in-the-absence-of-strong-non-linearity.R

USER

Wed Oct 23 11:59:41 2019

library(here)

## here() starts at C:/Users/user/Desktop/Cape-vs-SWA

source(here("draft-02/manuscript\_ver3/R/01\_setup.R"))

## -- Attaching packages ------------------------------------------------------------- tidyverse 1.2.1 --

## v ggplot2 3.1.0 v purrr 0.3.0   
## v tibble 2.0.1 v dplyr 0.8.0.1  
## v tidyr 0.8.2 v stringr 1.4.0   
## v readr 1.3.1 v forcats 0.4.0

## -- Conflicts ---------------------------------------------------------------- tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

##   
## Attaching package: 'glue'

## The following object is masked from 'package:dplyr':  
##   
## collapse

##   
## Attaching package: 'magrittr'

## The following object is masked from 'package:purrr':  
##   
## set\_names

## The following object is masked from 'package:tidyr':  
##   
## extract

## Loading required package: sp

## Warning: package 'sp' was built under R version 3.5.3

##   
## Attaching package: 'raster'

## The following object is masked from 'package:magrittr':  
##   
## extract

## The following object is masked from 'package:glue':  
##   
## trim

## The following object is masked from 'package:dplyr':  
##   
## select

## The following object is masked from 'package:tidyr':  
##   
## extract

## Warning: package 'rgdal' was built under R version 3.5.3

## rgdal: version: 1.4-4, (SVN revision 833)  
## Geospatial Data Abstraction Library extensions to R successfully loaded  
## Loaded GDAL runtime: GDAL 2.2.3, released 2017/11/20  
## Path to GDAL shared files: C:/Users/user/Documents/R/win-library/3.5/rgdal/gdal  
## GDAL binary built with GEOS: TRUE   
## Loaded PROJ.4 runtime: Rel. 4.9.3, 15 August 2016, [PJ\_VERSION: 493]  
## Path to PROJ.4 shared files: C:/Users/user/Documents/R/win-library/3.5/rgdal/proj  
## Linking to sp version: 1.3-1

## Warning: package 'canprot' was built under R version 3.5.3

## human\_base: 21006 proteins

## human\_additional: 71173 proteins

## human\_extra: 337 proteins

## Warning: package 'cowplot' was built under R version 3.5.3

##   
## \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

## Note: As of version 1.0.0, cowplot does not change the

## default ggplot2 theme anymore. To recover the previous

## behavior, execute:  
## theme\_set(theme\_cowplot())

## \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

## Warning: package 'ggfortify' was built under R version 3.5.3

## Warning: package 'rasterVis' was built under R version 3.5.3

## Loading required package: lattice

## Loading required package: latticeExtra

## Warning: package 'latticeExtra' was built under R version 3.5.3

## Loading required package: RColorBrewer

##   
## Attaching package: 'latticeExtra'

## The following object is masked from 'package:ggplot2':  
##   
## layer

##   
## Attaching package: 'scales'

## The following object is masked from 'package:purrr':  
##   
## discard

## The following object is masked from 'package:readr':  
##   
## col\_factor

library(MVN)

## Warning: package 'MVN' was built under R version 3.5.3

## sROC 0.1-2 loaded

data <- list(  
 QDS = read\_csv(glue("{data\_dir}/data-QDS-w-residuals.csv")),  
 HDS = read\_csv(glue("{data\_dir}/data-HDS-w-residuals.csv")),  
 DS = read\_csv(glue("{data\_dir}/data-DS-w-residuals.csv"))  
)

## Parsed with column specification:  
## cols(  
## Elevation = col\_double(),  
## MAP = col\_double(),  
## PDQ = col\_double(),  
## Surface\_T = col\_double(),  
## NDVI = col\_double(),  
## CEC = col\_double(),  
## Clay = col\_double(),  
## Soil\_C = col\_double(),  
## pH = col\_double(),  
## region = col\_character(),  
## PC1 = col\_double(),  
## lon = col\_double(),  
## lat = col\_double(),  
## QDS = col\_character(),  
## n\_EDS\_in\_region = col\_double(),  
## n\_collections = col\_double(),  
## QDS\_richness = col\_double(),  
## PC1\_residual = col\_double(),  
## multivariate\_residual = col\_double()  
## )

## Parsed with column specification:  
## cols(  
## .default = col\_double(),  
## region = col\_character(),  
## HDS = col\_character()  
## )

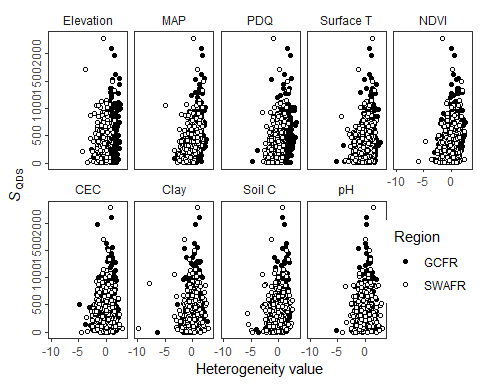
## See spec(...) for full column specifications.

## Parsed with column specification:  
## cols(  
## .default = col\_double(),  
## region = col\_character(),  
## DS = col\_character()  
## )

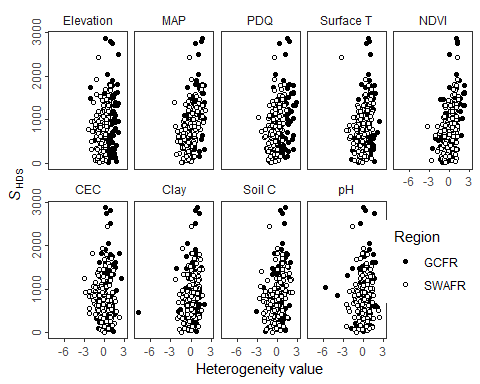
## See spec(...) for full column specifications.

data %>%  
 map(gather, environmental\_variable, heterogeneity\_value, Elevation:pH) %>%  
 map(mutate, environmental\_variable = environmental\_variable %>%  
 str\_replace\_all("\_", " ") %>%  
 factor(levels = var\_names)  
 ) %>%  
 imap(  
 ~ ggplot(.x) +  
 aes\_string(  
 "heterogeneity\_value", glue("{.y}\_richness"),  
 fill = "region"  
 ) +  
 geom\_point(shape = 21) +  
 scale\_fill\_manual(name = "Region", values = c("black", "white")) +  
 labs(x = "Heterogeneity value", y = bquote(italic("S")[.(.y)])) +  
 facet\_wrap(~environmental\_variable, nrow = 2) +  
 theme(  
 legend.position = c(0.9, 0.25),  
 axis.text.y = element\_text(angle = 90, hjust = 0.5)  
 )  
 )

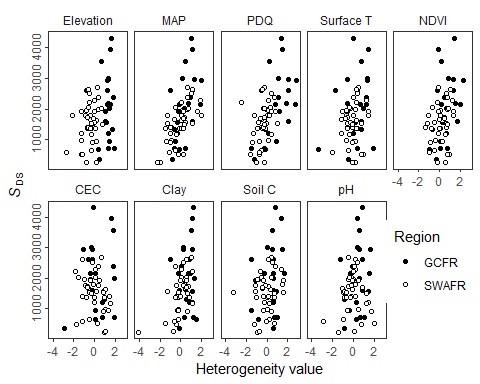
## $QDS



##   
## $HDS



##   
## $DS



mvn(data$QDS[, c(str\_replace\_all(var\_names, " ", "\_"), "QDS\_richness")])

## $multivariateNormality  
## Test Statistic p value Result  
## 1 Mardia Skewness 3399.28928632821 0 NO  
## 2 Mardia Kurtosis 65.1627466895385 0 NO  
## 3 MVN <NA> <NA> NO  
##   
## $univariateNormality  
## Test Variable Statistic p value Normality  
## 1 Shapiro-Wilk Elevation 0.9810 <0.001 NO   
## 2 Shapiro-Wilk MAP 0.9928 3e-04 NO   
## 3 Shapiro-Wilk PDQ 0.9817 <0.001 NO   
## 4 Shapiro-Wilk Surface\_T 0.9821 <0.001 NO   
## 5 Shapiro-Wilk NDVI 0.9763 <0.001 NO   
## 6 Shapiro-Wilk CEC 0.9837 <0.001 NO   
## 7 Shapiro-Wilk Clay 0.8870 <0.001 NO   
## 8 Shapiro-Wilk Soil\_C 0.9541 <0.001 NO   
## 9 Shapiro-Wilk pH 0.9916 1e-04 NO   
## 10 Shapiro-Wilk QDS\_richness 0.8595 <0.001 NO   
##   
## $Descriptives  
## n Mean Std.Dev Median Min  
## Elevation 897 7.730229e-03 1.0004493 -0.14329384 -4.398168  
## MAP 897 2.064284e-02 0.9890072 0.04116779 -4.921397  
## PDQ 897 4.132337e-03 1.0014037 -0.10273328 -4.789219  
## Surface\_T 897 1.562300e-02 0.9853468 0.03776854 -5.213080  
## NDVI 897 2.115264e-02 0.9670813 0.08506550 -5.910475  
## CEC 897 4.583554e-03 0.9978332 0.04116601 -4.797507  
## Clay 897 2.148703e-02 0.9836372 0.05969178 -9.815162  
## Soil\_C 897 2.508763e-03 0.9959214 0.09159413 -5.630048  
## pH 897 1.255236e-02 0.9913518 0.03181393 -5.212726  
## QDS\_richness 897 3.568919e+02 304.7498728 286.00000000 1.000000  
## Max 25th 75th Skew Kurtosis  
## Elevation 2.572127 -0.6049895 0.6817387 0.03499757 0.4177338  
## MAP 2.437376 -0.6507256 0.7277739 -0.24671730 0.3419298  
## PDQ 3.104414 -0.6473026 0.5733835 0.22975986 0.8504659  
## Surface\_T 2.729412 -0.5840353 0.6906098 -0.54334515 1.4145977  
## NDVI 2.506350 -0.5368687 0.6814544 -0.68395250 2.0262901  
## CEC 2.832029 -0.5852253 0.7031603 -0.54998341 1.0615750  
## Clay 2.920858 -0.4821128 0.5961381 -2.10006771 16.9725276  
## Soil\_C 2.457962 -0.5912595 0.6815475 -1.01807103 3.2945268  
## pH 3.202574 -0.6089060 0.6614978 -0.21678932 1.1472779  
## QDS\_richness 2287.000000 144.0000000 485.0000000 1.77995198 5.0135831

mvn(data$HDS[, c(str\_replace\_all(var\_names, " ", "\_"), "HDS\_richness")])

## $multivariateNormality  
## Test Statistic p value Result  
## 1 Mardia Skewness 1068.62015461087 1.66288132016507e-111 NO  
## 2 Mardia Kurtosis 18.7168624059074 0 NO  
## 3 MVN <NA> <NA> NO  
##   
## $univariateNormality  
## Test Variable Statistic p value Normality  
## 1 Shapiro-Wilk Elevation 0.9922 0.2379 YES   
## 2 Shapiro-Wilk MAP 0.9898 0.0926 YES   
## 3 Shapiro-Wilk PDQ 0.9708 1e-04 NO   
## 4 Shapiro-Wilk Surface\_T 0.9873 0.0334 NO   
## 5 Shapiro-Wilk NDVI 0.9877 0.0397 NO   
## 6 Shapiro-Wilk CEC 0.9923 0.2457 YES   
## 7 Shapiro-Wilk Clay 0.8496 <0.001 NO   
## 8 Shapiro-Wilk Soil\_C 0.9841 0.0093 NO   
## 9 Shapiro-Wilk pH 0.9414 <0.001 NO   
## 10 Shapiro-Wilk HDS\_richness 0.9434 <0.001 NO   
##   
## $Descriptives  
## n Mean Std.Dev Median Min  
## Elevation 238 4.362229e-02 0.9388588 -0.056715543 -2.487910  
## MAP 238 8.577895e-02 0.9739320 0.028552393 -2.533032  
## PDQ 238 5.378425e-02 0.9932871 -0.145091298 -2.770862  
## Surface\_T 238 7.708984e-02 0.9024968 0.109178785 -3.358693  
## NDVI 238 8.093076e-02 0.9145392 0.145224363 -3.315396  
## CEC 238 3.864205e-04 0.9349775 0.002564074 -2.980820  
## Clay 238 5.107935e-02 0.9484748 0.130334819 -7.856085  
## Soil\_C 238 6.823530e-02 0.9566463 0.166580249 -3.136899  
## pH 238 1.241093e-02 0.9890859 0.008392024 -5.635530  
## HDS\_richness 238 8.463571e+02 531.5058801 801.000000000 14.000000  
## Max 25th 75th Skew Kurtosis  
## Elevation 2.087566 -0.5876184 0.7140312 0.01568991 -0.4157390  
## MAP 2.107175 -0.5573021 0.7376676 -0.02917571 -0.5283062  
## PDQ 2.883094 -0.5760422 0.6023185 0.53452767 0.2184407  
## Surface\_T 2.348607 -0.4957149 0.7665361 -0.41083562 0.3083739  
## NDVI 2.264628 -0.5423857 0.7142780 -0.34213589 0.5924214  
## CEC 2.560445 -0.5223293 0.6705528 -0.31882885 0.2165780  
## Clay 1.957399 -0.4489745 0.6827627 -2.61464198 18.7921083  
## Soil\_C 2.244513 -0.4653847 0.7231860 -0.47723895 0.2609125  
## pH 2.431374 -0.4738467 0.6529132 -1.11199378 4.5515573  
## HDS\_richness 2875.000000 465.2500000 1135.0000000 0.97328103 1.4657348

mvn(data$DS[, c(str\_replace\_all(var\_names, " ", "\_"), "DS\_richness")])

## $multivariateNormality  
## Test Statistic p value Result  
## 1 Mardia Skewness 459.057395555042 5.82447555115049e-19 NO  
## 2 Mardia Kurtosis 5.21278787465366 1.86023571435712e-07 NO  
## 3 MVN <NA> <NA> NO  
##   
## $univariateNormality  
## Test Variable Statistic p value Normality  
## 1 Shapiro-Wilk Elevation 0.9633 0.0572 YES   
## 2 Shapiro-Wilk MAP 0.9885 0.8262 YES   
## 3 Shapiro-Wilk PDQ 0.9619 0.0486 NO   
## 4 Shapiro-Wilk Surface\_T 0.9537 0.0187 NO   
## 5 Shapiro-Wilk NDVI 0.9891 0.8509 YES   
## 6 Shapiro-Wilk CEC 0.9811 0.4453 YES   
## 7 Shapiro-Wilk Clay 0.8525 <0.001 NO   
## 8 Shapiro-Wilk Soil\_C 0.9489 0.0109 NO   
## 9 Shapiro-Wilk pH 0.9802 0.4046 YES   
## 10 Shapiro-Wilk DS\_richness 0.9682 0.103 YES   
##   
## $Descriptives  
## n Mean Std.Dev Median Min  
## Elevation 63 1.406010e-01 1.0283907 0.02172532 -2.684918  
## MAP 63 4.780904e-02 0.9705136 0.09226162 -2.287604  
## PDQ 63 4.074658e-02 1.0543135 -0.19988012 -2.485886  
## Surface\_T 63 4.720253e-02 0.9219765 0.03283641 -3.399848  
## NDVI 63 2.875597e-01 0.7885909 0.22097325 -1.333793  
## CEC 63 6.594642e-02 0.9888176 0.03174996 -2.910094  
## Clay 63 1.939684e-01 0.8596567 0.29014590 -4.085964  
## Soil\_C 63 9.884662e-02 0.9145478 0.23903996 -3.291785  
## pH 63 7.938959e-02 0.8860937 0.01432137 -2.872186  
## DS\_richness 63 1.728984e+03 854.6007083 1675.00000000 238.000000  
## Max 25th 75th Skew Kurtosis  
## Elevation 1.936882 -0.6232803 1.1815375 -0.18963554 -0.53148410  
## MAP 1.988253 -0.5897579 0.6846998 -0.01125138 -0.48420616  
## PDQ 2.829196 -0.5778798 0.6031533 0.57269517 0.27320382  
## Surface\_T 1.886663 -0.5504790 0.6928476 -0.73698764 1.77514417  
## NDVI 2.251177 -0.2004538 0.7190075 0.12362338 -0.37592439  
## CEC 1.900234 -0.5200396 0.8280049 -0.24591164 0.07654095  
## Clay 1.506906 -0.2193850 0.7822662 -1.97786880 7.65126511  
## Soil\_C 1.628105 -0.4358373 0.6352570 -0.79275910 1.35566267  
## pH 2.071172 -0.4437899 0.5787684 -0.21486468 0.70447233  
## DS\_richness 4318.000000 1212.0000000 2160.5000000 0.59998813 0.49406650