

Title

TODO

Running title

TODO

Authors

Ruan van Mazijk, Michael D. Cramer, and G. Anthony Verboom

Department of Biological Sciences, University of Cape Town, Rondebosch, Cape Town, 7701

ORCID IDs:

- TODO
- TODO
- TODO

Acknowledgements

TODO

Abstract

Aim

To quantify the explanatory power of heterogeneity in predicting plant species richness and turnover here in the Greater Cape Floristic Region and in the Southwest Australia Floristic Region. We compare the environmental heterogeneity in each region, how species richness and turnover interact in each region to produce the observed patterns of richness, and what different forms of environmental heterogeneity better predict richness in each region.

We expect the Cape to be more heterogeneous in most environmental axes, and at a finer grain, such that the consequent high levels of species turnover explain the Cape's greater species richness per unit

area. We also conjecture that edaphic heterogeneity will be an important factor in predicting richness in SW Australia.

Location

The Greater Cape Floristic Region (GCFR) in southwest Africa, and the Southwest Australian Floristic Region (SWAFR)

Taxon

Vascular plants

Methods

[Geospatial data, floral occurrence data → regression analyses, non-parametric statistics]

Results

TODO

Main conclusions

TODO

Keywords

TODO

Introduction

Biodiversity represents the variety of living things, and the variety of ecological and evolutionary processes responsible for it [Bohn2004]. Studying the distribution of biodiversity in space is a major avenue of biological research [Kerr2001; Kreft2007]. Regional-scale geographic patterns in species richness have long been studied, particularly in biodiversity hotspots [Cook2015]. The spatial distribution of species richness can be explained in terms of the physical environment. Properties of the environment have been suggested to

influence species richness in three ways: (i) productivity, water, and energy to enable organismal growth, and resources (i.e. niche space) to support a wider range of species [Kreft2007; Gaston2000; Mouchet2015]; (ii) stability, which enables species’ persistence; and (iii) heterogeneity, which enables ecological speciation and possible barriers to gene flow, and with a wider variety of environments to facilitate species’ co-existence [Thuiller2006; Mouchet2015; Cramer2016]. The physical environment, then, can be used to explain species richness in a local-deterministic sense, and in a historical context [Ricklefs1987].

The maintenance of species richness, particularly the coexistence of high numbers of species in biodiversity hotspots, is often regarded as “paradoxical” [Hart2017], and is a central problem in ecology [Hart2017; Kreft2007; Ricklefs1987]. Species richness is constrained by the ability of habitats to support a variety of species—its ecological carrying capacity [Mateo2017]. This is exemplified in approaches to modelling species richness as a function of environmental predictors in a correlative framework [“macro-ecological models”; Mateo2017]. Macro-ecological models of species richness implicitly assume that communities are saturated, following species-area and species-energy relationships, and at equilibrium with the environment [Mateo2017].

A solution to the paradox of species coexistence is environmental heterogeneity: a more heterogeneous environment has a larger environmental space, and can thus facilitate co-existence between species at the scale of that heterogeneity. Environmental heterogeneity can also stimulate ecological speciation, should the region be environmentally stable over evolutionary time-scales. Heterogeneity in the physical environment is known to be positively associated with species richness [VanRensburg2002; Hart2017], and has been demonstrated to do so across many taxa—e.g. Canadian butterflies [Kerr2001], European vertebrates [Mouchet2015], South African birds [VanRensburg2002], in communities along marine continental margins [Levin2010], French scarab beetles [Lobo2004], and for global terrestrial plants [Kreft2007]. The spatial scale of heterogeneity, or “grain” of the environment, is important to consider [Hart2017], in the same way that the spatial of absolute environmental conditions has also been considered [Baudena2015; Mouchet2015; Kerr2001]. Species co-existence and biodiversity maintenance is indeed suggested to be scale-dependent [Hart2017].

Environmental heterogeneity is often under-represented in macro-ecological models of species richness, and has recently been found to explain up to ca. 95% of biome level species richness across South Africa [Cramer2016]. Models that include environmental heterogeneity yield better estimates of the richness of the Cape flora, as they account for the role heterogeneous environments such as those in the Cape facilitate species coexistence [Cramer2016; Thuiller2006]. Mediterranean-type terrestrial biodiversity hotspots, such as the Cape flora included in the models by Cramer2016, present interesting study systems in which to investigate the relationship between the environment and species richness. These systems exhibit far greater species richness than predicted by their areas, productivities and latitudes [Kreft2007; Cowling1996]. There are five Mediterranean biodiversity hotspots on Earth: the California Floristic Province, the Mediterranean

Basin, the Chilean Winter Rainfall-Valdivian Forests, the Greater Cape Floristic Region, and the Southwest Australia Floristic Region [Cowling1996; Hopper2004; Cook2015]. These ecosystems have regular fire-cycles [Cowling1996], climatic buffering, and long term stability [Kreft2007], shrubby, sclerophyllous flora [Hopper2004]. Together, they account for ca. 20% of global vascular plant species, yet only ca. 5% of global land surface areas [Cowling1996]. Various hypotheses have been proposed to explain the high levels of plant species richness in these regions [Cook2015]. The species accumulation hypothesis states that the stability of these regions has allowed many species to accrue. The species co-existence hypothesis states that these hotspots may facilitate greater degrees of species co-existence in smaller spatial areas, due to fine-scale heterogeneity in their environments. Indeed, environmental heterogeneity has evolutionary implications too, stimulating ecological speciation across sharp environmental gradients.

Both the Southwest Australia Floristic Region (SWAFR) and the Greater Cape Floristic Region (GCFR) are Mediterranean-type biodiversity hotspots, particularly in terms of plant species. Where the GCFR (with an area of ca. 189,000 km²) is known to contain about 11,400 plant species (about 0.060 species per km²), the SWAFR (area of ca. 270,000 km²) has about 3,700 species (0.014 species per km²) [ManningandGoldblatt2012]. So, the GCFR has ca. 4.3 times as many species per km² as the SWAFR. The GCFR and SWAFR are appropriately often compared, due to the similarities between their environments (e.g. oligotrophic soils, an oceanically buffered moderate climate) and their plants' ecologies [Hopper2004]. These two regions present unique flora out of the five Mediterranean systems, with high levels of endemism [Cowling1996], and many obligate fire-adapted species [Cowling1996]. Similarities withstanding, the SWAFR is topographically and edaphically distinct from the GCFR. The SWAFR former is topographically rather uniform (i.e. flat)—uniquely so among the world's five Mediterranean-climate regions [Hopper2004]. The SWAFR possesses a mesoscale chronosequence dune system [Cook2015; Laliberte2014], while the GCFR is mountainous, topographically heterogeneous, and therefore associated with a large degree of spatial climatic variability, with a fine-scale mosaic of geologies and soils [Verboom2017; Cramer2014; Cowling1996]. Both regions have sources of edaphic heterogeneity, but at different scales. This edaphic variability may aid in explaining the species richness in these regions [Beard2000; Verboom2017]. Environmental heterogeneity can stimulate ecological speciation, should the region be stable over evolutionary time-scales, as is likely the case in both the GCFR and the SWAFR [Hopper2004; Lambers2010; Wardell-Johnson1996; Cramer2014; Laliberte2014; Cook2015]. For the GCFR, this richness is largely known to result from long term climatic stability, and fine grain variation in geology and soils [Cramer2014]. The question thus arises whether heterogeneity is a significant contributor to SWAFR species richness as is likely the case in the GCFR. In the absence of topographic variability in the SWAFR, it is proposed that the heterogeneity of that region is due to the juxtaposition of soil types [Cook2015; Laliberte2014], creating extreme edaphic variation.

New hypothesis para

Aim

This study investigates the role environmental heterogeneity plays in explaining vascular plant species richness in the Cape and SWA. We compare the relative importance of heterogeneity between the two regions, as heterogeneity has the evolutionary role of facilitating speciation, and the ecological role of supporting diverse species assemblages. Spatial scale of that heterogeneity is also considered, as the heterogeneity-richness relationship can vary with habitat grain-size .

Hypothesis

v1

Our main hypothesis is that the greater abiotic heterogeneity in the Cape, and the finer grain of that heterogeneity, compared to that of the SWA, accounts for the Cape's greater species richness per unit area. We expect the relationships between environmental heterogeneity, species richness, and species turnover in these two regions to demonstrate this. As stated above, heterogeneous environments can (i) support diverse species assemblages, and (ii) stimulate ecological speciation. Thus, we expect species richness to covary with heterogeneity. Additionally, as one moves across a heterogeneous landscape, we expect to find greater turnover in community composition, as different environments support different species. Thus, areas of greater turnover should also be more rich, due to potential complementarity between neighbouring communities increasing total richness. Thus, we expect that environmental heterogeneity positively influences species richness and species turnover, and that species turnover itself positively influences species richness.

v2

Our hypotheses concern the Cape and SWA's environments and floras. Our main hypothesis is that the Cape possesses greater abiotic heterogeneity, and at finer grain, compared to SWA, such as to explain the Cape's greater species richness per unit area, and proposed greater levels of species turnover between areas. We also conjecture that the heterogeneity that predicts species richness in SWA will be more pronounced in terms of edaphic variables.

Predictions

Here we attempt to assess six key predictions of this hypothesis, additionally investigating a seventh prediction to test the conjectured role of edaphic heterogeneity in SWA.

Dealing with the two regions' environments, we assess whether:

- (i) the Cape environment is more heterogeneous than that of SWA;
- (ii) the Cape environment has finer grain heterogeneity than that of SWA.

Dealing with the distribution of species in the two regions, we assess whether:

- (iii) the Cape exhibits greater levels of species turnover between areas, such as to...
- (iv) ... explain the greater species richness per unit area of the Cape compared to SWA.

Relating each regions' environment and flora, we finally assess whether:

- (v) Species richness and species turnover are adequately predicted by environmental heterogeneity in both regions;
- (vii) Species richness and species turnover are better predicted by different forms of environmental heterogeneity in either region (e.g. the importance of edaphic heterogeneity in SWA).

Methods

We employ classical statistical methods to analyse publicly available geospatial and species occurrence datasets.

Results

Describing environmental heterogeneity numerically across scales

Discussion

I have provided support for the hypothesis that the difference in plant species richness between the GCFR and SWAFR is accounted for by the fact that the GCFR is more abiotically heterogeneous than the SWAFR. As expected, the GCFR is shown to possess (i) a quantifiably more heterogeneous environment, and (ii) is heterogeneous at a finer spatial scale than the SWAFR. I have shown that vascular plant species richness (iii) can be explained in terms of environmental conditions, including environmental heterogeneity, in both the GCFR and SWAFR. Also, I have shown that (iv) the set of environmental axes that explain plant species richness, both absolute and as heterogeneity, differs between the GCFR and SWAFR. These findings contribute towards an understanding of the ecological conditions that facilitate species coexistence (and likely stimulate ecological speciation) in these two regions.

These two regions present differentiable environmental spaces, each with heterogeneity varying across spatial scales. The clear separation of the regions' topographic features is as expected (Figures

??A, ??). Indeed, topography seems to be the most striking distinction between the regions. The Cape region has been found previously to have the second highest median topographic heterogeneity of the five Mediterranean-climate regions [Bradshaw2014]. The GCFR has a much wider range of scales exhibited in the heterogeneity across its environmental axes. Notably, each region has finer scale heterogeneity in some variables, and coarser scale in others—neither region is necessarily more fine or coarse than the other, as it depends on the variable concerned. BRT-models of species richness in both regions reveal species richness to depend on those environmental axes that differentiate the two regions (Figures ??, ??). The importance of variables is also shown to vary with spatial scale (Figure ??), as previously suggested may be the case when modelling geographic patterns of biodiversity [Baudena2015]. Indeed, as Cowling1996 describes differing patterns of species richness across spatial scales, so do the predictors of those patterns vary with scale [Hart2017].

The fact that a combination of absolute and roughness variables is also as predicted by the hypothesis in this study. In the models developed by Cramer2016 for South Africa, roughness in topography was largely superseded as an important predictor of species richness by other roughness variables. My models, however, did not show this. Similar to the study by VanRensburg2002, my models revealed roughness in topography and other variables to be important. Although, VanRensburg2002 considered differences within pixels, as opposed to this study, which considered differences between pixels. My models, those of Cramer2016, and those of VanRensburg2002, do not all concur as to the role of roughness in elevation vs. more biologically meaningful variables in explaining species richness. The source of these discrepancies is unclear, though no doubt complex. The complements of environmental variables and methodologies used in these studies do differ, limiting extensive comparison between these analyses.

The determinants of vascular plant species are shown to be region specific (Figures ??, ??, ??). The importance of MAP and roughness in rainfall seasonality (PCV) in predicting richness in the SWAFR (Figure ??I, ??J), aligns with the steep climatic gradients observed there [Cook2015]. The soil variables that determine plant species richness in the model for the SWAFR (Figures ??K, ??L) differ to those that determine richness in the GCFR (Figures ??G, ??H), further highlighting the edaphic differences between these two regions. Although both are nutrient leached systems, the SWAFR is flat, with soil-chronosequences [Cook2015; Laliberte2014], while the GCFR is mountainous [Verboom2017; Cramer2014; Cowling1996]. The importance of roughness in soil density, and absolute texture, in the SWAFR (Figures ??K, ??L) highlights the changes in soil that are associable with age of the substrate (e.g. particle size) as being biologically relevant to species richness. The positive effect of soil clay content on species richness in the SWAFR aligns with the findings of Laliberte2014 that richness in the SWAFR increases with soil age.

NDVI is more heterogeneous across the GCFR than the SWAFR (Figures ??A). The fact that thermal variables tend to be more rough in the GCFR (Figure ??A) is likely due to possible covariance of

the MODIS/Terra products with topography, as MODIS data used here describes land surface temperature. As the GCFR is topographically rugged, the roughness of NDVI may arise from this. Despite this, NDVI is an integrating variable, which captures information about productivity, light availability, and soil nutrients [Power2017]. The fact that absolute NDVI contributes to predicting species richness in the GCFR, especially at finer spatial scales (Figure ??E) demonstrates the role of ecological productivity in facilitating the coexistence diverse species assemblages. Environmental heterogeneity, then, is integral to explaining patterns of species richness, but must be considered along with resource- and energy-availability axes. In so much as a diverse environmental space supports more species, the materials and productivity required for biota to thrive are also needed to support species [Bohn2004; Kerr2001; Kreft2007; Gaston2000]. As such, my findings, along with those of previous studies [Cramer2016; VanRensburg2002; Thuiller2006; Kreft2007], suggest that there is ecological and evolutionary consequence to resource availability *and* environmental heterogeneity, in that they tend to be positively associated with species richness.

The combined BRT-model of species richness for both regions reveals soil clay content as an important predictor, at coarse spatial scales, despite this variable not being particularly important within each region separately (Figure ??). Though this model does not strictly consider the regions as separate, this finding may indicate that the relationship between clay content and species richness differs between the regions. So far as clay content can be used to predict species richness, it matters more to those predictions when applied to large sections (i.e. coarse scales) of each regions.

Kreft2007 modelled global terrestrial vascular plant species richness, which focussed on primarily absolute environmental values, underestimated the richness of the Cape flora. Though Kreft2007 did include topographic heterogeneity in their predictor set, topography is often a proxy for more biologically meaningful variables [Cramer2016]. This explains why the inclusion of these variables (e.g. roughness in mean annual precipitation) yields more accurate predictions of species richness. Indeed, Thuiller2006 also included topographic heterogeneity. Cramer2016 described 68% of species richness at the QDS scale across South Africa. Regarding the GCFR, depending on whether one consults pseudo- R^2 (Table 3), the ratio of mean predicted to observed richness per grid-cell (Table 5), or the distributions of predicted vs. observed richness values per grid-cell (Figure ??), I have achieved a similarly suitable level of predictive accuracy. There is, though, still unexplained species richness in light of my models. As Cramer2016, VanRensburg2002, Thuiller2006, and Mouchet2015 have done, these macro-ecological models are a-historical. Evolutionary considerations of species richness in geographic space are worthwhile, especially in regions with environments stable over evolutionary time.

The findings here are correlative. There are, however, many proposed mechanisms to explain the correlative signals demonstrated here. My findings support the hypothesis that Mediterranean systems' plant species richness is a function of spatial variability in environmental conditions. This can stimulate diversification, and maintain that diversity by providing a range of habitats for species co-existence. Oligotrophic

soils can stimulate an increase in functional diversity, through the evolution of diverse nutrient acquisition strategies [Lambers2010; Verboom2017] (e.g. sclerophylly [Cook2015; Cramer2014]). An aspect of the environment I have neglected to consider is fire, shown to also contribute to predictions here in the GCFR [Cramer2016]. Cardillo2012 have shown the structuring forces behind species co-occurrence patterns, and thus likely species richness, differ between species-pairs with different post-fire responses and those with similar post-fire responses.

Though the GCFR was correctly predicted to have, on average, more species per grid-cell at HDS and 3QDS scales than the SWAFR, this was not the case for QDS grid-cells (Table 5). This demonstrates that the GCFR is indeed overall more rich in plant species than the SWAFR, but a given HDS in the SWAFR contains fewer species than a given GCFR HDS. Thus, the greater richness in the GCFR is a product of greater turnover in species at spatial scales no more coarse than the HDS. Species turnover is an interesting aspect to species richness studies, as it species turnover is implicit to species-area and co-existence-area relationships [Hart2017]. One could expect patterns of endemism and species turnover to concur with patterns in environmental heterogeneity to some degree.

Following from the understanding that functionally diverse assemblages, which are more likely to be more species rich, are likely to arise and/or occur in areas with diverse ecological pressures [Molina-Venegas2015], one would expect, then, heterogeneous habitats such as those in Mediterranean-type biodiversity hotspots to exhibit high levels functional beta diversity along steep environmental gradients [Molina-Venegas2015]. If the niches concerning these functions are phylogenetically conserved among those biota, then one would also expect high levels of species and phylogenetic beta diversity along these gradients [Molina-Venegas2015]. This concurs with the notion put forward by Power2017, wherein megadiverse systems such as these represent the results of “phylogenetic niche conservatism on a heterogeneous landscape”. Thus, species and phylogenetic turnover should covary with environmental heterogeneity in some way. Indeed, endemism, at certain scales, could also follow this pattern. Thuiller2006 demonstrated that there is phylogenetic and biome related determinants of species richness. This makes sense, in light of the difficulty of crossing biome boundaries in Mediterranean systems [Power2017]. NDVI and light availability, and the heterogeneity therein, are associated with high levels of floristic turnover [Power2017]. This may be indicative of ecological specialisation precluding species from crossing these boundaries, thus increasing the level of endemism within a region, while also increasing the level of turnover, and thus likely species richness, along environmental gradients. Although, this may be debated. Beard2000 state that the high levels of endemism in SWAFR are function of habitat specialisation to soil mosaics. Cf. Laliberte2014, who say that this endemism is likely due to environmental filtering along these soil turnover sequences, as opposed to the juxtaposition of specialised species along soil gradients.

I have demonstrated support for the idea that environmental heterogeneity is positively associated with species richness, particularly Mediterranean systems. In the SWAFR and the GCFR, high levels

of endemism and biodiversity are also likely the results of long-term landscape and climatic stability [Hopper1979]. Thus, the roles of environmental variability through space, and stability through time, are the two main ways in which the environment relates to biodiversity in these regions.

Tables

Table 1: Data sources used in this study. Abbreviations are as follows: ... **TODO**

Variable	Source (Dataset code)	Citation
Elevation	SRTM (<code>)	@Farr2007
Climate & vegetation variables		
NDVI	MODIS (<product code>)	<ref>
Surface T	MODIS (<product code>)	<ref>
MAP	CHIRPS (<product code>)	<ref>
PDQ	CHIRPS (<product code>)	<ref>
Soil variables (all from SoilGrids250m)		
pH	<code>	@Hengl2017
CEC	<code>	
Soil C	<code>	
Clay	<code>	
Plant species occurrence	GBIF	<ref>

Table 2: Estimated regression coefficients and significances (P-value) following multiple linear regression of HDS richness (S_{HDS}) against the mean QDS richness ($\overline{S_{QDS}}$) and turnover ($\overline{J_{QDS}}$; Jaccards distance) within a given HDS, of the form $S_{HDS} \sim \overline{S_{QDS}} + \overline{J_{QDS}} + SWAFR + \overline{S_{QDS}} \times SWAFR + \overline{J_{QDS}} \times SWAFR$. The GCFR was fit as the baseline, hence the SWAFR represents the categorical term here. This was model was better fitting than a similar model without a region category ($\Delta AIC = 90.56$). Note, this model does not represent those curves plot in Figure 4 (there, the curves are from simple linear regressions of the variables in each panel, separated by region).

Term	Estimate	P-value
Intercept	-3062.297	< 0.001
$\log(\overline{S_{QDS}} + 1)$	595.574	< 0.001
$\overline{J_{QDS}}$	338.589	0.534

Term	Estimate	P-value
SWAFR	2905.604	0.001
$\log(\overline{S_{QDS}} + 1) \times \text{SWAFR}$	-333.908	< 0.001
$\overline{J_{QDS}} \times \text{SWAFR}$	-1246.109	0.073

Figures

References

Biosketches

Appendices

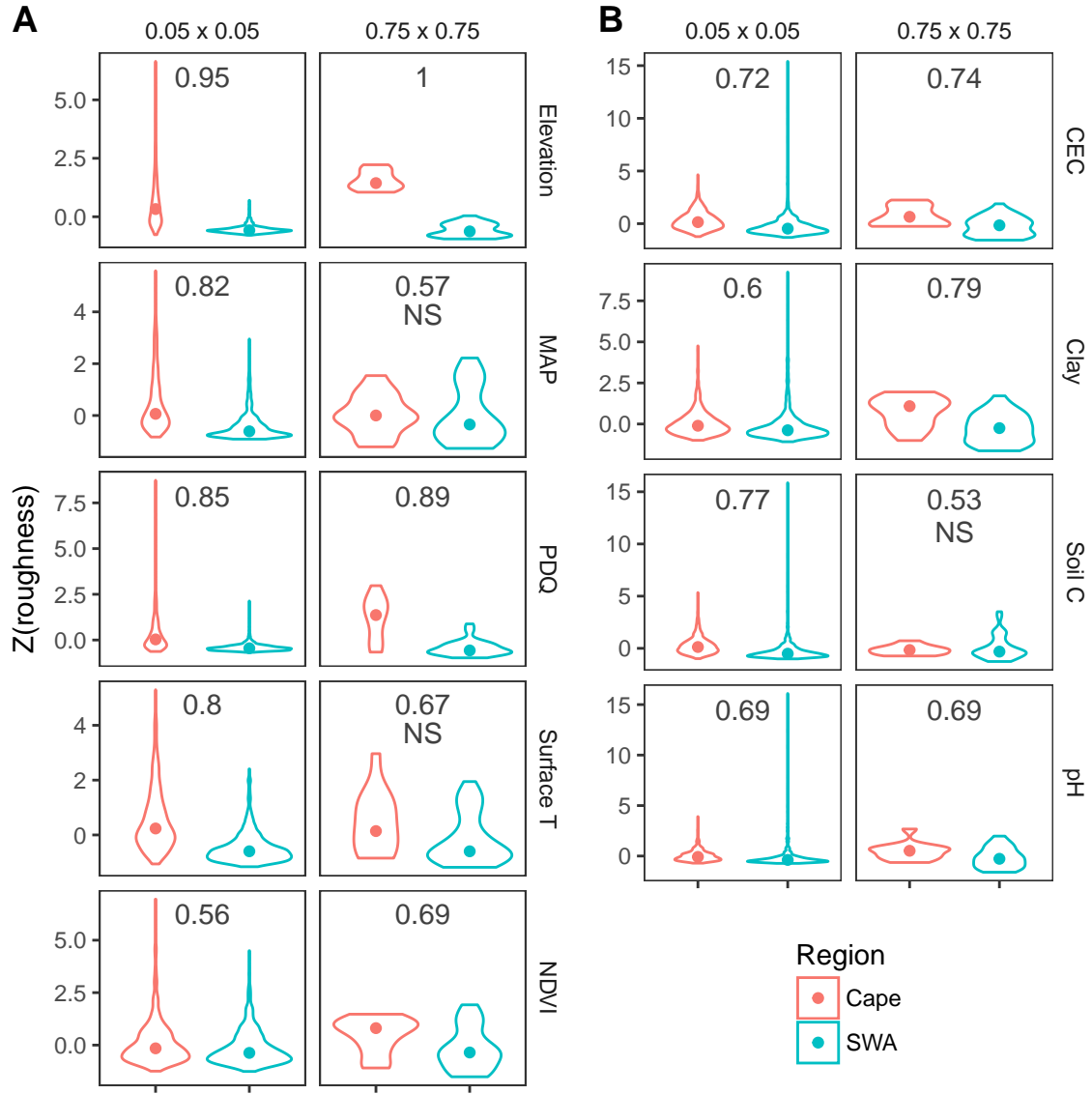


Figure 1: Violin-plots of the distribution of roughness values (Z-scaled) for nine environmental variables in the GCFR and SWAFR, presented at 0.05x and 0.75x (= 3QDS) square cell scales. Roughness values were calculated following Equation ???. Points represent the median of each distribution. Violin-plots were constructed with Gaussian distributions on data, with bandwidth following Silverman's 'rule of thumb' (??). Roughness values were found to be generally greater in the GCFR in all cases, unless marked as non-significant (NS), following Mann-Whitney U tests ($P < 0.05$).

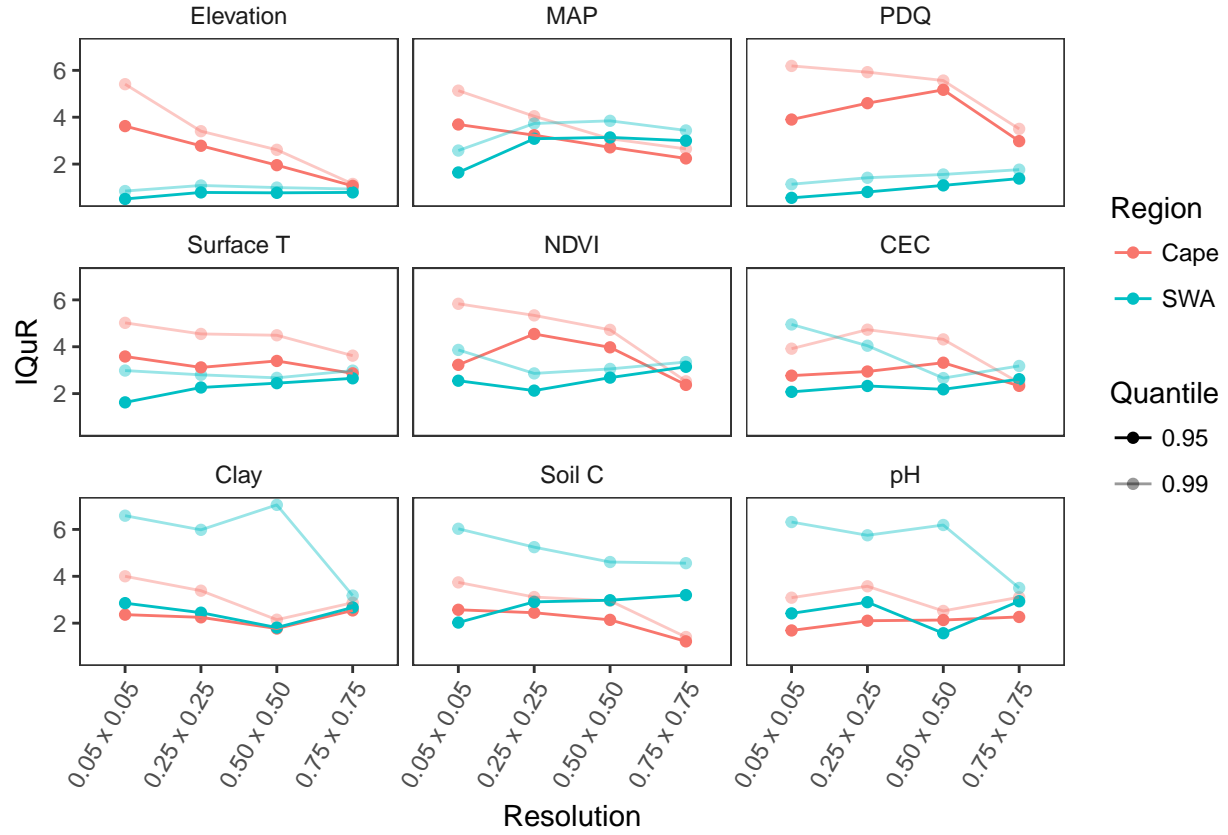


Figure 2: Scatter-plots of the size of the 95%-to-5% and 99%-to-1% interquantile ranges ($IQ_{0.95}R$ and $IQ_{0.99}$ respectively, as keyed) of Z-scaled roughness values (Equation ??) against spatial resolution (cells as $0.05\tilde{r}$ to $0.75\tilde{r}$ squares ($=3QDS$)), for the GCFR (red) and SWAFR (blue). Lines connect points for illustration.

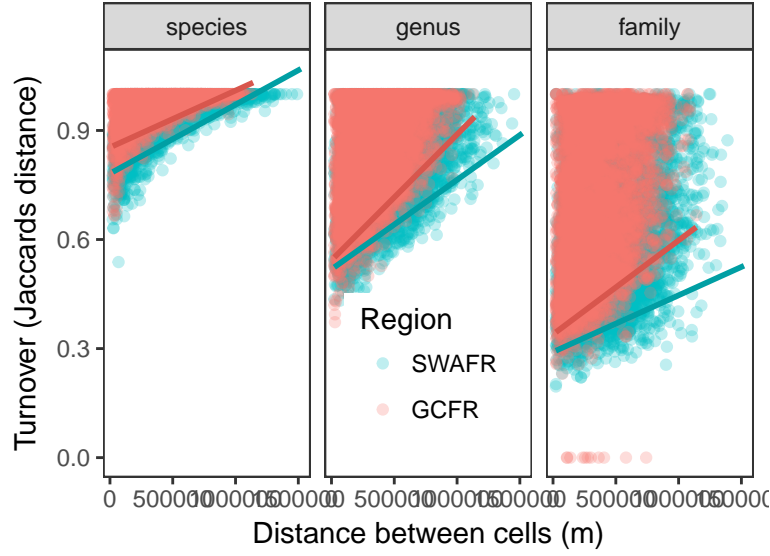


Figure 3: Scatter-plot species turnover (as Jaccards distance) between cells against the geographic distance (m) between them, from both regions (coloured as keyed). Species turnover was calculated for all possible pairs of cells, but only the turnover values for a random 5000 pairs in each region have been plotted, for clarity. Fitted lines represent the 5%-quantile regressions of turnover as a function of log-distance for each region. Following a single 5%-quantile regression of turnover as a function of log-distance with region as a categorical variable, a significant interaction between distance and region was found ($P < 0.05$), such that the SWAFR negatively impacted the distance slope term.

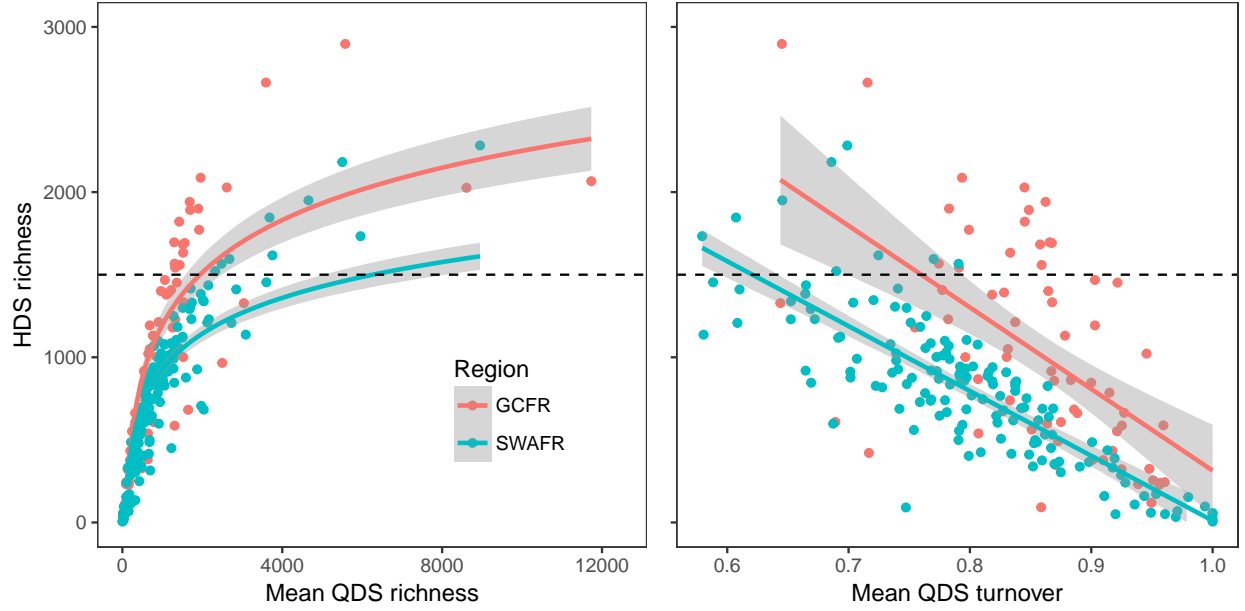


Figure 4: Scatter-plots of HDS-scale species richness against the average QDS-scale richness in a given HDS (left), and the average species turnover (Jaccards distance) between QDS in a given HDS (right). Curves represent simple linear regressions of HDS richness against the two respective independent variables (note, mean QDS richness was $\log(x + 1)$ -transformed), separately for each region, for illustration of the two regions' differences. The dashed horizontal line indicates HDS richness of 1500 species [discussed in text].