Bio3018F Practical 2022

Biodiversity and Ecosystem Function in the Cape Floristic Region

Jasper Slingsby

Introduction

Our understanding of and approach to studying the relationship between biodiversity and ecosystem function (B-EF) has evolved over the past few decades (Figure 1; van der Plas 2019). The historical view was one of biodiversity as the response variable, being determined by environmental and anthropogenic factors, with little feedback to ecosystem function. In the early 1990s, this shifted (and perhaps overcompensated) to focus on the causal effects of variation in biodiversity on ecosystem functioning, with little emphasis on the role of environmental variation. More recently, there is recognition that biodiversity both responds to the environment and partly drives ecosystem function in concert with environmental variation. The current focus of most B-EF research is on the relative importance of abiotic drivers (natural and anthropogenic) versus biotic variation in determining various ecosystem functions.

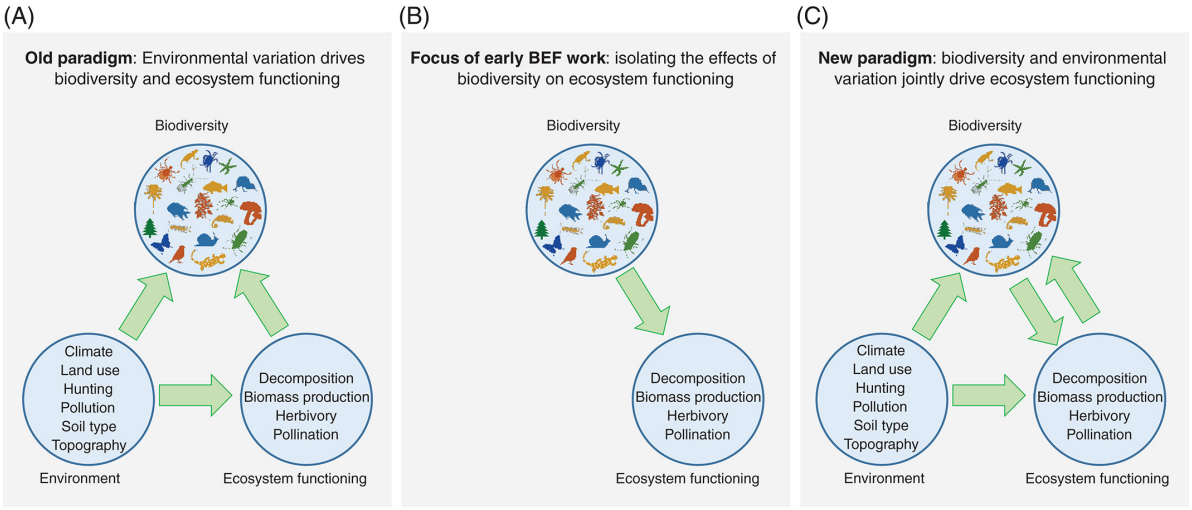


Figure 1: (from van der Plas 2019)

The Cape Floristic Region (CFR) of South Africa is one of the most botanically diverse areas on the planet. The indigenous flora of the CFR has several components with different evolutionary and biogeographic origins (Figure 2; Bergh et al. 2014), and distinct differences in a range of ecosystem functions. The CFR also has a long history of global change impacts, from direct anthropogenic disturbance (e.g. land use / land cover) to the introduction of invasive alien species. This provides a range of highly varied ecosystem types within close proximity, that often share or contrast in their biotic composition (species, functional and phylogenetic diversity) and abiotic conditions - climate, soils and disturbance regimes (fire, herbivory).

In this practical, we will explore how variation in a set of measures of biodiversity (species, functional and phylogenetic α and β diversity) and environmental conditions relate to a set of metrics of ecosystem function derived from satellite time-series.

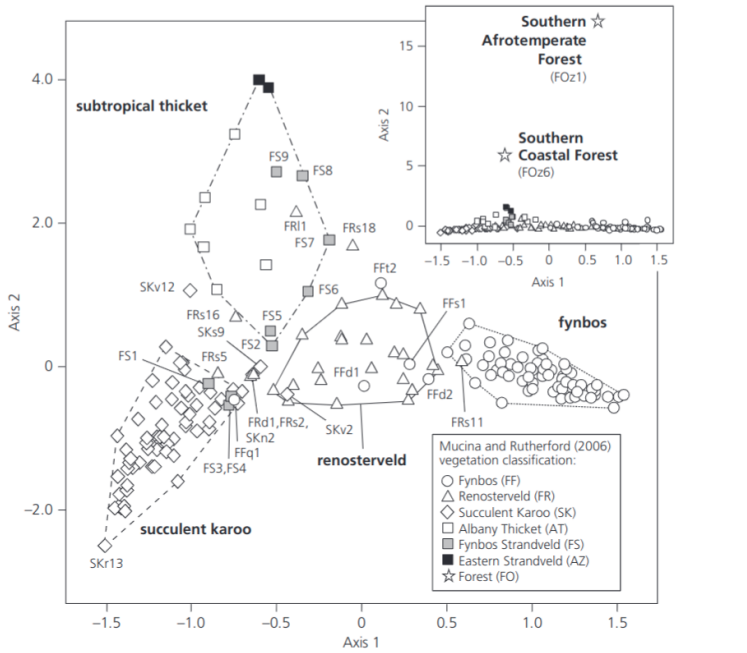


Figure 2: Ordination of genus-level floristic similarities of GCFR vegetation units sensu Mucina and Rutherford (2006), as inferred on the basis of the 'important species' lists provided in The vegetation of South Africa, Lesotho and Swaziland (Mucina and Rutherford 2006). Figure from Bergh et al. 2014.

The broad questions are:

- Does the variation in environmental conditions potentially explain the observed variation in biodiversity (species and functional α and β diversity)?
- To what degree does the variation in species and functional diversity potentially explain the observed variation in our measures of ecosystem function?
- What is the relative role of the environment versus biodiversity in determining the observed variation in ecosystem function?
- Finally, this practical is a largely descriptive observational study. In your discussion, describe a follow-up study that you would perform to discern cause from correlation and partition the relative influence of environmental conditions versus biodiversity on our measures of ecosystem function.

Methods:

The sites are selected to represent contrasting vegetation, but to all be the same (or similar) post-fire age (time since last fire). We will split into 4 teams of 2 or 3. Each pair will survey a point location (towards the corners) within each of six 250 by 250m MODIS satellite mission pixels that make up our sites.

Each team will need:

- One or more smartphones
- 2 x 10m measuring tape (or longer)
- 1 x clipboard, paper and pen or pencil
- 1 x densiometer
- 1 x metre rule

- 1 x Vernier calipers
- 8 x large plastic bags for carrying soil and plant samples for each point location
- 8 x soil sample bags
- 2 x masking tape
- 2 x marker pen

At each site, navigate to your team's point location, lay out the two tapes at 90 degrees, crossing at 3.2m and ending at 6.4m. Consider this your guide for a 3.2m radius (~20m2) circle. Within the circle we will measure:

1. *Environmental conditions:*

- Estimate % projected cover (think the area you would see from above) bare soil
- Take a densiometer reading at ground level. To do this, hold the densiometer level on the soil surface. Pick a spot close to the centre, but try not to pick an obviously open (or closed) patch, it should be representative of the location. To take a reading, count up the number of squares where no vegetation is visible (i.e. sky only) to a maximum of 24 squares, and write this down. We will convert this to canopy cover later, applying the formula $100 - 4.16 * X$, where X is your reading.
- Take a **soil sample** (as per Mike's prac, but it can be smaller as we are going to pool the 4 samples per site). These we will process for soil colour and pH as per Mike's prac.
- Do a **dung count**, scoring dung for the point location by the number of quarters where you find dung (i.e. a single score from 0-4 for the point location). Try to avoid scoring obvious single scat ("dung creation") events in more than one quadrat, unless it's an impressive pile.
- Take a few notes (and photos) on any other features that you think may be important or interesting (slope, rockiness, whatever) to assist you in the write-up.

2. *Diversity sampling:*

- First, make sure your smartphone is fully charged and set to record GPS location with your photographs!
- With your team, decide on which are the top 5 species by % projected cover. If your site is dominated by fewer than 5 species, count up as many species as make up 80% of the vegetation.
  - Take diagnostic photos for the 5 species (habit (whole plant), leaf, base, flowers and/or fruits).
  - For each of three individuals of your 5 target species, measure the height and collect a shoots for measuring leaf length and leaf width (mark with masking tape and put in sample bag labelled for the point location and site).
- Finally, set a timer and take as many photos of new species (other than your 5 target species) within or near your plot as you can before the alarm goes off. Make sure to take a photo of your site label on your sample bag between sites so you know which photos were collected at which sites. When we get back to base you will upload the photos to folders in the intranet labeled by site and identify them as far as possible (using morphospecies if needed).

Analyses

Species Alpha Diversity

For this I have just taken the count of species encountered at each point during our rapid photographic surveys, and the aggregated set of unique species for each site (i.e. no rarefaction etc).

> *Discussion hint: Do you think this is an issue for the method we used? Are there any biases we may have introduced? Justify your answer.*

Sites

Site	Species Number
grass	13
invasion	13
sandstone	24

Point locations

PointName	Species Number
grass_NE	17
grass_NW	12
grass_SE	13
grass_SW	14
invasion_SE	13
renosterveld_NE	17
renosterveld_NW	3
renosterveld_SW	12
sandstone_SE	24

Species Beta Diversity

Here I calculate species beta diversity using Sorenson's coefficient.

Discussion hint: Why can't we do Bray-Curtis? What would we gain if we could?

Sites

	invasion	renosterveld	sandstone	grass
invasion				
renosterveld	1.000			
sandstone	0.946	0.885		
grass	0.957	0.645	0.966	

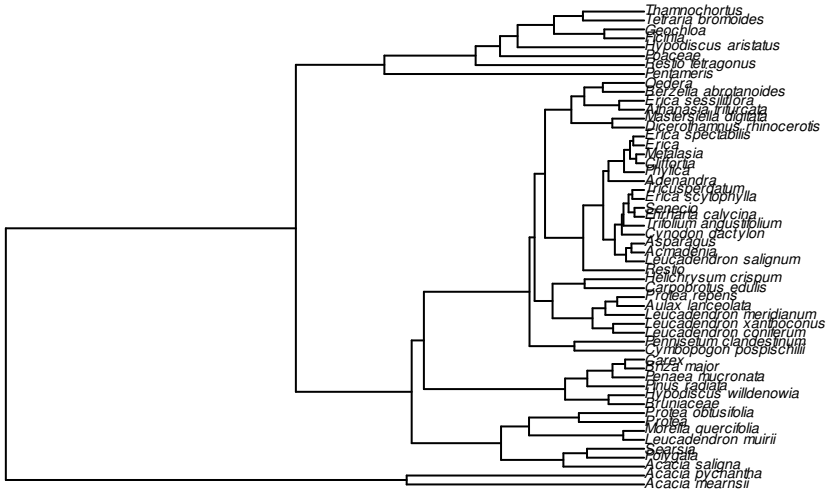
Point locations

	invasion_SE	renosterveld_SW	sandstone_SE	grass_NE	grass_NW	grass_SE	grass_SW	renosterveld_NE	renosterveld_NW
invasion_SE									
renosterveld_SW	1.000								
sandstone_SE	0.946	0.944							
grass_NE	0.933	0.931	0.951						
grass_NW	1.000	0.833	1.000	0.724					
grass_SE	0.923	0.920	1.000	0.600	0.680				
grass_SW	0.926	0.923	1.000	0.613	0.692	0.556			
renosterveld_NE	1.000	0.793	0.902	0.706	0.793	0.867	0.677		
renosterveld_NW	1.000	0.867	1.000	1.000	0.733	1.000	0.882	0.9	

Functional Alpha Diversity

Here I've estimated Functional Diversity (FD) according to the method of Petchey and Gaston (2002) for each of the points and aggregated sites using only the dominant species for which we measured traits. To

apply Petchey and Gaston's method, we first need to estimate a functional dendrogram representing the functional similarity among species based on the traits we measured.



FD by Site

	FD	SR
invasion	12.181	11
limestone	5.815	11
sandstone	6.964	9
grass	7.369	9
renosterveld	7.483	11
sand	6.876	9

FD by point location:

	FD	SR
grass_SE	3.200	2
renosterveld_SE	4.181	5
invasion_SE	7.422	4
sand_SE	4.650	5
sandstone_SE	5.231	5
limestone_SE	4.599	5
grass_NE	3.729	3
renosterveld_NE	4.756	4

	FD	SR
invasion_NE	5.856	5
sand_NE	4.136	4
sandstone_NE	4.344	3
limestone_NE	4.193	5
grass_SW	4.778	5
renosterveld_SW	5.410	5
invasion_SW	7.150	4
sand_SW	4.873	4
limestone_SW	4.599	5
sandstone_SW	4.344	3
grass_NW	4.680	3
renosterveld_NW	4.018	5
invasion_NW	6.413	3
sand_NW	2.713	1
sandstone_NW	5.840	5
limestone_NW	4.193	5

Functional Beta Diversity

I've also estimated functional turnover between sites using the method of Bryant et al. 2008. See help file `?picante::phylosor` in R for details to help know how to interpret. Note that while the function was written for phylogenetic turnover, we've used it for functional turnover.

	invasion	limestone	sandstone	grass	renosterveld	sand
invasion						
limestone	0.401					
sandstone	0.431	0.487				
grass	0.487	0.481	0.643			
renosterveld	0.378	0.412	0.560	0.603		
sand	0.460	0.504	0.800	0.647	0.564	

What about turnover between point locations?

Environmental Similarity among sites

First, let's calculate % canopy cover from the densiometer readings and then look at site-level averages for most variables.

```
## [1] "Site"            "Point"           "PercentBareSoil" "SoilPH"
## [5] "SoilColour"      "Dung"            "Densiometer"
```

Site	BareSoil	Soil pH	Dung	% Canopy Cover
grass	2.50	4.85	0.25	43.84
invasion	5.00	3.88	0.50	94.80
limestone	30.00	6.33	0.00	65.68
renosterveld	21.25	5.47	2.50	39.68
sand	12.00	4.81	0.50	72.96
sandstone	31.25	4.76	0.50	86.48

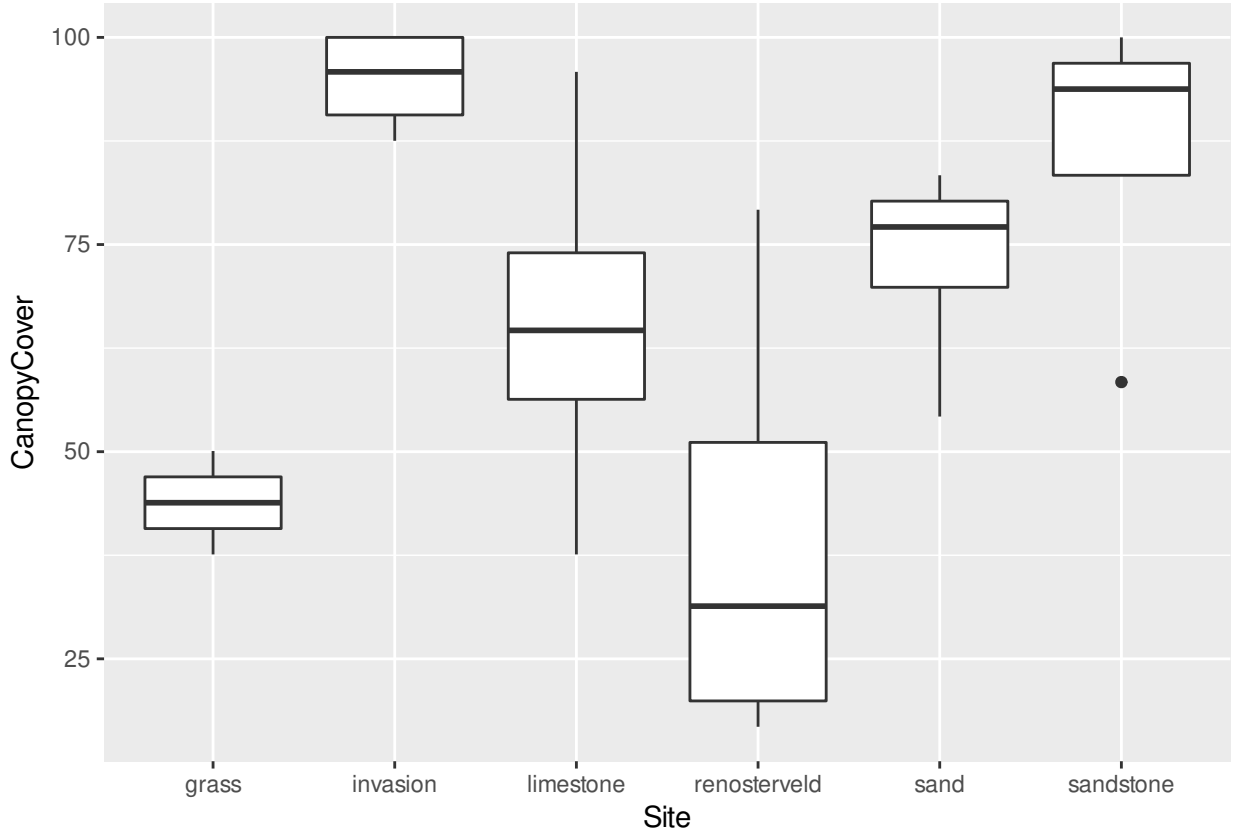
Or we can look at environmental similarity among sites

	grass	invasion	limestone	renosterveld	sand	sandstone
grass						
invasion	2.604					
limestone	3.034	3.876				
renosterveld	3.022	4.070	3.267			
sand	1.543	1.606	2.438	2.894		
sandstone	3.023	2.401	2.207	3.282	1.668	

Mike has shown you how to explore soil colour. Feel free to explore and add it in... You may also want to drop or explore subsets of variables to explore different aspects of environmental dissimilarity.

Plotting

Here's an example of how to make a boxplot of point location measures by site for one variable. You can make a panel of plots if you convert `pdat` into long format using `pivot_longer()` and adding `+ facet_wrap(~name, scales = "free")` to the end of the plotting call (where `name` is whatever you provided for the `names_to =` argument in `pivot_longer()`).



Test for significant difference among sites

Here we use the non-parametric Kruskal-Wallis rank sum test.

```
##
##  Kruskal-Wallis rank sum test
```

```
##
## data:  CanopyCover by Site
## Kruskal-Wallis chi-squared = 14.539, df = 5, p-value = 0.01252
```

And use Dunn's posthoc test to explore where the differences lie.

Comparison	Z	P.unadj	P.adj
grass - invasion	-2.881	0.004	0.030
grass - limestone	-1.152	0.249	0.340
invasion - limestone	1.728	0.084	0.252
grass - renosterveld	0.050	0.960	0.960
invasion - renosterveld	2.931	0.003	0.051
limestone - renosterveld	1.202	0.229	0.344
grass - sand	-1.428	0.153	0.288
invasion - sand	1.453	0.146	0.313
limestone - sand	-0.276	0.783	0.839
renosterveld - sand	-1.478	0.139	0.349
grass - sandstone	-2.405	0.016	0.061
invasion - sandstone	0.476	0.634	0.732
limestone - sandstone	-1.252	0.210	0.351
renosterveld - sandstone	-2.455	0.014	0.070
sand - sandstone	-0.977	0.329	0.411

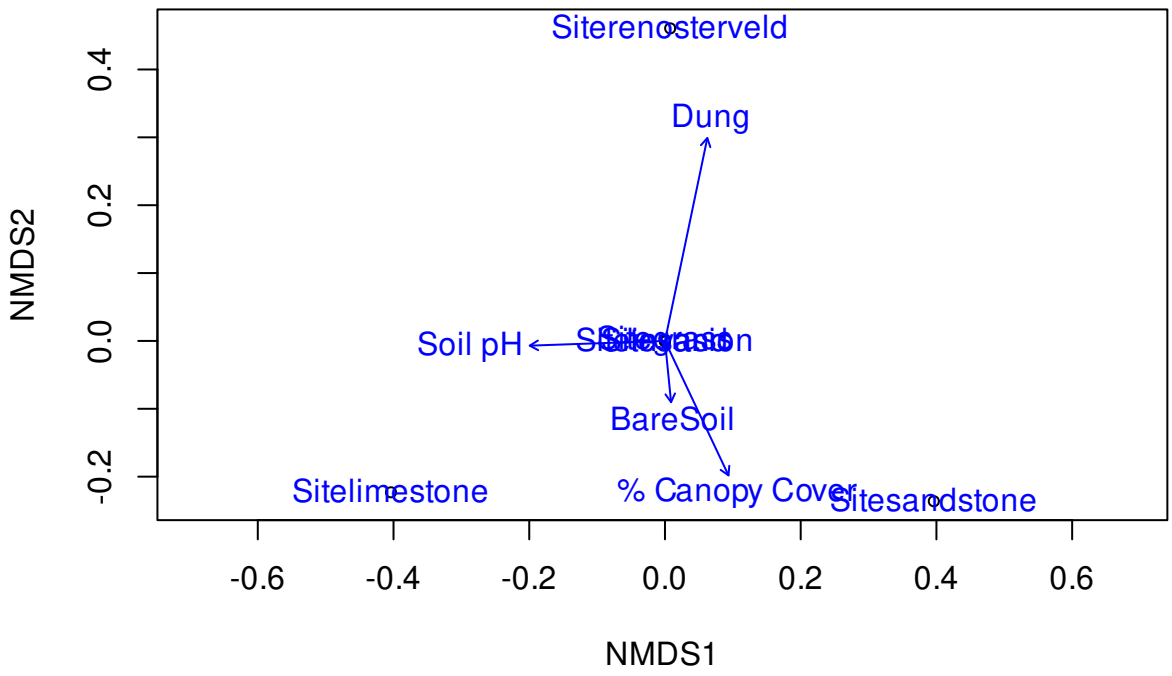
Ordination

You could also use ordination as you were taught in Timm's prac. I'll leave it to you to check assumptions etc as Hana taught you. Note you have different measures of biodiversity that you can apply at different levels (point location vs site).

Using species data at site level

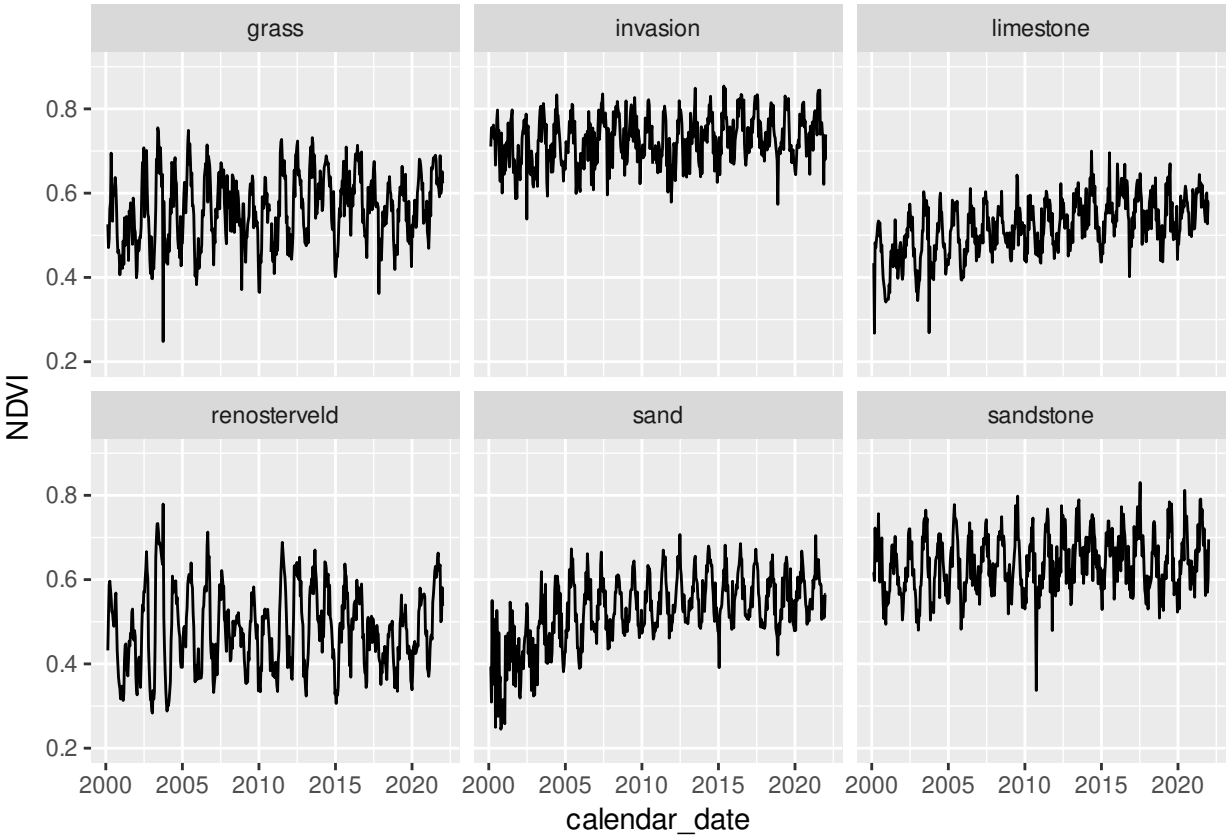
Using FD at site level

```
## Warning in metaMDS(msampBFD, trymax = 999, trace = FALSE): stress is (nearly)
## zero: you may have insufficient data
```

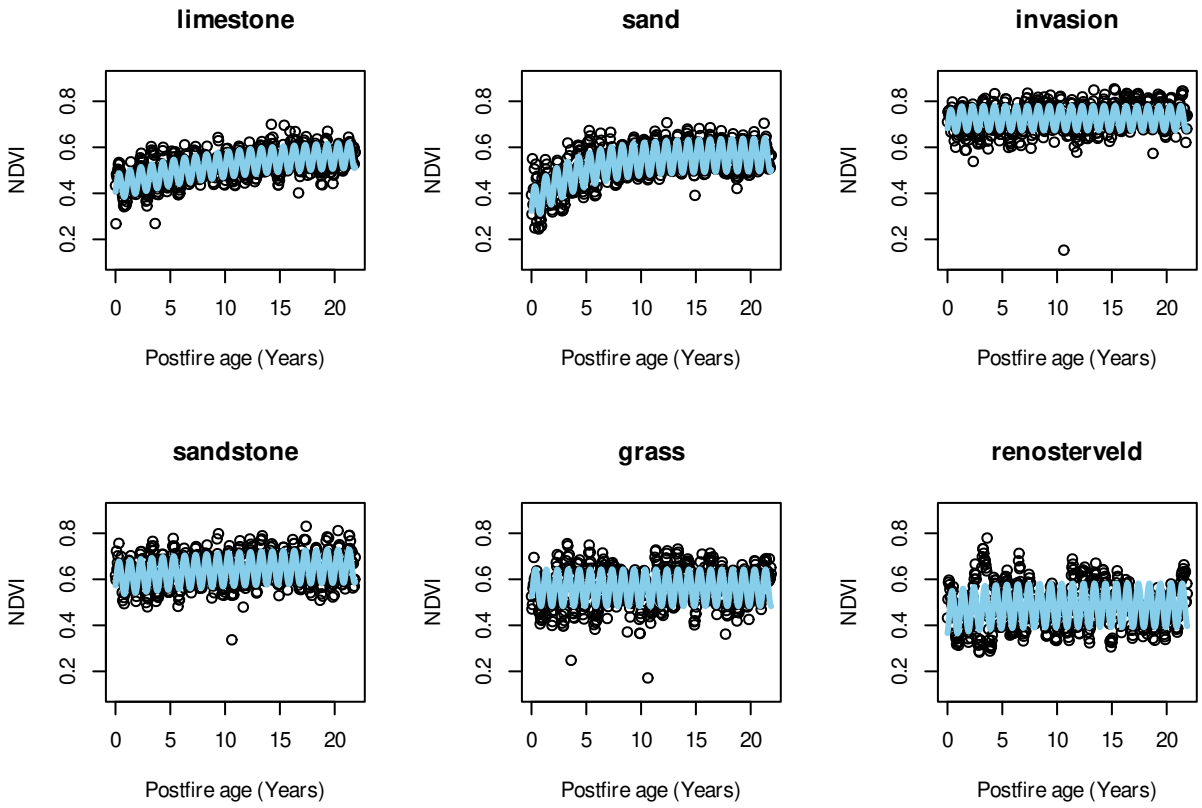


Ecosystem function

To explore ecosystem function we looked at the 20-year time-series of the Normalized Difference Vegetation Index (NDVI) recorded by the MODIS satellite mission. From these we used the post-fire recovery trajectory modelling framework developed by Wilson et al (2015) to derive estimates of the mean maximum NDVI (alpha + gamma), and the amplitude and timing of seasonality (big alpha and phi) as our measures of ecosystem function. Here I've provided a table of these parameters by site and plots of the model fits. Are the models good fits? Do they miss anything? Is it relevant to the questions we're asking?



Now let's fit the model and plot the model fit (blue) over the observed data (black circles).



And view a table of the parameter estimates

	alpha	gamma	lambda	A	phi
initial	0.200	0.400	0.500	0.100	-1.000
limestone	0.427	0.156	10.057	0.056	-1.441
sand	0.343	0.225	4.150	0.072	-1.326
invasion	0.695	0.031	0.459	0.054	-1.296
sandstone	0.603	0.062	10.539	0.073	-1.470
grass	0.579	-0.016	0.003	0.080	-1.762
renosterveld	0.449	0.038	2.670	0.095	-2.101

alpha = starting NDVI after a fire (lower stippled line above)

gamma = the difference between alpha and the asymptote of NDVI (i.e. alpha + gamma = estimated steady long-term NDVI (upper stippled line above))

lambda = 1/recovery rate after fire (i.e. the smaller the value the faster the rate of recovery)

A = the magnitude of seasonality (the absolute value indicates the magnitude, ignore the sign (+/-) for now)

phi = ignore

> *Discussion hint: Are there any obvious properties of the time-series that are not captured by these parameters? How do they differ between sites and what may the causes be?*

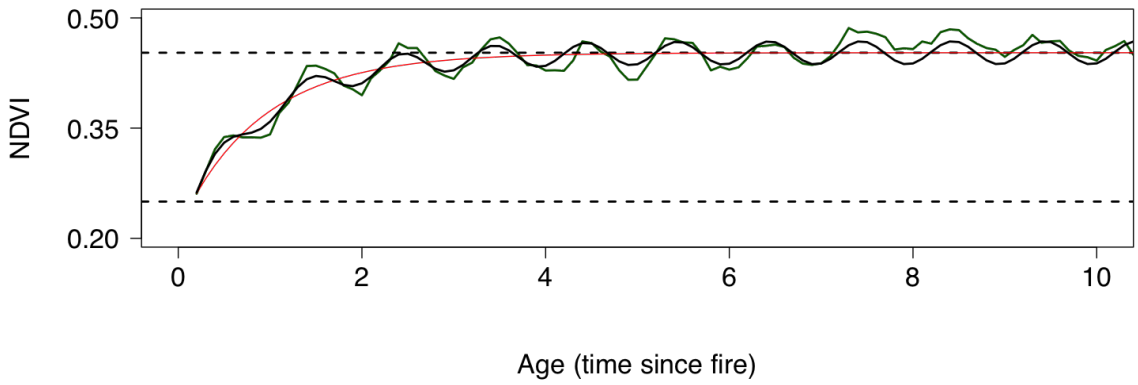
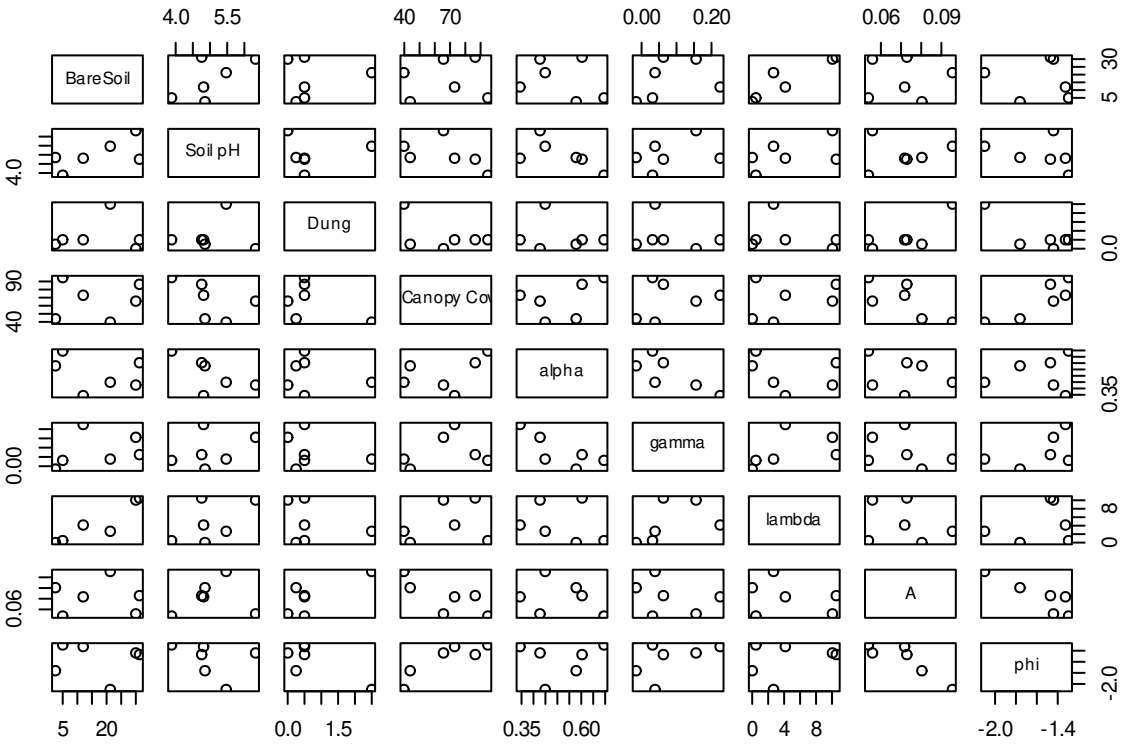


Figure 3: Stylized postfire recovery curve from Slingsby et al. (2020).

Explore relationships between EF and measures of biodiversity or environmental properties

This practical was a learning exercise and constrained by time and available resources. Perhaps the biggest drawback was the low number of (or lack of replication within) ecosystem types sampled. That said, plotting the relationships between the different site level variables gives us some indication of whether further sampling is likely to produce strong evidence of clear relationships. For this prac write up, I'll forgive you for discussing significant ($p<0.05$) or near-significant ($p<0.1$) relationships based on a sample size of 6... Usually, this would not be okay...



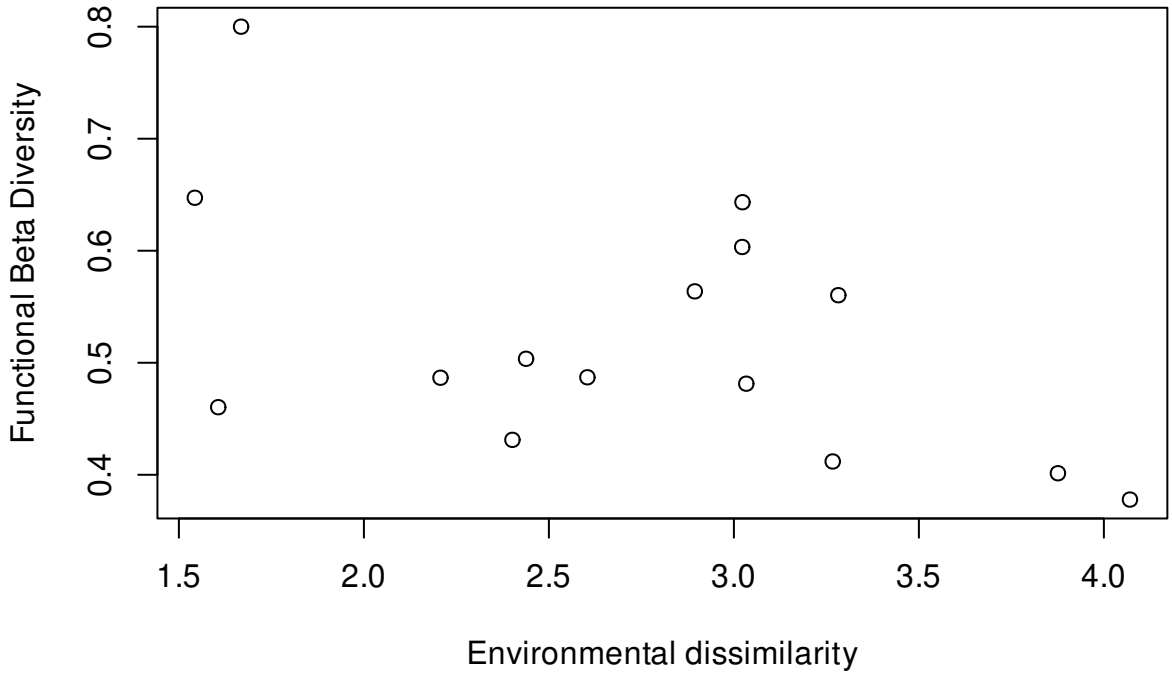

```
## $r
```

```
##                   BareSoil     Soil pH        Dung % Canopy Cover      alpha
## BareSoil        1.00000000  0.63638054  0.10074287     0.08464995 -0.2945215
## Soil pH         0.63638054  1.00000000  0.07931414    -0.52929548 -0.6492444
## Dung            0.10074287  0.07931414  1.00000000    -0.48449472 -0.1792223
## % Canopy Cover  0.08464995 -0.52929548 -0.48449472     1.00000000  0.4224203
## alpha          -0.29452149 -0.64924437 -0.17922233     0.42242028  1.0000000
## gamma           0.32444995  0.34489853 -0.24950373     0.24189140 -0.7690976
## lambda          0.92590650  0.54290949 -0.26371838     0.29094318 -0.2602762
## A               0.00053158  0.13433730  0.76099818    -0.75205975 -0.2710008
## phi            -0.04601876 -0.33669799 -0.77719679     0.86472596  0.1371301
##                      gamma     lambda           A          phi
## BareSoil         0.3244500  0.9259065  0.00053158 -0.04601876
## Soil pH          0.3448985  0.5429095  0.13433730 -0.33669799
## Dung            -0.2495037 -0.2637184  0.76099818 -0.77719679
## % Canopy Cover   0.2418914  0.2909432 -0.75205975  0.86472596
## alpha           -0.7690976 -0.2602762 -0.27100076  0.13713011
## gamma            1.0000000  0.4644813 -0.30837354  0.50662409
## lambda           0.4644813  1.0000000 -0.25277146  0.27296716
## A               -0.3083735 -0.2527715  1.00000000 -0.87408259
## phi              0.5066241  0.2729672 -0.87408259  1.00000000
##
## $df
## [1] 4
##
## $P
##                    BareSoil   Soil pH       Dung % Canopy Cover      alpha
## BareSoil        0.000000000 0.1742899 0.84939693     0.87332836 0.57099159
## Soil pH         0.174289943 0.0000000 0.88127826     0.28019883 0.16296762
## Dung            0.849396925 0.8812783 0.00000000     0.33012189 0.73404487
## % Canopy Cover  0.873328361 0.2801988 0.33012189     0.00000000 0.40405769
## alpha           0.570991586 0.1629676 0.73404487     0.40405769 0.00000000
## gamma           0.530402133 0.5031659 0.63351047     0.64423960 0.07381846
## lambda          0.008031389 0.2656472 0.61359289     0.57589910 0.61840177
## A               0.999202630 0.7997062 0.07885669     0.08459057 0.60345019
## phi             0.931020580 0.5140380 0.06893179     0.02621090 0.79559418
##                       gamma      lambda           A         phi
## BareSoil        0.53040213 0.008031389 0.99920263 0.93102058
## Soil pH         0.50316590 0.265647241 0.79970620 0.51403800
## Dung            0.63351047 0.613592893 0.07885669 0.06893179
## % Canopy Cover  0.64423960 0.575899099 0.08459057 0.02621090
## alpha           0.07381846 0.618401772 0.60345019 0.79559418
## gamma           0.00000000 0.353382270 0.55210196 0.30508095
## lambda          0.35338227 0.000000000 0.62891803 0.60071879
## A               0.55210196 0.628918026 0.00000000 0.02278457
## phi             0.30508095 0.600718791 0.02278457 0.00000000
```

Test relationship among distance matrices

When working with measures of beta diversity you are automatically dealing with distance matrices (i.e. the difference between two or more samples) rather than tables of point estimates (i.e site specific measures of biodiversity etc). A major problem with distance matrices is that each sample is represented more than once (usually the total number of samples minus 1) in the distance matrix, which means that the entries in the distance matrix are not independent. This violates a major assumption of most traditional statistics. For example, one cannot apply a linear model to the graph below. All three points with environmental

dissimilarity < 1.75 include the Sand Fynbos site.



In this case we need to use methods that can explore correlations among distance matrices. Perhaps the most common of these is Mantel's test, which in this case gives:

```
##
## Mantel statistic based on Pearson's product-moment correlation
##
## Call:
## mantel(xdis = msampBFD[rownames(edis), colnames(edis)], ydis = edis)
##
## Mantel statistic r: -0.5158
##       Significance: 0.92083
##
## Upper quantiles of permutations (null model):
##   90%   95% 97.5%   99%
## 0.484 0.592 0.640 0.725
## Permutation: free
## Number of permutations: 719
```

References (papers accessible at this link)

Bergh, Nicola G., G. A. Verboom, Mathieu Rouget, and Richard M. Cowling. 2014. "Vegetation Types of the Greater Cape Floristic Region." in Fynbos: Ecology, Evolution, and Conservation of a Megadiverse Region. Allsopp et al. (Eds)

Mucina, Ladislav, Michael C. Rutherford, and Others. 2006. The Vegetation of South Africa, Lesotho and

Swaziland. South African National Biodiversity Institute. https://www.cabdirect.org/cabdirect/abstract/20073221934.

Petchey, Owen L., and Kevin J. Gaston. 2002. "Functional Diversity (FD), Species Richness and Community Composition." Ecology Letters 5 (3): 402–11. https://doi.org/10.1046/j.1461-0248.2002.00339.x.

Plas, Fons van der. 2019. "Biodiversity and Ecosystem Functioning in Naturally Assembled Communities." Biological Reviews of the Cambridge Philosophical Society 94 (4): 1220–45. https://doi.org/10.1111/brv.12499.

Slingsby, Jasper A., Glenn R. Moncrieff, and Adam M. Wilson. 2020. "Near-Real Time Forecasting and Change Detection for an Open Ecosystem with Complex Natural Dynamics." ISPRS Journal of Photogrammetry and Remote Sensing. 166 (August): 15–25. https://doi.org/10.1016/j.isprsjprs.2020.05.017.

Wilson, Adam M., Andrew M. Latimer, and John A. Silander Jr. 2015. "Climatic Controls on Ecosystem Resilience: Postfire Regeneration in the Cape Floristic Region of South Africa." Proceedings of the National Academy of Sciences of the United States of America 112 (29): 9058–63. https://doi.org/10.1073/pnas.1416710112.