 453–470.

Woodward, F. I. and Lomas, M. R. 2004. Vegetation dynamics – simulating responses to climatic change. – Biol. Rev. 79: 643–670.

Wright, D. H. et al. 1993. Energy supply and patterns of species richness on local and regional scales. – In: Ricklefs, R. E. and Schluter, D. (eds), Species diversity in ecological communities: historical and geographical perspectives. Chicago Univ. Press, pp. 66–74.

Subject Editor: Jens-Christian Svenning.