**SUPPLEMENTARY ELECTRONIC MATERIAL**

**ARDEOLA 68(1)**

# WINTER BIRD RICHNESS DISTRIBUTION IN THE SOUTH-WESTERN PALEARCTIC: CURRENT PATTERNS AND POTENTIAL CHANGES

## DISTRIBUCIÓN DE LA RIQUEZA INVERNAL DE AVES EN EL PALEÁRTICO OCCIDENTAL: PATRONES ACTUALES Y POSIBLES CAMBIOS

José Luis TELLERÍA[1 *], Guillermo FANDOS[1, 2] and Javier FERNÁNDEZ-LÓPEZ[1, 3]

[1] Department of Biodiversity, Ecology and Evolution, Facultad de Ciencias Biológicas, Universidad Complutense, 28040 Madrid, Spain.

[2] Geography Department, Humboldt-Universität zu Berlin, Berlin, Germany.

[3] Real Jardín Botánico de Madrid, Plaza de Murillo 2, 28014 Madrid, Spain.

* Corresponding author: telleria@bio.ucm.es

**Appendix 4.** R script for analyses [*Guión en R de los análisis*].

# Winter bird richness distribution

José Luis Tellería, Guillermo Fandos and Javier Fernández-López

January 2020

## Load packages

```
library(nlme)
library(MuMIn)
library(ape)
```

First of all, download the following table and prepare for the analysis

```
data_com <- read.csv("species_richness.csv")
head(data_com)

##   ID longitude latitude total_abundance total_richness abund_6sp
## 1  1   -5.3581  43.5310              14              6         2
## 2  2   -5.3567  43.5283              19             10         3
## 3  3   -5.3535  43.5231              70             16         7
## 4  4   -5.3494  43.5212              36             13        10
## 5  5   -5.3202  43.5139             119             15        16
## 6  6   -5.3082  43.5110               8              3         1
##   richness_6sp habitat woodcover temperature precipitation
## 1            1   0.043     2.168       0.827         2.097
## 2            1   2.119     2.168       0.827         2.097
## 3            4   0.792     1.612       0.657         2.413
## 4            3   0.427     1.413       0.676         2.431
## 5            4  -0.233     1.532       0.714         2.082
## 6            1   1.438     1.532       0.705         2.058

names(data_com)

##  [1] "ID"              "longitude"       "latitude"
##  [4] "total_abundance" "total_richness"  "abund_6sp"
##  [7] "richness_6sp"    "habitat"         "woodcover"
## [10] "temperature"     "precipitation"

options(na.action = "na.omit")
options(na.action = "na.fail")

#Let's clean the table
data_com <- na.omit(data_com)
#we eliminate duplicate records to avoid problems with distances 0
duplicados<-duplicated(data_com[ , c("latitude", "longitude")])
#How many duplicates are there?
length(duplicados[duplicados==TRUE])

## [1] 16
```

```r
#select the non-duplicates
data_com <- data_com[!duplicados, ]
#we check that there are no duplicates
duplicados<-duplicated(data_com[ , c("latitude", "longitude")])
length(duplicados[duplicados==TRUE])

## [1] 0
```

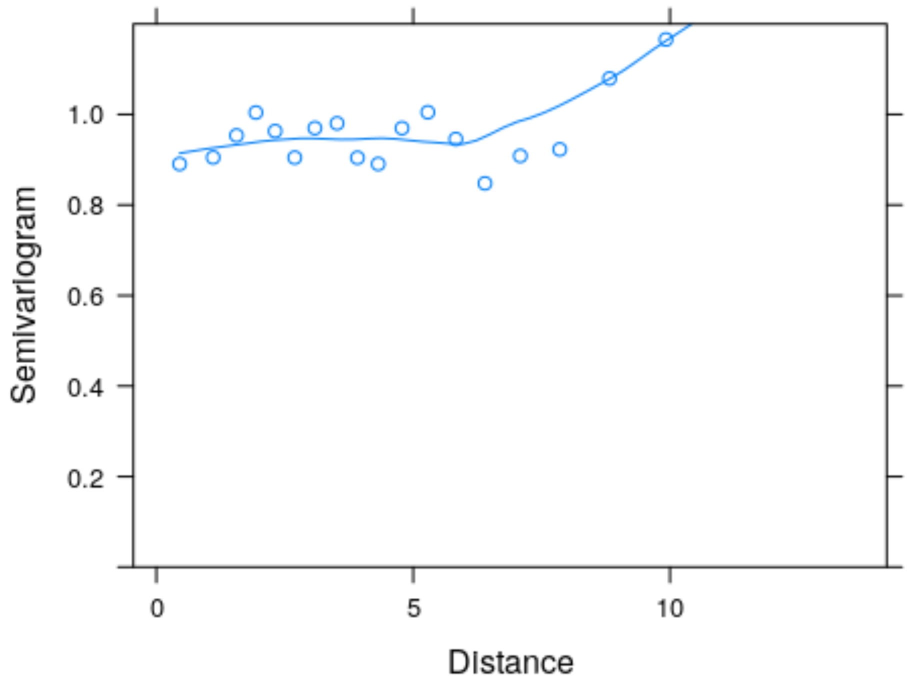
## Moran's Test

```r
data_com.dists <- as.matrix(dist(cbind(data_com$longitude, data_com$latitude)))
data_com.dists.inv <- 1/data_com.dists
diag(data_com.dists.inv) <- 0
data_com.dists.inv[1:5, 1:5]

##           1         2         3         4        5
## 1   0.00000 328.79797 109.38929  76.30920 24.05056
## 2 328.79797   0.00000 163.78045  98.19980 25.48559
## 3 109.38929 163.78045   0.00000 221.29527 28.94565
## 4  76.30920  98.19980 221.29527   0.00000 33.22406
## 5  24.05056  25.48559  28.94565  33.22406  0.00000

Moran.I(data_com$total_richness, data_com.dists.inv)

## $observed
## [1] 0.237838
##
## $expected
## [1] -0.001162791
##
## $sd
## [1] 0.01318638
##
## $p.value
## [1] 0
```

## Models

First we are going to run the model without spatial correlation correction.

```r
m1 <- gls(log1p(total_richness) ~
            temperature + precipitation + woodcover + habitat, data = data_com)
vario1 <- Variogram(m1, form = ~ latitude + longitude, resType = "pearson")
plot(vario1, smooth = TRUE, ylim = c(0, 1.2))
```




## Spatial correlation correction with six spatial correlation structures

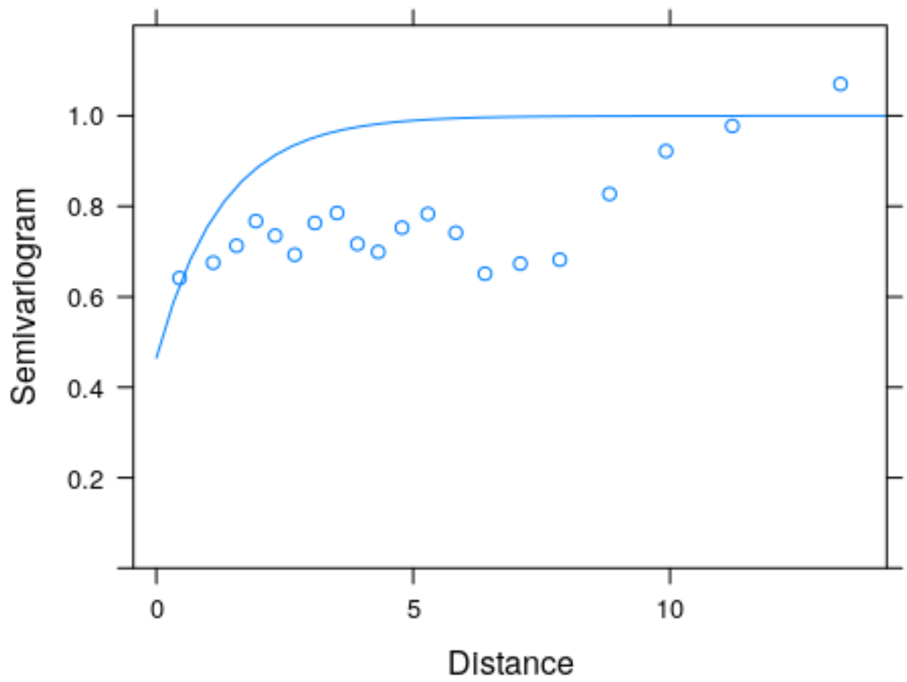
Function using the correlation argument. We fit our model using different correlation structures, and we then use AIC to choose the best model (we also compare the initial model without correlation).

```
m2 <- gls(log1p(total_richness) ~ temperature + precipitation + woodcover +
            habitat, correlation = corExp(form =  ~ latitude + longitude ,
                                  nugget = TRUE), data = data_com)
m3 <- gls(log1p(total_richness)  ~ temperature + precipitation + woodcover +
            habitat, correlation = corGaus(form =  ~ latitude + longitude,
                                   nugget = TRUE), data = data_com)
m4 <- gls(log1p(total_richness) ~ temperature + precipitation + woodcover +
            habitat, correlation = corSpher(form = ~latitude + longitude,
                                    nugget = TRUE), data = data_com)
m5 <- gls(log1p(total_richness)  ~ temperature + precipitation + woodcover +
            habitat, correlation = corLin(form = ~latitude + longitude,
                                  nugget = TRUE), data = data_com)
m6 <- gls(log1p(total_richness) ~ temperature + precipitation + woodcover +
            habitat, correlation = corRatio(form = ~latitude + longitude,
                                    nugget = TRUE), data = data_com)

#Check what model has best AIC
AIC(m1, m2, m3, m4, m5, m6)

##    df      AIC
## m1  6 1556.180
## m2  8 1365.932
```

```
## m3   8 1560.180
## m4   8 1560.180
## m5   8 1560.180
## m6   8 1377.639
```

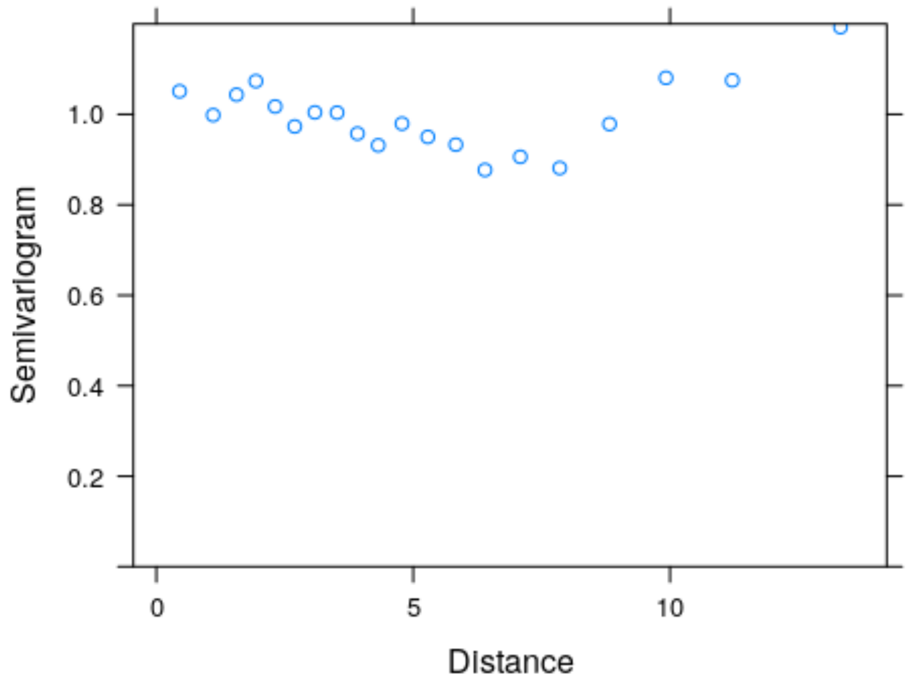
In this case: m2 (put the best model below) correlation structure seems to be the most appropriate one (the corRation structure is a close runner-up). We can also plot the fitted variogram to check if the fit is OK.

```
vario2 <- Variogram(m2, form = ~latitude + longitude, resType = "pearson")
plot(vario2, smooth = FALSE, ylim = c(0, 1.2))
```




As a last check, we plot a sample variogram of the normalized residuals. These residuals are standardized residuals pre-multiplied by the inverse square-root factor of the estimated error correlation matrix. Hence we shouldn't see any trend in this new variogram if the residual spatial autocorrelation was properly accounted for.

```
vario2 <- Variogram(m2, form = ~latitude + longitude, resType = "normalized",
                    maxDist = 30)
plot(vario2, smooth = FALSE, ylim = c(0, 1.2))
```




```
summary(m2)

## Generalized least squares fit by REML
##   Model: log1p(total_richness) ~ temperature + precipitation + woodcover +
habitat
##   Data: data_com
##        AIC      BIC    logLik
##   1365.932 1403.95 -674.9658
##
## Correlation Structure: Exponential spatial correlation
##  Formula: ~latitude + longitude
##  Parameter estimate(s):
##     range    nugget
## 1.2714653 0.4647591
##
## Coefficients:
##                    Value  Std.Error   t-value p-value
## (Intercept)    1.6793926 0.14469079 11.606769  0.0000
## temperature    0.3383195 0.05071851  6.670534  0.0000
## precipitation  0.1035088 0.06974525  1.484099  0.1382
## woodcover     -0.0196839 0.03214560 -0.612337  0.5405
## habitat        0.1247302 0.02347186  5.314029  0.0000
##
##  Correlation:
##               (Intr) tmprtr prcptt wodcvr
## temperature   -0.109
## precipitation  0.069  0.275
## woodcover     -0.007  0.068 -0.202
## habitat       -0.009  0.039 -0.048 -0.325
##
## Standardized residuals:
```

```
##        Min         Q1        Med         Q3        Max
## -2.9058758 -0.2978511  0.2706680  0.8219704  2.3386984
##
## Residual standard error: 0.6767225
## Degrees of freedom: 861 total; 856 residual
```