Spatial Dependence Homeowrk

Mikala

February 22, 2016

library(vegan)

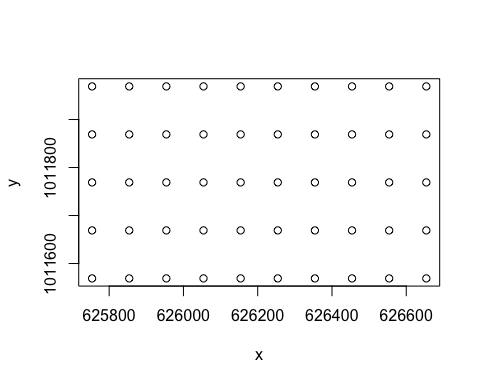
## Loading required package: permute

## Loading required package: lattice

## This is vegan 2.3-3

library(nlme)  
data(BCI)  
BCI\_xy = data.frame(x = rep(seq(625754, 626654, by=100), each=5),   
 y = rep(seq(1011569, 1011969, by=100), len=50))  
dist(BCI\_xy)

xy\_dist = dist(BCI\_xy)  
max\_dist = max(xy\_dist) / 2  
# Oenocarpus.mapora: common Pouteria.reticulata: rare  
plot(BCI\_xy)



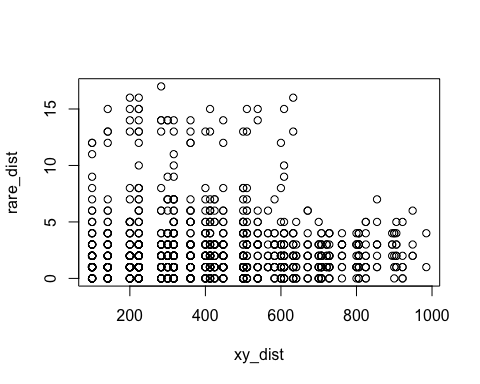
common = BCI$Oenocarpus.mapora  
rare = BCI$Pouteria.reticulata  
common\_dist = dist(common)  
rare\_dist = dist(rare)  
com = mantel(xy\_dist, common\_dist)  
com

##   
## Mantel statistic based on Pearson's product-moment correlation   
##   
## Call:  
## mantel(xdis = xy\_dist, ydis = common\_dist)   
##   
## Mantel statistic r: 0.1935   
## Significance: 0.001   
##   
## Upper quantiles of permutations (null model):  
## 90% 95% 97.5% 99%   
## 0.0646 0.0843 0.0946 0.1126   
## Permutation: free  
## Number of permutations: 999

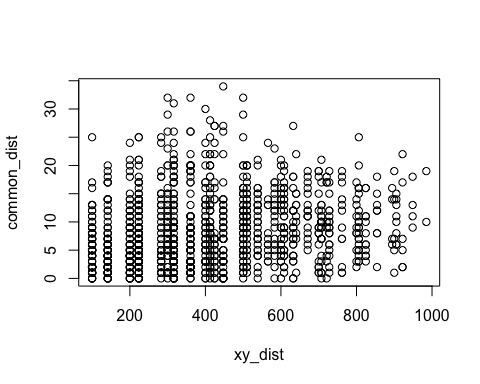
rare = mantel(xy\_dist, rare\_dist)  
rare

##   
## Mantel statistic based on Pearson's product-moment correlation   
##   
## Call:  
## mantel(xdis = xy\_dist, ydis = rare\_dist)   
##   
## Mantel statistic r: -0.04323   
## Significance: 0.73   
##   
## Upper quantiles of permutations (null model):  
## 90% 95% 97.5% 99%   
## 0.0925 0.1107 0.1236 0.1454   
## Permutation: free  
## Number of permutations: 999

plot(xy\_dist, rare\_dist)



plot(xy\_dist, common\_dist)



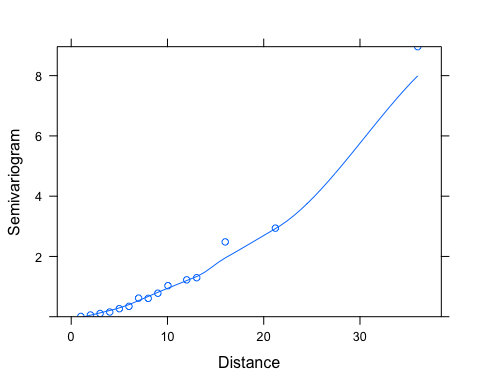
# The Mantel statistic for the common species indicates there is a positive correlation between spatial dependance and the common species. The Mantel statistic for the rare species indicates there is only a very slight correlation between the species and spatial dependance.   
  
# Model 1  
DS = BCI$Drypetes.standleyi  
QA = BCI$Quassia.amara  
mod\_DS1 <- glm(DS ~ QA, data= BCI)  
mod\_DS1

##   
## Call: glm(formula = DS ~ QA, data = BCI)  
##   
## Coefficients:  
## (Intercept) QA   
## 4.963 9.215   
##   
## Degrees of Freedom: 49 Total (i.e. Null); 48 Residual  
## Null Deviance: 3290   
## Residual Deviance: 2469 AIC: 342.9

summary(mod\_DS1)

##   
## Call:  
## glm(formula = DS ~ QA, data = BCI)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -4.963 -4.963 -2.963 2.037 34.037   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.963 1.031 4.815 1.51e-05 \*\*\*  
## QA 9.215 2.305 3.998 0.000219 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for gaussian family taken to be 51.42777)  
##   
## Null deviance: 3290.5 on 49 degrees of freedom  
## Residual deviance: 2468.5 on 48 degrees of freedom  
## AIC: 342.86  
##   
## Number of Fisher Scoring iterations: 2

mod\_DS1 <- gls(DS ~ QA, data= BCI)  
par(mfrow=c(1,1))  
plot(Variogram(mod\_DS1, form= ~ DS + QA))



x = BCI\_xy$x   
y = BCI\_xy$y  
sp\_ids1 = c("Drypetes.standleyi", "Quassia.amara")  
abn\_xy = data.frame(BCI, BCI\_xy)  
BCI\_sub1 = subset(abn\_xy, select=sp\_ids1)  
gls\_multi1 = gls(Drypetes.standleyi ~ Quassia.amara, data = BCI\_sub1)  
multi\_rationalnug1 = update(gls\_multi1, corr=corRatio(form=~'x' + 'y', nugget=T))  
gls\_multi1

## Generalized least squares fit by REML  
## Model: Drypetes.standleyi ~ Quassia.amara   
## Data: BCI\_sub1   
## Log-restricted-likelihood: -165.7644  
##   
## Coefficients:  
## (Intercept) Quassia.amara   
## 4.962810 9.214876   
##   
## Degrees of freedom: 50 total; 48 residual  
## Residual standard error: 7.171316

multi\_rationalnug1

## Generalized least squares fit by REML  
## Model: Drypetes.standleyi ~ Quassia.amara   
## Data: BCI\_sub1   
## Log-restricted-likelihood: -159.1761  
##   
## Coefficients:  
## (Intercept) Quassia.amara   
## 5.993552 7.149095   
##   
## Correlation Structure: Rational quadratic spatial correlation  
## Formula: ~"x" + "y"   
## Parameter estimate(s):  
## range nugget   
## 49.1419743 0.2943482   
## Degrees of freedom: 50 total; 48 residual  
## Residual standard error: 11.07303

summary(gls\_multi1)

## Generalized least squares fit by REML  
## Model: Drypetes.standleyi ~ Quassia.amara   
## Data: BCI\_sub1   
## AIC BIC logLik  
## 337.5287 343.1423 -165.7644  
##   
## Coefficients:  
## Value Std.Error t-value p-value  
## (Intercept) 4.962810 1.030804 4.814503 0e+00  
## Quassia.amara 9.214876 2.304948 3.997867 2e-04  
##   
## Correlation:   
## (Intr)  
## Quassia.amara -0.179  
##   
## Standardized residuals:  
## Min Q1 Med Q3 Max   
## -0.6920362 -0.6920362 -0.4131473 0.2840748 4.7462963   
##   
## Residual standard error: 7.171316   
## Degrees of freedom: 50 total; 48 residual

summary(multi\_rationalnug1)

## Generalized least squares fit by REML  
## Model: Drypetes.standleyi ~ Quassia.amara   
## Data: BCI\_sub1   
## AIC BIC logLik  
## 328.3522 337.7082 -159.1761  
##   
## Correlation Structure: Rational quadratic spatial correlation  
## Formula: ~"x" + "y"   
## Parameter estimate(s):  
## range nugget   
## 49.1419743 0.2943482   
##   
## Coefficients:  
## Value Std.Error t-value p-value  
## (Intercept) 5.993552 8.335829 0.719011 0.4756  
## Quassia.amara 7.149095 1.994282 3.584797 0.0008  
##   
## Correlation:   
## (Intr)  
## Quassia.amara -0.044  
##   
## Standardized residuals:  
## Min Q1 Med Q3 Max   
## -0.54127474 -0.54127474 -0.36065574 0.09089177 2.98079586   
##   
## Residual standard error: 11.07303   
## Degrees of freedom: 50 total; 48 residual

anova(gls\_multi1, multi\_rationalnug1)

## Model df AIC BIC logLik Test L.Ratio  
## gls\_multi1 1 3 337.5287 343.1423 -165.7644   
## multi\_rationalnug1 2 5 328.3522 337.7082 -159.1761 1 vs 2 13.17651  
## p-value  
## gls\_multi1   
## multi\_rationalnug1 0.0014

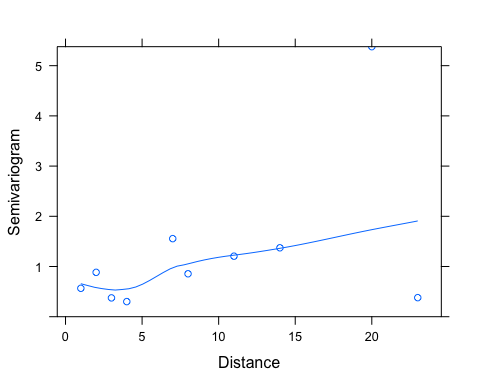
# When comparing the coefficients for the species QA that was used to predict DS, the coefficients essentially didn't change when comparing the general linear model, the gls model, and the rational quadratic error model with a nugget. This makes sense because only one species is being used to predict the abundance of another species.   
# The anova comparing the gls model to the rational quadratic error model with a nugget shows that the rational quadratic error model with nugget has a slightly lower AIC value, implying that it may better model the data than the gls model. The very low P-value indicates that there is a significant difference between the models, so we can assume the rational quadratic error model (w/ nugget) is the model that can better explain the spatial data  
# When comparing the summaries of the general linear model predicitng DS by one other species, QA, to the rational quadratic error model with a nugget, the AIC value for the rational model was slightly lower than the AIC value for the general linear model. This indicates that taking spatial dependence into consideration slightly improved the model. When comparing the glm and the gls models predicting DS by QA, they had the same AIC and P-values, indicating that the gls didn't do a radically better job of modelling the data.   
  
#Model 2  
group\_sp = BCI$Cordia.lasiocalyx + BCI$Hirtella.triandra + BCI$Picramnia.latifolia + BCI$Quassia.amara +BCI$Tabernaemontana.arborea + BCI$Trattinnickia.aspera + BCI$Xylopia.macrantha  
mod\_DS2 = glm(DS ~ Cordia.lasiocalyx + Hirtella.triandra + Picramnia.latifolia + Quassia.amara + Tabernaemontana.arborea + Trattinnickia.aspera + Xylopia.macrantha, data = BCI)  
mod\_DS2

##   
## Call: glm(formula = DS ~ Cordia.lasiocalyx + Hirtella.triandra + Picramnia.latifolia +   
## Quassia.amara + Tabernaemontana.arborea + Trattinnickia.aspera +   
## Xylopia.macrantha, data = BCI)  
##   
## Coefficients:  
## (Intercept) Cordia.lasiocalyx Hirtella.triandra   
## -1.0518 0.4289 0.1223   
## Picramnia.latifolia Quassia.amara Tabernaemontana.arborea   
## 0.6623 4.0857 -0.2497   
## Trattinnickia.aspera Xylopia.macrantha   
## 1.3493 0.5488   
##   
## Degrees of Freedom: 49 Total (i.e. Null); 42 Residual  
## Null Deviance: 3290   
## Residual Deviance: 865.6 AIC: 302.5

summary(mod\_DS2)

##   
## Call:  
## glm(formula = DS ~ Cordia.lasiocalyx + Hirtella.triandra + Picramnia.latifolia +   
## Quassia.amara + Tabernaemontana.arborea + Trattinnickia.aspera +   
## Xylopia.macrantha, data = BCI)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -8.5214 -1.9385 -0.1831 1.0721 15.3791   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.05175 2.11753 -0.497 0.622000   
## Cordia.lasiocalyx 0.42892 0.20393 2.103 0.041474 \*   
## Hirtella.triandra 0.12228 0.08026 1.523 0.135138   
## Picramnia.latifolia 0.66226 0.63589 1.041 0.303617   
## Quassia.amara 4.08566 2.28428 1.789 0.080894 .   
## Tabernaemontana.arborea -0.24973 0.14912 -1.675 0.101431   
## Trattinnickia.aspera 1.34932 0.71474 1.888 0.065968 .   
## Xylopia.macrantha 0.54883 0.14688 3.737 0.000557 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for gaussian family taken to be 20.60899)  
##   
## Null deviance: 3290.50 on 49 degrees of freedom  
## Residual deviance: 865.58 on 42 degrees of freedom  
## AIC: 302.46  
##   
## Number of Fisher Scoring iterations: 2

mod\_DS2 <- gls(DS ~ Cordia.lasiocalyx + Hirtella.triandra + Picramnia.latifolia + Quassia.amara + Tabernaemontana.arborea + Trattinnickia.aspera + Xylopia.macrantha, data= BCI)  
par(mfrow=c(1,1))  
plot(Variogram(mod\_DS2, form= ~ DS + Cordia.lasiocalyx + Hirtella.triandra + Picramnia.latifolia + Quassia.amara + Tabernaemontana.arborea + Trattinnickia.aspera + Xylopia.macrantha))



x = BCI\_xy$x   
y = BCI\_xy$y  
sp\_ids = c("Drypetes.standleyi", "Cordia.lasiocalyx", "Hirtella.triandra",  
 "Picramnia.latifolia", "Quassia.amara",  
 "Tabernaemontana.arborea", "Trattinnickia.aspera",   
 "Xylopia.macrantha")  
abn\_xy = data.frame(BCI, BCI\_xy)  
BCI\_sub = subset(abn\_xy, select=sp\_ids)  
gls\_multi = gls(Drypetes.standleyi ~ Cordia.lasiocalyx + Hirtella.triandra + Picramnia.latifolia + Quassia.amara + Tabernaemontana.arborea + Trattinnickia.aspera + Xylopia.macrantha, data = BCI\_sub)  
multi\_rationalnug = update(gls\_multi, corr=corRatio(form=~'x' + 'y', nugget=T))  
multi\_rationalnug

## Generalized least squares fit by REML  
## Model: Drypetes.standleyi ~ Cordia.lasiocalyx + Hirtella.triandra + Picramnia.latifolia + Quassia.amara + Tabernaemontana.arborea + Trattinnickia.aspera + Xylopia.macrantha   
## Data: BCI\_sub   
## Log-restricted-likelihood: -141.5708  
##   
## Coefficients:  
## (Intercept) Cordia.lasiocalyx Hirtella.triandra   
## -0.60827809 0.26403751 0.06303651   
## Picramnia.latifolia Quassia.amara Tabernaemontana.arborea   
## 0.48843289 3.05570050 -0.03163622   
## Trattinnickia.aspera Xylopia.macrantha   
## 1.74999828 0.59566469   
##   
## Correlation Structure: Rational quadratic spatial correlation  
## Formula: ~"x" + "y"   
## Parameter estimate(s):  
## range nugget   
## 26.4249097 0.4179927   
## Degrees of freedom: 50 total; 42 residual  
## Residual standard error: 6.233545

anova(gls\_multi, multi\_rationalnug)

## Model df AIC BIC logLik Test L.Ratio  
## gls\_multi 1 9 307.1163 322.7554 -144.5582   
## multi\_rationalnug 2 11 305.1415 324.2559 -141.5708 1 vs 2 5.974822  
## p-value  
## gls\_multi   
## multi\_rationalnug 0.0504

# For the model that includes multiple species to predict abundance of DS, the coefficients did change between the general linear model and the rational quadratic error model with a nugget. This indicates that accounting for spatial dependence is important when using multiple species to predict the abundance of DS  
# When looking at the anova outputs, the AIC value for the rational quadratic error model with a nugget is slightly lower, indicating that it may be a slightly better model for the data than the gls model. Having a non-origin y-intercept by adding the nugget improves the model by allowing for a better fit and a more accurate starting point.  
# Adding the spatial error term had more of an impact on model 2 because it includes multiple species as predictor variables, which means more variables are able to display spatial dependence. Since we are addressing species locations and abundance, the majority of the species should have some degree of spacial dependence, so the more species that are included in the model, the larger the influence of spatial depence will be. Model 1 had less of a change when spatial dependence was taken into consideration because only 1 other species was used as a predictive variable.