Supporting Information for "The optimal species richness environments for human populations." By Freeman et al. 2018 Submitted to PNAS July 2018

Jacob Freeman [1]

[1]Anthropology Program, Utah State University

This supporting document should allow one to recreate the analysis performed as part of "The optimal species richness environments for human populations, " By Freeman et al. 2018 Submitted to PNAS July 2018. The annotated scripts below contain code to replicate the analysis, as well as code for additional analyses not included in the main paper or supporting information. To replicate the analysis, one can either analyze the data files provided or build their own data set. As discussed in the main body of the text, we built three data sets following the procedures outlined by Tallavaara et al. (2017) for linking species richness values, net primary productivity and pathogen stress to each ethnographic case. We do not replicate the scripts provided by Tallavaara et al. (2017) as these are available, clear and should be cited when used.

To replicate our analysis, one needs to set their working directory in R to the file location that contains the data files. There are 11 files that follow the naming convention "name.csv." The 11 files are "MainFinal.csv". "AGPOP3Eco.csv", "HGFEM4R.csv", "AGPOPClass.csv", "CountryMeansEco2.csv", "AGPOP3EcoH.csv", "AGPOP3EcoL.csv", "HiHG.csv", "LowHG,csv", "CountryMeansEco2H.csv", and "CountryMeansEco2L.csv". The first five files are the main files, the second six files are divided into high and low species richness environments by economy type for convenience. In each file, the variables are defined as follows:

1. Group/Country–name of the ethnographic society of country

2. Latitude–the latitude at the geographic center of a group's territory or a country's territory.

3. Longitude–the longitude at the geographic center of a group's territory or a country's territory.

4. Class–an ordinal ranking of wealth and status differentiation among the hunter-gatherer and agriculturalists societies (see main text for more

5. Class2–an binary ranking of wealth and status differentiation among the hunter-gatherer and agriculturalists societies (see main text for more details).ails).

6. ECI–The average economic complexity index since 1973 as measured among modern countries.

7. DENSITY–Population density in people per square kilometer. This is a point in time estimate for hunter-gatherer and agricultural groups and an average density since 1973 among nation states.

8. LnDENSITY–The natural log of population density

9. npp–net primary productivity estimated at the center of each group's territory

10. npp2-Net primary productivity squared

11. biodiv–Standardized estimate of species richness at the center of each group's range.

12. biodiv2–Species richness *100 ad squared.

13. pathos–Index of pathogen stress at the center of a group's territory.

14. DivDiff–The absolute value of species richness-the species richness value of peak population density (values identified in Fig. 2 of the main manuscript).

15. ID–A nominal variable that denotes economy type. HG=hunter-gatherer, AG=subsistence agriculturalist, IND=modern nation state

Analysis

1. Explore the relationship between species richness and human population density, controlling for basic economic orientation (e.g., nation state, subsistence agriculture, hunter-gatherer). Set working directory to the location of the data files above. And then run the following code.

```
##Load packages.
library(ggplot2)
library(plyr)
library(scales)
library(effects)
library(splines)
library(sandwich)
library(lmtest)
library(boot)
library(breakpoint)
library(lavaan)
library(piecewiseSEM)
library(semPlot)
library(segmented)
library(reshape2)
library(xtable)

#####Set working directory to the source file location
#setwd()

  #Load data
AllSoc<-read.csv(file="MainFinal.csv", header=T)

  #A.#Make scatter plots
###Scatterplot of species richness against the log of population density
#controlling for economy type
p <- ggplot(AllSoc, aes((biodiv), log(DENSITY)))
p + aes(shape = factor(ID), color = factor(ID)) +
  theme_bw() +
  geom_point(aes(size = factor(ID)))+
  scale_size_manual(values=c(3,1,2))+
  theme(axis.text = element_text(size = rel(1.5), colour = "black"),
  axis.title=element_text(size=16))+
  labs(x = "Standardized species richness", y="log Population density")+
  #geom_smooth(method='lm', formula=y ~ ns(x, 2), se=0.95)+
  geom_smooth(se=FALSE)+
  geom_vline(xintercept=0.23)+
  geom_vline(xintercept=0.28)+
  geom_vline(xintercept=0.38)
```

```
###Scatterplot of npp against the log of population density
#controlling for economy type
p <- ggplot(AllSoc, aes((npp), log(DENSITY)))
p + aes(shape = factor(ID), color = factor(ID)) +
  theme_bw() +
  geom_point(aes(size = factor(ID)))+
  scale_size_manual(values=c(3,1,2))+
  theme(axis.text = element_text(size = rel(1.5), colour = "black"),
  axis.title=element_text(size=16))+
  labs(x = "Net primary productivity", y="log Population density")+
  #geom_smooth(method='lm', formula=y ~ ns(x, 2), se=0.95)+
  geom_smooth(se=FALSE)+
  #geom_vline(xintercept=1100)+
  #geom_vline(xintercept=0.28)+
  geom_vline(xintercept=1100)


###Scatterplot of npp agaist species richness
#controlling for economy type
p <- ggplot(AllSoc, aes((npp), (biodiv)))
p + aes(shape = factor(ID), color = factor(ID)) +
  theme_bw() +
  geom_point(aes(size = factor(ID)))+
  scale_size_manual(values=c(3,1,2))+
  theme(axis.text = element_text(size = rel(1.5), colour = "black"),
  axis.title=element_text(size=16))+
  labs(x = "Net primary productivity", y="Standardized species richness")+
  #geom_smooth(method='lm', formula=y ~ ns(x, 2), se=0.95)+
  geom_smooth(method="lm")+
  #geom_vline(xintercept=1100)+
  #geom_vline(xintercept=0.28)+
  geom_vline(xintercept=1100)


###Scatterplot of species richness against the log of population density
#controlling for economy type
p <- ggplot(AllSoc, aes((biodiv),log(DENSITY)))
p + aes(shape = factor(ID), color = factor(ID)) +
  theme_bw() +
  geom_point(aes(size = factor(ID)))+
  scale_size_manual(values=c(3,1,2))+
  theme(axis.text = element_text(size = rel(1.5), colour = "black"),
  axis.title=element_text(size=18))+
  labs(x = "Standardized species richness", y="log Population density")+
  #geom_smooth(method='lm', formula=y ~ ns(x, 2), se=0.95)+
 facet_wrap(~factor(ID))

  #A_2###Fit a quadratic regression controlling for economy type

    fit <- lm((DENSITY) ~ biodiv+biodiv2+factor(ID),data=AllSoc)
    BioQuad<-summary(fit)
    BioQuad
    xtable(BioQuad)

  #A_3.##Scatterplots by economy type, fitting a quadratic function
    keep<-read.csv(file="AGPOP3Eco.csv", header=T)
    keep2<-read.csv(file="HGFEM4R.csv", header=T)
    keep3<-read.csv(file="CountryMeansEco2.csv", header=T)

    ##Agriculturalists
    p <- ggplot(keep, aes((biodiv),DENSITY))
    p  + geom_point(color="steelblue4", size=2.5)+
      theme_bw() +
      theme(axis.text = element_text(size = rel(1.5), colour = "black"),
      axis.title=element_text(size=20))+
      labs(x = "Standardized species richness", y="Population density")+
      geom_smooth(method='lm', formula=y ~ ns(x, 2), se=0.95)
```

```
    ##Hunter-Gatherers
    p <- ggplot(keep2, aes((biodiv),DENSITY))
    p  + geom_point(color="steelblue4", size=2.5)+
      theme_bw() +
      theme(axis.text = element_text(size = rel(1.5), colour = "black"),
      axis.title=element_text(size=20))+
      labs(x = "Standardized species richness", y="Population density")+
      geom_smooth(method='lm', formula=y ~ ns(x, 2), se=0.95)

    ##Nation States
    p <- ggplot(keep3, aes((biodiv),DENSITY))
    p  + geom_point(color="steelblue4", size=2.5)+
      theme_bw() +
      theme(axis.text = element_text(size = rel(1.5), colour = "black"),
      axis.title=element_text(size=20))+
      labs(x = "Standardized species richness", y="Population density")+
      geom_smooth(method='lm', formula=y ~ ns(x, 2), se=0.95)
```

Breakpoint analysis

Here, conduct the breakpoint analysis using a segmented regression and graph the results using
ggplot2. Be sure to have loaded all of the packages above.

```
#2. BREAKPOINT ANALYSIS OF SPECIES RCINESS=============================================

  #A. Agriculturalist break point analysis
keep<-read.csv(file="AGPOP3Eco.csv", header=T)

  ###Quadratic Model
fit <- lm((DENSITY) ~ biodiv+biodiv2,data=keep)
BioQuadAg<-summary(fit)
xtable(BioQuadAg)


  ######Breakpoint analysis

DenBio<- lm(log(DENSITY) ~ biodiv, data=keep)
set.seed(1234)
DenBio.s <- segmented(DenBio, seg.Z=~biodiv,
control=seg.control(n.boot=1000,stop.if.error=T,it.max=1000))
summary(DenBio.s)
pscore.test(DenBio,~Biodiv)

DenBio.predict <- cbind(keep, predict(DenBio.s, interval = 'confidence'))


    ## Estimated Break-Point:

    davies.test(DenBio,~biodiv,k=20)
    pscore.test(DenBio,~Biodiv)

    plot(keep$biodiv,log(keep$DENSITY))
    plot(DenBio.s,add=TRUE,link=FALSE,lwd=2,col=2:3, lty=c(1,3))
    lines(DenBio.s,col=2,pch=19,bottom=FALSE,lwd=2)
    points(DenBio.s,col=4, link=FALSE)

    DenBio.s

    # plot the points (actual observations), regression line
    #and confidence interval
    p <- ggplot(DenBio.predict, aes(biodiv,log(DENSITY)))
    p <- p + geom_point(shape=15, color="steelblue4", size=2.5)
    p <- p + geom_line(aes(biodiv, fit))
    p <- p + geom_ribbon(aes(ymin=lwr,ymax=upr), alpha=0.3)+
```

```
    theme_bw() +
    theme(axis.text = element_text(size = rel(1.5), colour = "black"),
    axis.title=element_text(size=18))+
    labs(x = "Standardized species richness", y="log Population density")
  P


#B. Hunter-gatherer breakpoint analysis
  keep2<-read.csv(file="HGFEM4R.csv", header=T)

  ###Quadratic Model
  fit <- lm((DENSITY) ~ biodiv+biodiv2,data=keep2)
  summary(fit)

  ######Breakpoint analysis

  DenBio<- lm(log(DENSITY) ~ biodiv, data=keep2)
  set.seed(1234)
  DenBio.s <- segmented(DenBio, seg.Z=~biodiv,
  control=seg.control(n.boot=1000,stop.if.error=T,it.max=1000))
  summary(DenBio.s)


  DenBio.predict <- cbind(keep2, predict(DenBio.s, interval = 'confidence'))


  ## Estimated Break-Point:

  davies.test(DenBio,~biodiv,k=20)

  plot(keep2$biodiv,log(keep2$DENSITY))
  plot(DenBio.s,add=TRUE,link=FALSE,lwd=2,col=2:3, lty=c(1,3))
  lines(DenBio.s,col=2,pch=19,bottom=FALSE,lwd=2)
  points(DenBio.s,col=4, link=FALSE)

  DenBio.s

  # plot the points (actual observations), regression line
  #and confidence interval
  p <- ggplot(DenBio.predict, aes(biodiv,log(DENSITY)))
  p <- p + geom_point(shape=17, color="steelblue4", size=2)
  p <- p + geom_line(aes(biodiv, fit))
  p <- p + geom_ribbon(aes(ymin=lwr,ymax=upr), alpha=0.3)+
    theme_bw() +
    theme(axis.text = element_text(size = rel(1.5), colour = "black"),
    axis.title=element_text(size=18))+
    labs(x = "Standardized species richness", y="log Population density")
  P


#C. Nation state breakpoint analysis

  keep3<-read.csv(file="CountryMeansEco2.csv", header=T)

  ###Quadratic Model
  fit <- lm((DENSITY) ~ biodiv+biodiv2,data=keep3)
  summary(fit)

  ######Breakpoint analysis

  DenBio<- lm(log(DENSITY) ~ biodiv, data=keep3)
  set.seed(1234)
  DenBio.s <- segmented(DenBio, seg.Z=~biodiv,
  control=seg.control(n.boot=1000,stop.if.error=T,it.max=1000))
  summary(DenBio.s)


  DenBio.predict <- cbind(keep3, predict(DenBio.s, interval = 'confidence'))
```

```
## Estimated Break-Point:

davies.test(DenBio,~biodiv,k=20)

plot(keep3$biodiv,log(keep3$DENSITY))
plot(DenBio.s,add=TRUE,link=FALSE,lwd=2,col=2:3, lty=c(1,3))
lines(DenBio.s,col=2,pch=19,bottom=FALSE,lwd=2)
points(DenBio.s,col=4, link=FALSE)

DenBio.s

# plot the points (actual observations), regression line
#and confidence interval
p <- ggplot(DenBio.predict, aes(biodiv,log(DENSITY)))
p <- p + geom_point(color="steelblue4", size=2)
p <- p + geom_line(aes(biodiv, fit))
p <- p + geom_ribbon(aes(ymin=lwr,ymax=upr), alpha=0.3)+
  theme_bw() +
  theme(axis.text = element_text(size = rel(1.5), colour = "black"),
  axis.title=element_text(size=18))+
  labs(x = "Standardized species richness", y="log Population density")
p

#D.Breakpoint comparison======================================
#Here we compare the biodiversity breakpoints from each data set
BreakComp<-read.csv(file="BreakComp.csv", header=T)


ggplot(BreakComp, aes(Economy, BioBreak)) +
  stat_summary(fun.y=mean, geom="bar", fill="steelblue4",
  position="dodge", colour='black')+
  geom_errorbar(aes(ymin=BioBreak-2*SE, ymax=BioBreak+2*SE),
  width=.2) +
  theme_bw() +
  theme(axis.text = element_text(size = rel(1.5), colour = "black"),
  axis.title=element_text(size=18))+
  labs(x = "Economy type",
  y="Species richness breakpoint (peak human density)")
```

Path models split into low and high species richness environments

```
#3. CONTROLLING FOR NPP: Split Path models
  #Nation state path model low species richness
    ========================================================

  CountryLow<-read.csv(file="CountryMeansEco2L.csv", header=T)

  model_Area<-'

  LnDENSITY~a2*npp+a3*biodiv
  biodiv~b1*npp

  IndNppBio:=a3*b1
  TotNppBio:=(a3*b1)+a2
  '
  model.fit<-sem(model_Area, se="bootstrap", data=CountryLow)
  summary(model.fit, standardized=TRUE, fit.measures=TRUE)
  semPaths(model.fit, whatLabels="std", layout="spring", nCharNodes=8)
  ParDiv<-parameterEstimates(model.fit)
  ParDiv
  xtable(ParDiv)
  fitTable<-fitMeasures(model.fit, c("cfi","rmsea","srmr", "aic"))
```

```
    write.table(fitTable, file = "fit1.csv", sep = ",")
    fit_base<-read.csv(file="fit1.csv", header=T)
    xtable(fit_base)

#======#Agricultural path model low speies richness========================================

    AgLow<-read.csv(file="AGPOP3EcoL.csv", header=T)

    model_Area<-'

    LnDENSITY~a2*npp+a3*biodiv
    biodiv~b1*npp

    IndNppBio:=a3*b1
    TotNppBio:=(a3*b1)+a2
    '
    model.fit<-sem(model_Area, se="bootstrap", data=AgLow)
    summary(model.fit, standardized=TRUE, fit.measures=TRUE)
    semPaths(model.fit, whatLabels="std", layout="spring", nCharNodes=8)
    fitMeasures(model.fit, c("cfi","rmsea","srmr", "aic"))
    ParDiv<-parameterEstimates(model.fit)
    ParDiv
    xtable(ParDiv)
    fitTable<-fitMeasures(model.fit, c("cfi","rmsea","srmr", "aic"))
    write.table(fitTable, file = "fit1.csv", sep = ",")
    fit_base<-read.csv(file="fit1.csv", header=T)
    xtable(fit_base)

    #Hunter-gatherer Low species richness
        ====================================================================
    HGLow<-read.csv(file="LowHG.csv", header=T)

    model_Area<-'

    LnDENSITY~a2*npp+a3*biodiv
    biodiv~b1*npp

    IndNppBio:=a3*b1
    TotNppBio:=(a3*b1)+a2
    '
    model.fit<-sem(model_Area, se="bootstrap", data=HGLow)
    summary(model.fit, standardized=TRUE, fit.measures=TRUE)
    semPaths(model.fit, whatLabels="std", layout="spring", nCharNodes=8)
    fitMeasures(model.fit, c("cfi","rmsea","srmr", "aic"))
    ParDiv<-parameterEstimates(model.fit)
    ParDiv
    xtable(ParDiv)
    fitTable<-fitMeasures(model.fit, c("cfi","rmsea","srmr", "aic"))
    write.table(fitTable, file = "fit1.csv", sep = ",")
    fit_base<-read.csv(file="fit1.csv", header=T)
    xtable(fit_base)

    ###Hunter-Gatherer High Species Richness Environments
        =========================================
    HGHow<-read.csv(file="HiHG.csv", header=T)

    model_Area<-'

    LnDENSITY~a2*npp+a3*biodiv+a4*pathos
    biodiv~b1*npp
    pathos~c1*biodiv

    IndNppBio:=a3*b1
    TotNppBio:=(a3*b1)+a2
    IndBio:=c1*a4
    TotBio:=(c1*a4)+a3
    '
```

```
model.fit<-sem(model_Area, se="bootstrap", data=HGHow)
summary(model.fit, standardized=TRUE, fit.measures=TRUE)
semPaths(model.fit,"std", layout="spring", nCharNodes=8)
ParDiv<-parameterEstimates(model.fit)
ParDiv
xtable(ParDiv)
fitTable<-fitMeasures(model.fit, c("cfi","rmsea","srmr", "aic"))
write.table(fitTable, file = "fit1.csv", sep = ",")
fit_base<-read.csv(file="fit1.csv", header=T)
xtable(fit_base)

###Agricultural path model High biodiversity environments
    ========================================

AgHow<-read.csv(file="AGPOP3EcoH.csv", header=T)

model_Area<-'

LnDENSITY~a2*npp+a3*biodiv+a4*pathos
biodiv~b1*npp
pathos~c1*biodiv+c2*npp

IndNppBio:=a3*b1
TotNppBio:=(a3*b1)+a2
IndBio:=c1*a4
TotBio:=(c1*a4)+a3
'
model.fit<-sem(model_Area, se="bootstrap", data=AgHow)
summary(model.fit, standardized=TRUE, fit.measures=TRUE)
semPaths(model.fit,"std", layout="spring", nCharNodes=8)
ParDiv<-parameterEstimates(model.fit)
ParDiv
xtable(ParDiv)
fitTable<-fitMeasures(model.fit, c("cfi","rmsea","srmr", "aic"))
write.table(fitTable, file = "fit1.csv", sep = ",")
fit_base<-read.csv(file="fit1.csv", header=T)
xtable(fit_base)

####Nation sates high biodiveristy environments
CountryHow<-read.csv(file="CountryMeansEco2H.csv", header=T)

model_Area<-'

LnDENSITY~a2*npp+a3*biodiv+a4*pathos
biodiv~b1*npp
pathos~c1*biodiv

IndNppBio:=a3*b1
TotNppBio:=(a3*b1)+a2
IndBio:=c1*a4
TotBio:=(c1*a4)+a3
'

model.fit<-sem(model_Area, se="bootstrap", data=CountryHow)
summary(model.fit, standardized=TRUE, fit.measures=TRUE)
semPaths(model.fit, "std", layout="spring", nCharNodes=8)
ParDiv<-parameterEstimates(model.fit)
ParDiv
fitTable<-fitMeasures(model.fit, c("cfi","rmsea","srmr", "aic"))
write.table(fitTable, file = "fit1.csv", sep = ",")
fit_base<-read.csv(file="fit1.csv", header=T)
xtable(fit_base)
```

Path models of the effect of distance from optimal species richness on social complexity

```
#4.EFFECTS OF DISTANCE FROM PEAK RICHNESS ON COMPLEXITY VIA DENSITY

  #Modern nation states
   keep3<-read.csv(file="CountryMeansEco2.csv", header=T)


   model_Pop<-'

   ECI~ a1*DivDiff+a2*LnDENSITY
   LnDENSITY~b1*DivDiff

   IndDiv:=a2*b1
   TotDiv:=(a2*b1)+a1
   '
   model.fit<-sem(model_Pop, se="bootstrap", data=keep3)
   summary(model.fit, standardized=TRUE, fit.measures=TRUE)
   semPaths(model.fit, whatLabels="std", layout="spring", nCharNodes=8)
   ParDiv<-parameterEstimates(model.fit)
   ParDiv
   xtable(ParDiv)
   fitTable<-fitMeasures(model.fit, c("cfi","rmsea","srmr", "aic"))
   write.table(fitTable, file = "fit1.csv", sep = ",")
   fit_base<-read.csv(file="fit1.csv", header=T)
   xtable(fit_base)

 #Hunter-gathere societies--Main analysis is on class2, binary ordinal
   keep2<-read.csv(file="HGFEM4R.csv", header=T)

   #order potential ordinal response variables for estimating social complexity
   keep2[,c("headm",
            "class","class2")] <-
     lapply(keep2[,c("class", "class2")], ordered)


   model_Pop<-'

   class2~ a1*DivDiff+a2*LnDENSITY
   LnDENSITY~b1*DivDiff

   IndDiv:=a2*b1
   TotDiv:=(a2*b1)+a1
   '
   model.fit<-sem(model_Pop, se="bootstrap",estimator = "DWLS", data=keep2)
   #model.fit<-sem(model_Pop, data=keep2)
   summary(model.fit, standardized=TRUE, fit.measures=TRUE)
   semPaths(model.fit, whatLabels="std", layout="spring", nCharNodes=8)
   ParDiv<-parameterEstimates(model.fit)
   ParDiv
   xtable(ParDiv)
   fitTable<-fitMeasures(model.fit, c("cfi","rmsea","srmr", "aic"))
   write.table(fitTable, file = "fit1.csv", sep = ",")
   fit_base<-read.csv(file="fit1.csv", header=T)
   xtable(fit_base)

   #Agrculturalist---Main analysis is on class2--binary ordinal
   keep<-read.csv(file="AGPOP3Eco.csv", header=T)

   #order potential ordinal response variables for estimating social complexity
   keep[,c("class2",
           "class")] <-
     lapply(keep[,c("class2", "class")], ordered)


   model_Pop<-'
```

```
    class2~ a1*DivDiff+a2*LnDENSITY
    LnDENSITY~b1*DivDiff

    IndDiv:=a2*b1
    TotDiv:=(a2*b1)+a1
    '
    model.fit<-sem(model_Pop, se="bootstrap", estimator = "DWLS", data=keep)
    #model.fit<-sem(model_Pop, data=keep)
    summary(model.fit, standardized=TRUE, fit.measures=TRUE)
    semPaths(model.fit, whatLabels="std", layout="spring", nCharNodes=8)
    ParDiv<-parameterEstimates(model.fit)
    ParDiv
    xtable(ParDiv)
    fitTable<-fitMeasures(model.fit, c("cfi","rmsea","srmr", "aic"))
    write.table(fitTable, file = "fit1.csv", sep = ",")
    fit_base<-read.csv(file="fit1.csv", header=T)
    xtable(fit_base)
```

Correcting for spatial autocorrelation

```
##Updated spatial correct function to account for
#infinite values bug in the inverse distance matrix.
##First, run the spatial correction function
spatialCorrectL2 <- function(obj,xvar,yvar, alpha=0.05){
  require(ape)
  require(fossil)
  if(class(obj)=="lm" || class(obj)=="glm"){
    resids <- residuals(obj)
    distMat <- as.matrix(earth.dist(cbind(xvar, yvar), dist=F))
    distsInv <- 1/distMat
    distsInv[is.infinite(distsInv)] <- 0
    diag(distsInv) <- 0
    mi <- Moran.I(resids, distsInv)
    if(mi$p.value > alpha){
      mi$n.eff <- length(resids)
    }else{
      mi$n.eff <- length(resids)*(1-mi$observed)/(1+mi$observed)
    }
    v <- diag(vcov(obj))
    n <- length(resids)
    v2 <- v*n/mi$n.eff
    ret <- data.frame(Estimate=coef(obj),
                      n.eff = mi$n.eff, Std.err = sqrt(v2))
    ret[["Z-value"]] <- ret$Estimate/ret$Std.err
    ret[["P(>|z|)"]] <- 2*pnorm(abs(ret[["Z-value"]]), lower.tail=F)
    list1 <- list(mi, ret)
    names(list1) <- c(names(obj$model[1]),names(obj$model[1]))
    return(list1)
  } else {
    if(class(obj)=="list"){
      sp.corr.list <- list()
      for(i in 1:length(obj)){
        resids <- residuals(obj[[i]])
        distMat <- as.matrix(earth.dist(cbind(xvar, yvar),dist=F))
        distsInv <- 1/distMat
        distsInv[is.infinite(distsInv)] <- 0
        diag(distsInv) <- 0
        mi <- Moran.I(resids, distsInv)
        if(mi$p.value > alpha){
          mi$n.eff <- length(resids)
        }else{
          mi$n.eff <- length(resids)*(1-mi$observed)/(1+mi$observed)
        }
```

```
          v <- diag(vcov(obj[[i]]))
          n <- length(resids)
          v2 <- v*n/mi$n.eff
          ret <- data.frame(Estimate=coef(obj[[i]]),
                            n.eff = mi$n.eff, Std.err = sqrt(v2))
          ret[["Z-value"]] <- ret$Estimate/ret$Std.err
          ret[["P(>|z|)"]] <- 2*pnorm(abs(ret[["Z-value"]]), lower.tail=F)
          list1 <- list(mi, ret)
          names(list1) <- c(names(obj[[i]]$model[1]),names(obj[[i]]$model[1]))
          sp.corr.list[[i]] <- list1
        }
        return(sp.corr.list)
      } else {
        "Warning: object is not lm or glm model object or list of
        such model objects"
      }
  }}

#1. Quadratic regressions
#A. #All Societies
AllSoc<-read.csv(file="MainFinal.csv", header=T)

modelList1 <- list(
  lm(DENSITY ~ biodiv + biodiv2+factor(ID), data = AllSoc))
sem.coefs(modelList1, data=AllSoc, standardize = "none")

#Run spatial correction function
sem.model.fits(modelList1)
spatialCorrectL2(modelList1,AllSoc$Longitude,AllSoc$Latitude)[[1]][[2]]

#   B.##Nation States Quadratic
keep3<-read.csv(file="CountryMeansEco2.csv", header=T)

modelList1 <- list(
  lm(DENSITY ~ biodiv + biodiv2, data = keep3))
sem.coefs(modelList1, data=keep3, standardize = "none")
sem.model.fits(modelList1)

#Run spatial correction function
spatialCorrectL2(modelList1,keep3$Longitude,keep3$Latitude)[[1]][[2]]

##C. Agricultural societies quadratic.
keep<-read.csv(file="AGPOP3Eco.csv", header=T)

modelList1 <- list(
  lm(DENSITY ~ biodiv + biodiv2, data = keep))
sem.coefs(modelList1, data=data.low, standardize = "none")
sem.model.fits(modelList1)

#run spatial correction
spatialCorrectL2(modelList1,keep$Longitude,keep$Latitude)[[1]][[2]]


##D. Hunter-gatherer societies quadratic
keep2<-read.csv(file="HGFEM4R.csv", header=T)

###Quadratic Model
modelList1 <- list(
  lm(DENSITY ~ biodiv + biodiv2, data = keep2))
sem.coefs(modelList1, data=keep2, standardize = "none")
sem.model.fits(modelList1)
#run spatial correct
spatialCorrectL2(modelList1,keep2$Longitude,keep2$Latitude)[[1]][[2]]


#2. Spatial correction on population density path models in
#High and Low species richness environments
```

```
###A. NATION STATE HIGH SPECIES RICHNESS
data.low <-read.csv(file="CountryMeansEco2H.csv", header=T)

modelList1 <- list(
  lm(LnDENSITY ~ biodiv + npp+pathos, data = data.low),
  lm(biodiv ~ npp, data=data.low),
  lm(pathos~biodiv,data=data.low))

sem.coefs(modelList1, data=data.low, standardize = "none")
sem.coefs(modelList1, data=data.low, standardize = "scale")
sem.fit(modelList1, data=data.low, .progressBar = FALSE)

sem.model.fits(modelList1)
spatialCorrectL2(modelList1,data.low$Longitude,data.low$Latitude)[[1]][[2]]
spatialCorrectL2(modelList1,data.low$Longitude,data.low$Latitude)[[2]][[2]]
spatialCorrectL2(modelList1,data.low$Longitude,data.low$Latitude)[[3]][[2]]

###B.HG HIGH DIVERSITY
data.low <-read.csv(file="HiHG.csv", header=T)

modelList1 <- list(
  lm(LnDENSITY ~ biodiv + npp+pathos, data = data.low),
  lm(biodiv ~ npp, data=data.low),
  lm(pathos~biodiv,data=data.low))

sem.coefs(modelList1, data=data.low, standardize = "none")
sem.coefs(modelList1, data=data.low, standardize = "scale")
sem.fit(modelList1, data=data.low, .progressBar = FALSE)

sem.model.fits(modelList1)
spatialCorrectL2(modelList1,data.low$Longitude,data.low$Latitude)[[1]][[2]]
spatialCorrectL2(modelList1,data.low$Longitude,data.low$Latitude)[[2]][[2]]
spatialCorrectL2(modelList1,data.low$Longitude,data.low$Latitude)[[3]][[2]]

#####C.Agriculturalist high species richness environments
data.low <-read.csv(file="AGPOP3EcoH.csv", header=T)

modelList1 <- list(
  lm(LnDENSITY ~ biodiv + npp+pathos, data = data.low),
  lm(biodiv ~ npp, data=data.low),
  lm(pathos~biodiv,data=data.low))

sem.coefs(modelList1, data=data.low, standardize = "none")
sem.coefs(modelList1, data=data.low, standardize = "scale")
sem.fit(modelList1, data=data.low, .progressBar = FALSE)

sem.model.fits(modelList1)
spatialCorrectL2(modelList1,data.low$Longitude,data.low$Latitude)[[1]][[2]]
spatialCorrectL2(modelList1,data.low$Longitude,data.low$Latitude)[[2]][[2]]
spatialCorrectL2(modelList1,data.low$Longitude,data.low$Latitude)[[3]][[2]]

#D. Nationstate Low species richness
CountryLow<-read.csv(file="CountryMeansEco2L.csv", header=T)

modelList1 <- list(
  lm(LnDENSITY ~ biodiv + npp, data =CountryLow),
  lm(biodiv ~ npp, data=CountryLow))

sem.coefs(modelList1, data=CountryLow, standardize = "none")
sem.coefs(modelList1, data=CountryLow, standardize = "scale")
sem.fit(modelList1, data=CountryLow, .progressBar = FALSE)

sem.model.fits(modelList1)
spatialCorrectL2(modelList1,CountryLow$Longitude,CountryLow$Latitude)[[1]][[2]]
spatialCorrectL2(modelList1,CountryLow$Longitude,CountryLow$Latitude)[[2]][[2]]

#E. Agriculturalists low species richness
AgLow<-read.csv(file="AGPOP3EcoL.csv", header=T)
```

```r
modelList1 <- list(
  lm(LnDENSITY ~ biodiv + npp, data =AgLow),
  lm(biodiv ~ npp, data=AgLow))

sem.coefs(modelList1, data=AgLow, standardize = "none")
sem.coefs(modelList1, data=AgLow, standardize = "scale")
#sem.fit(modelList1, data=AgLow, .progressBar = FALSE)

sem.model.fits(modelList1)
spatialCorrectL2(modelList1,AgLow$Longitude,AgLow$Latitude)[[1]][[2]]
spatialCorrectL2(modelList1,AgLow$Longitude,AgLow$Latitude)[[2]][[2]]


#F. Hunter-gatherers low species richness
HGLow<-read.csv(file="LowHG.csv", header=T)

modelList1 <- list(
  lm(LnDENSITY ~ biodiv + npp, data =HGLow),
  lm(biodiv ~ npp, data=HGLow))

sem.coefs(modelList1, data=HGLow, standardize = "none")
sem.coefs(modelList1, data=HGLow, standardize = "scale")
#sem.fit(modelList1, data=AgLow, .progressBar = FALSE)

sem.model.fits(modelList1)
spatialCorrectL2(modelList1,HGLow$Longitude,HGLow$Latitude)[[1]][[2]]
spatialCorrectL2(modelList1,HGLow$Longitude,HGLow$Latitude)[[2]][[2]]

#2. COMPLEXITY SPATIAL CORRECTION
NS.complex <-read.csv(file="CountryMeansEco2.csv", header=T)

modelList1 <- list(
  lm(ECI ~ DivDiff + LnDENSITY, data = NS.complex),
  lm(LnDENSITY ~ DivDiff, data=NS.complex))

sem.coefs(modelList1, data=NS.complex, standardize = "none")
sem.coefs(modelList1, data=NS.complex, standardize = "scale")
sem.fit(modelList1, data=NS.complex, .progressBar = FALSE)

sem.model.fits(modelList1)
spatialCorrectL2(modelList1,NS.complex$Longitude,NS.complex$Latitude)[[1]][[2]]
spatialCorrectL2(modelList1,NS.complex$Longitude,NS.complex$Latitude)[[2]][[2]]


###AGRICULTUALISTS
Ag.complex <-read.csv(file="AGPOPClass.csv", header=T)

modelList1 <- list(
  glm(class2 ~ DivDiff + LnDENSITY, family=binomial(link="logit"),
  data = Ag.complex),
  lm(LnDENSITY ~ DivDiff, data=Ag.complex))

sem.coefs(modelList1, data=Ag.complex, standardize = "none")
#sem.coefs(modelList1, data=data.Ag, standardize = "scale")
sem.fit(modelList1, data=Ag.complex, .progressBar = FALSE)

sem.model.fits(modelList1)
spatialCorrectL2(modelList1,Ag.complex$Longitude,Ag.complex$Latitude)[[1]][[2]]
spatialCorrectL2(modelList1,Ag.complex$Longitude,Ag.complex$Latitude)[[2]][[2]]

###HUNTER-GATHERERS
Hg.complex <-read.csv(file="HGFEM3R.csv", header=T)

modelList1 <- list(
  glm(class2 ~ DivDiff + LnDENSITY, family=binomial(link="logit"),
  data = Hg.complex),
  lm(LnDENSITY ~ DivDiff, data=Hg.complex))
```

```
sem.coefs(modelList1, data=Hg.complex, standardize = "none")
#sem.coefs(modelList1, data=Hg.complex, standardize = "scale")
sem.fit(modelList1, data=Hg.complex, .progressBar = FALSE)

sem.model.fits(modelList1)
spatialCorrectL2(modelList1,Hg.complex$Longitude,Hg.complex$Latitude)[[1]][[2]]
spatialCorrectL2(modelList1,Hg.complex$Longitude,Hg.complex$Latitude)[[2]][[2]]
```

```
#5. SUPPORTING ANALYSES ==============================================================.
 # A. #Hunter-gathere societies--Supporting analysis of
#and kinship augmentation,
   #which is a measure of social networks created to augment kinship and
   #possibly to reduce risk
   keep2<-read.csv(file="HGFEM4R.csv", header=T)

   #order potential ordinal response variables for estimating social complexity
   keep2[,c("headm",
            "augmen2","class")] <-
     lapply(keep2[,c( "augmen2", "class")], ordered)


   model_Pop<-'

   class~ a1*DivDiff+a2*LnDENSITY
   LnDENSITY~b1*DivDiff

   IndDiv:=a2*b1
   TotDiv:=(a2*b1)+a1
   '
   model.fit<-sem(model_Pop, se="bootstrap",estimator = "DWLS", data=keep2)
   #model.fit<-sem(model_Pop, data=keep2)
   summary(model.fit, standardized=TRUE, fit.measures=TRUE)
   #semPaths(model.fit, whatLabels="std", layout="spring", nCharNodes=8)
   ParAug<-parameterEstimates(model.fit)
   ParAug
   xtable(ParAug)

 ##Alternative complexity variable  Augmen2:
 #The number of weak ties that augment kinship
   model_Pop<-'

   augmen2~ a1*DivDiff+a2*LnDENSITY
   LnDENSITY~b1*DivDiff

   IndDiv:=a2*b1
   TotDiv:=(a2*b1)+a1
   '
   model.fit<-sem(model_Pop, se="bootstrap",estimator = "DWLS", data=keep2)
   #model.fit<-sem(model_Pop, data=keep2)
   summary(model.fit, standardized=TRUE, fit.measures=TRUE)
   #semPaths(model.fit, whatLabels="std", layout="spring", nCharNodes=8)
   ParAug<-parameterEstimates(model.fit)
   ParAug
   xtable(ParAug)



   #Agrculturalist---supporting analysis on Class variable
   #(three ordered classes instead of two)
   keep<-read.csv(file="AGPOP3Eco.csv", header=T)

   #order potential ordinal response variables for estimating social complexity
   keep[,c("class2",
           "class")] <-
     lapply(keep[,c("class2", "class")], ordered)
```

```
   model_Pop<-'

   class~ a1*DivDiff+a2*LnDENSITY
   LnDENSITY~b1*DivDiff

   IndDiv:=a2*b1
   TotDiv:=(a2*b1)+a1
   '
   model.fit<-sem(model_Pop, se="bootstrap", estimator = "DWLS", data=keep)
   #model.fit<-sem(model_Pop, data=keep)
   summary(model.fit, standardized=TRUE, fit.measures=TRUE)
 # semPaths(model.fit, whatLabels="std", layout="spring", nCharNodes=8)
   ParAug2<-parameterEstimates(model.fit)
   ParAug2
   xtable(ParAug2)
```

References

Tallavaara, M., J. T. Eronen, and M. Luoto

  2017. Supporting data and script for "productivity, biodiversity, and pathogens influence the global hunter-gatherer population density" (Tallavaara et al. pnas 2018). *https://doi.org/10.5281/zenodo.1167852.*