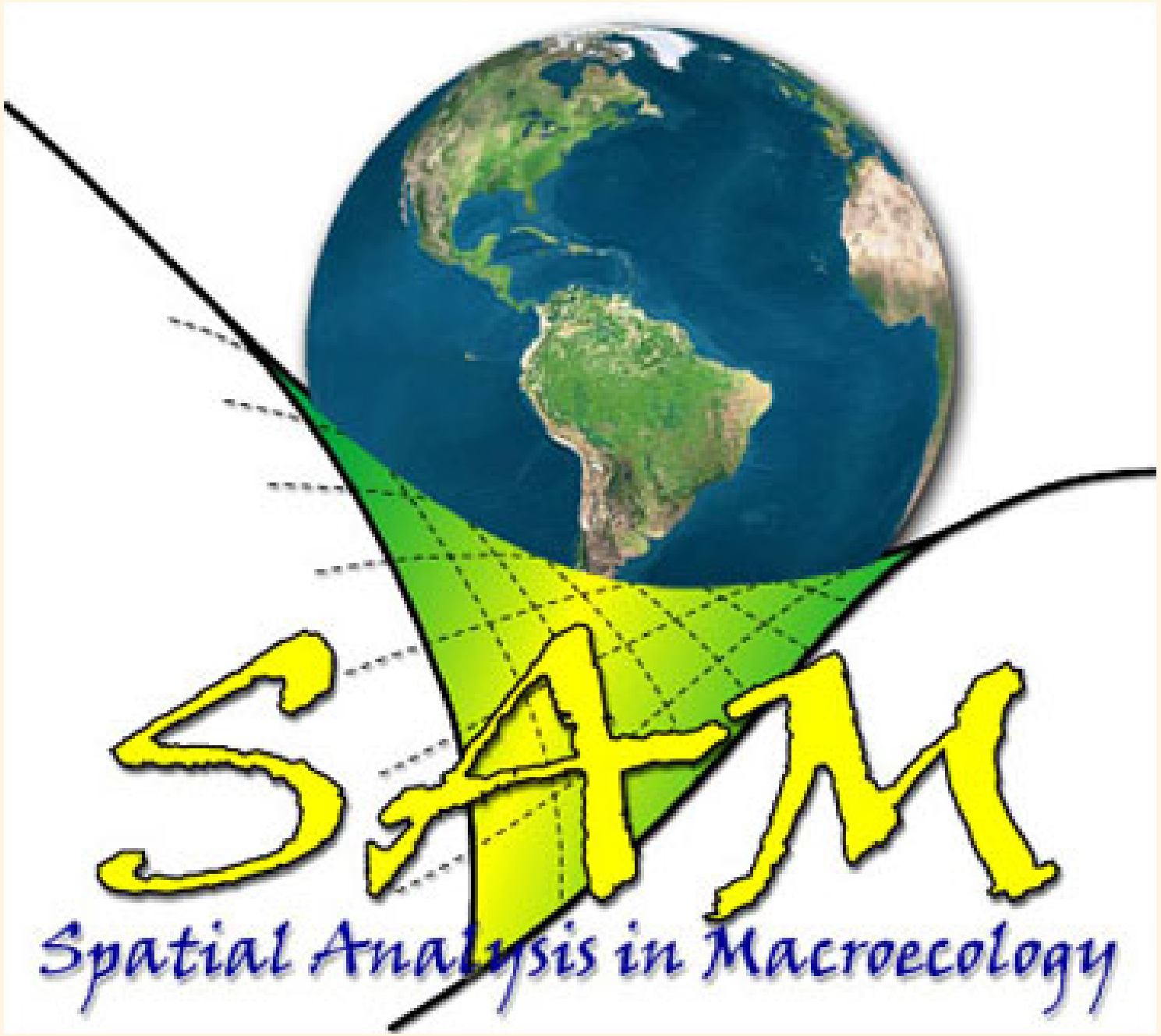
SPATIAL MODELLING

DAVID ORME

SPATIAL MODELLING TOOLS

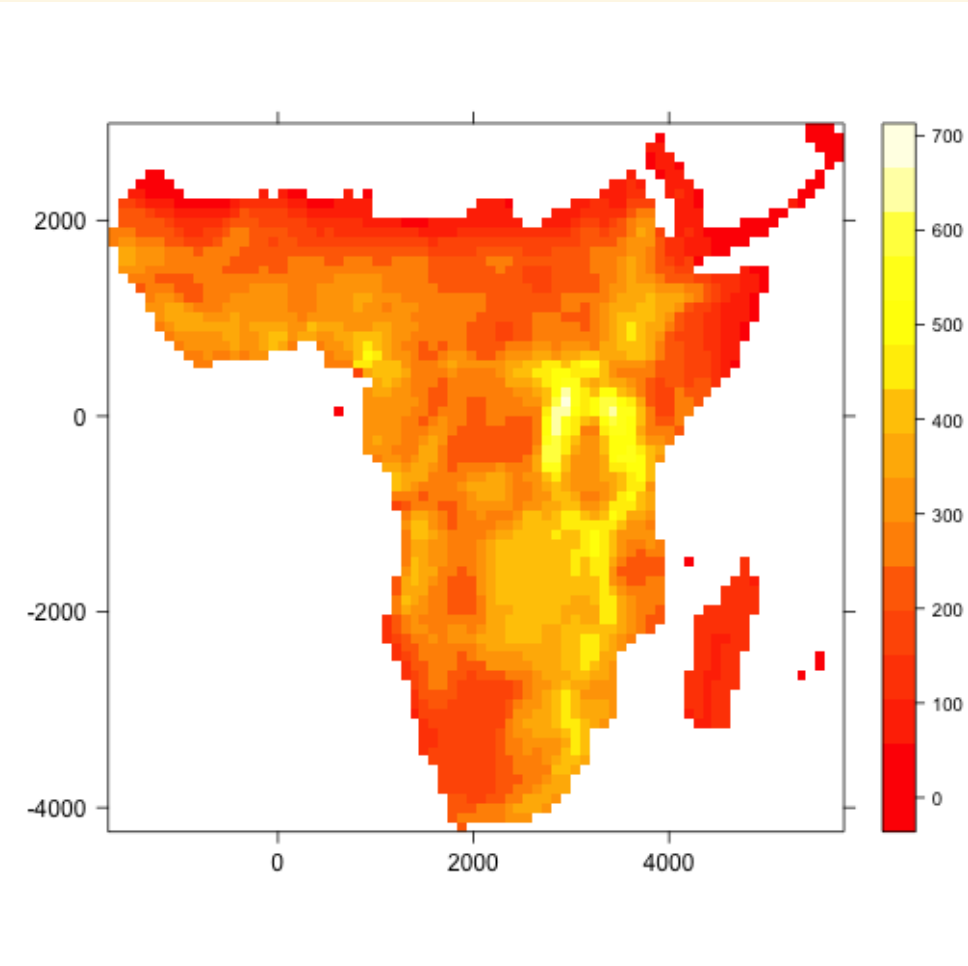
- The examples presented here use R
- Another excellent program with a nice GUI interface:
  - Spatial Analysis in Macroecology
  - http://www.ecoevol.ufg.br/sam/


SAM
Spatial Analysis in Macroecology

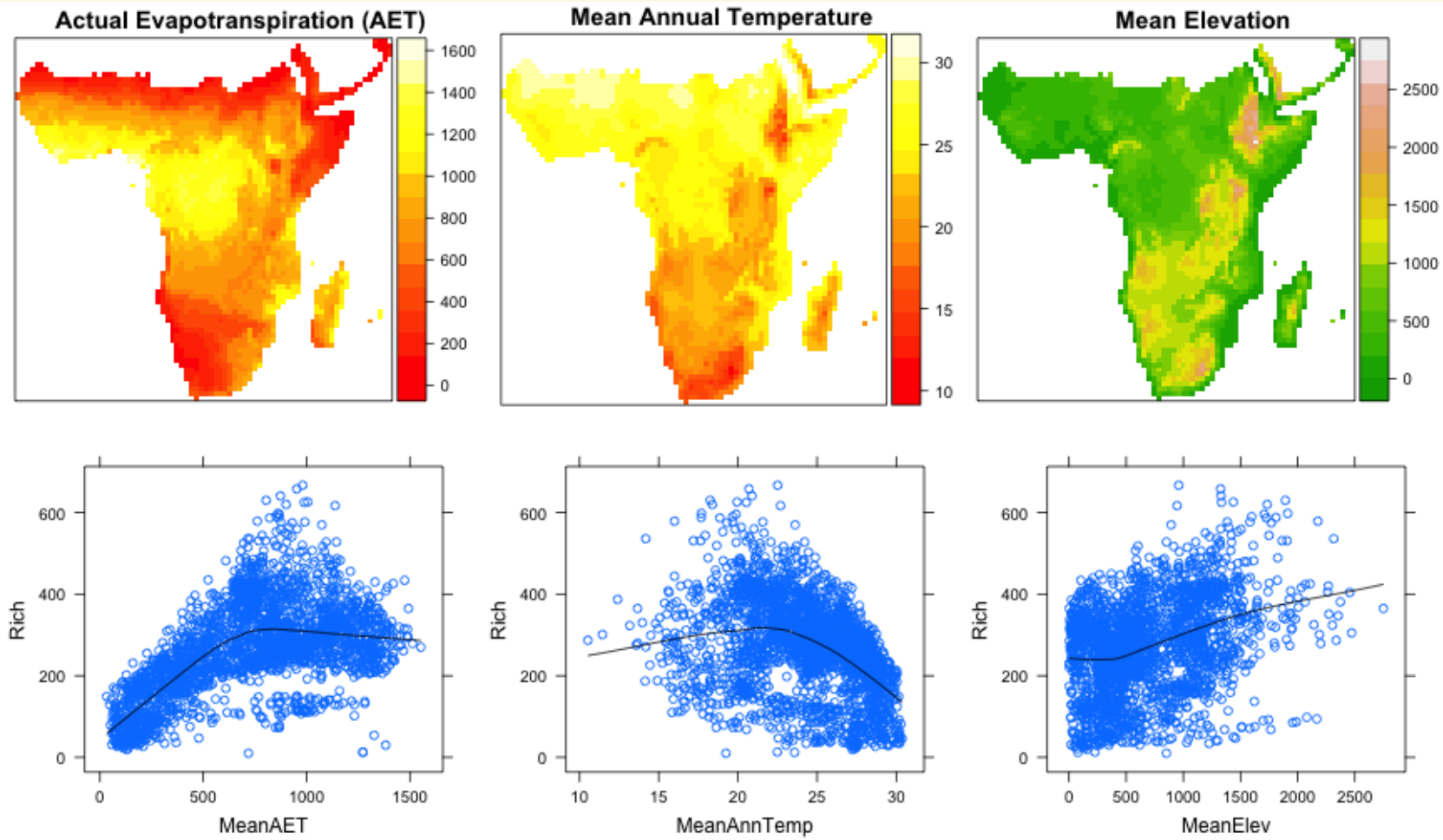
OVERVIEW

- Example data: Afrotropical bird diversity
- Naive models
- Describing spatial autocorrelation
- Accounting for spatial autocorrelation

AFROTROPICAL BIRD SPECIES RICHNESS



EXPLANATORY VARIABLES



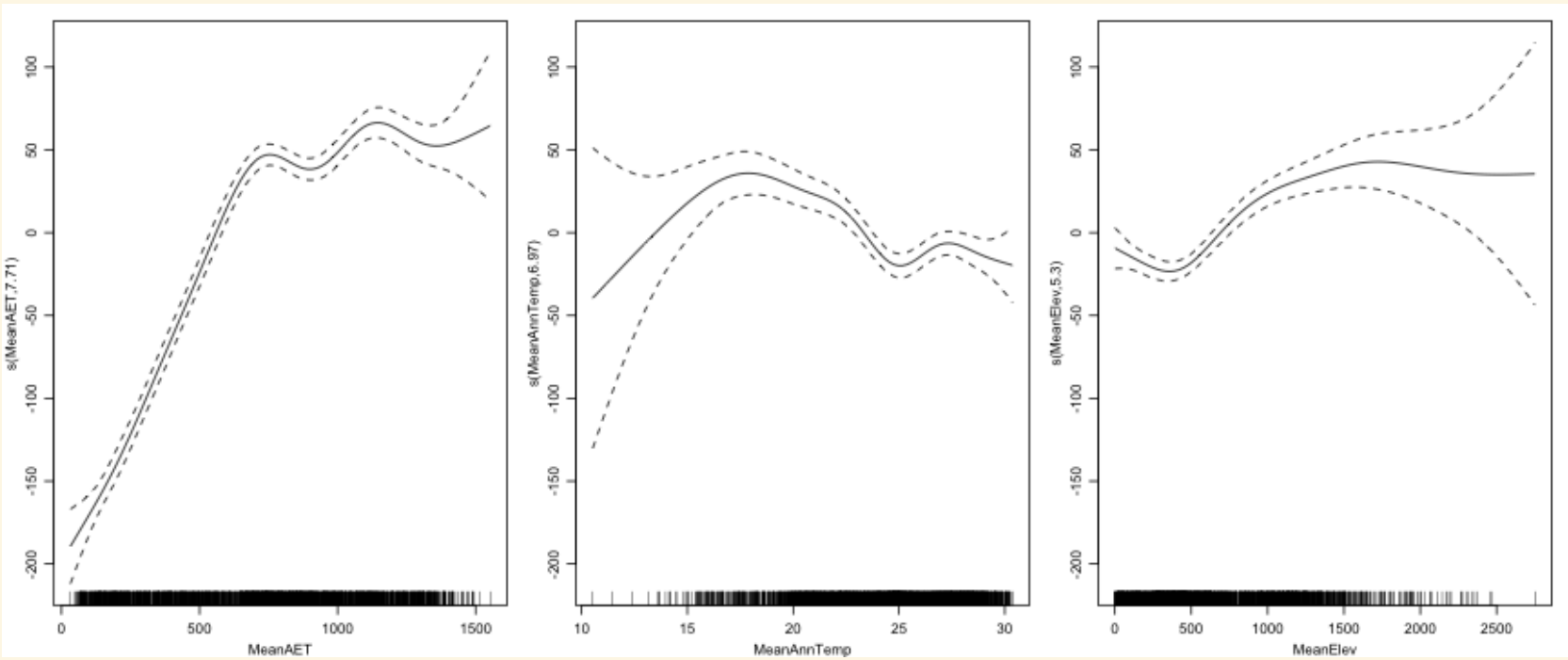
A SIMPLE LINEAR MODEL

Richness ~ AET + Temperature + Elevation

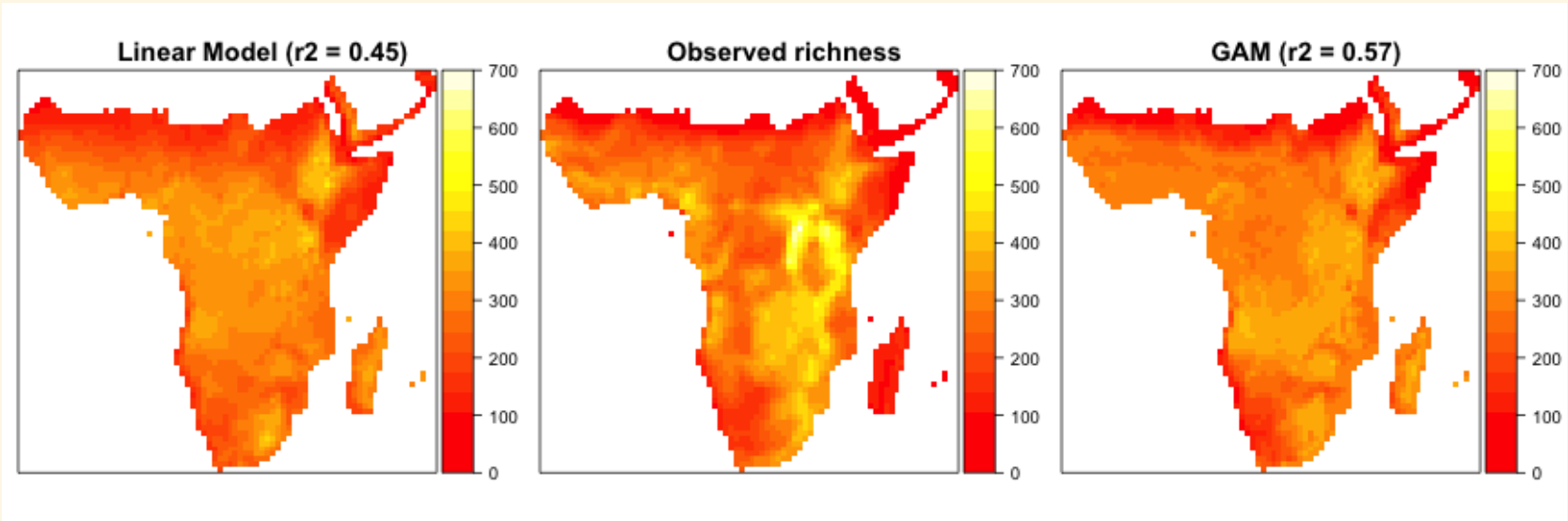
	Est	SE	t	p
(Intercept)	189.45	21.33	8.88	< 0.001
MeanAET	0.18	0.00	37.34	< 0.001
MeanAnnTemp	-4.18	0.72	-5.79	< 0.001
MeanElev	0.08	0.01	13.85	< 0.001

A SIMPLE GAM

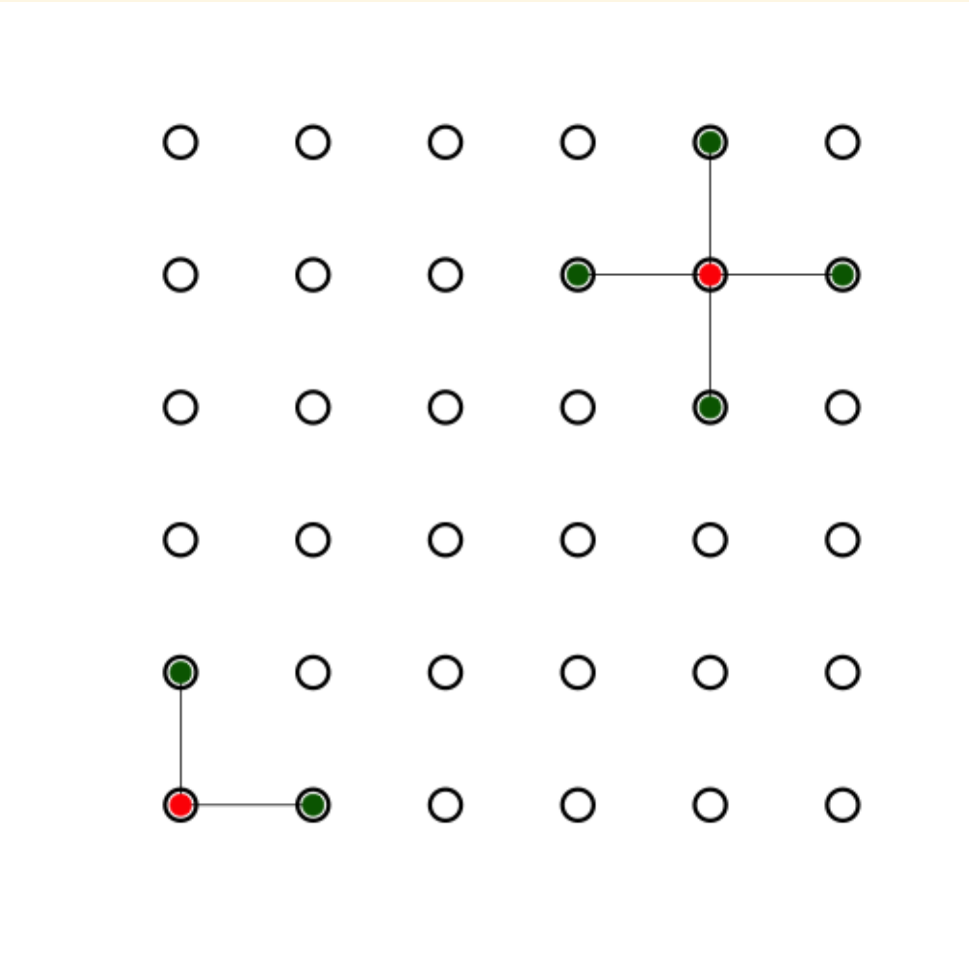
Richness ~ s(AET) +s(Temperature) + s(Elevation)



MODEL PREDICTIONS



NEIGHBOURHOODS

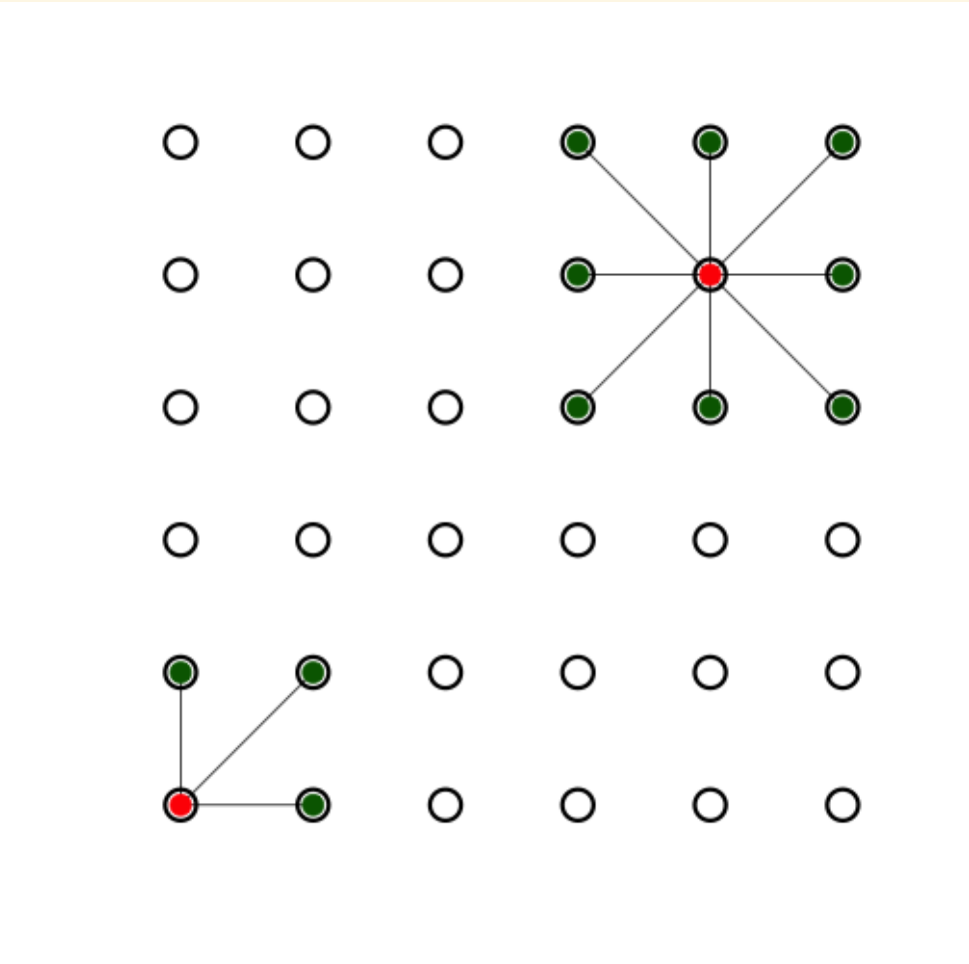


Rooks move

All cells within one step:

- vertically or
- horizontally

NEIGHBOURHOODS



Queens move

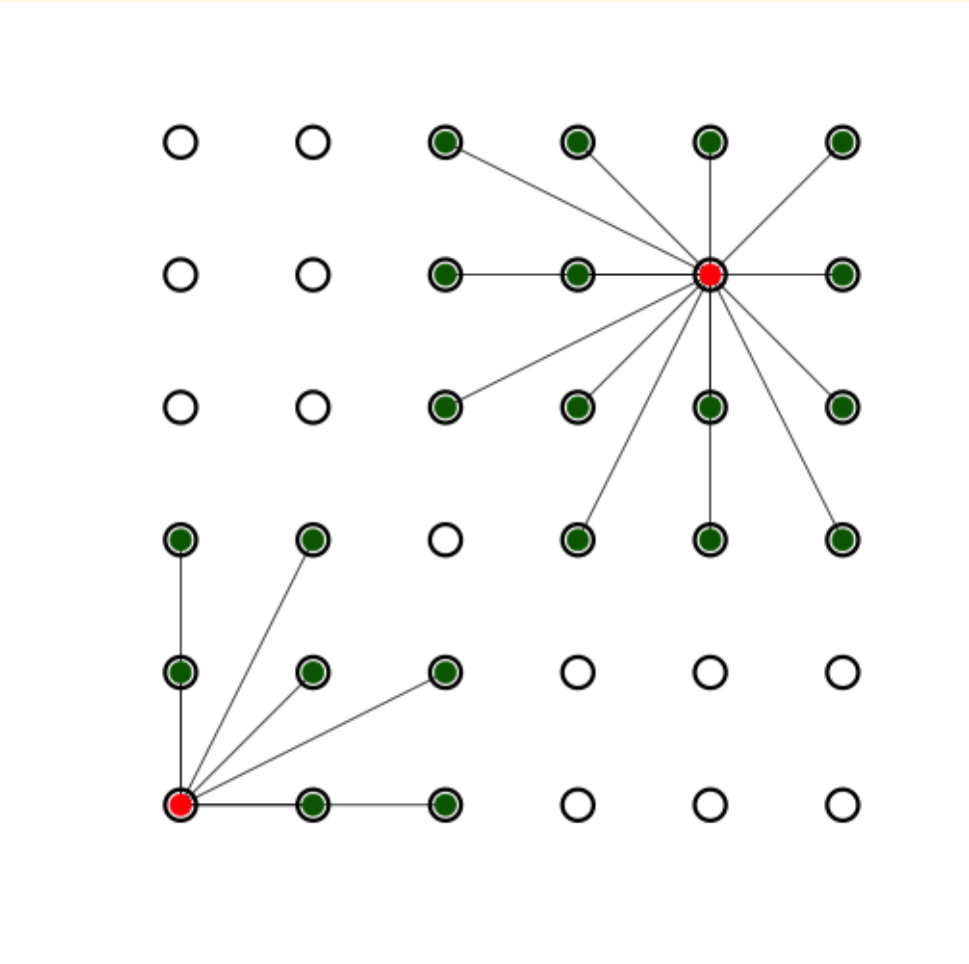
All cells within one step:

- vertically,
- horizontally

or

- diagonally

NEIGHBOURHOODS

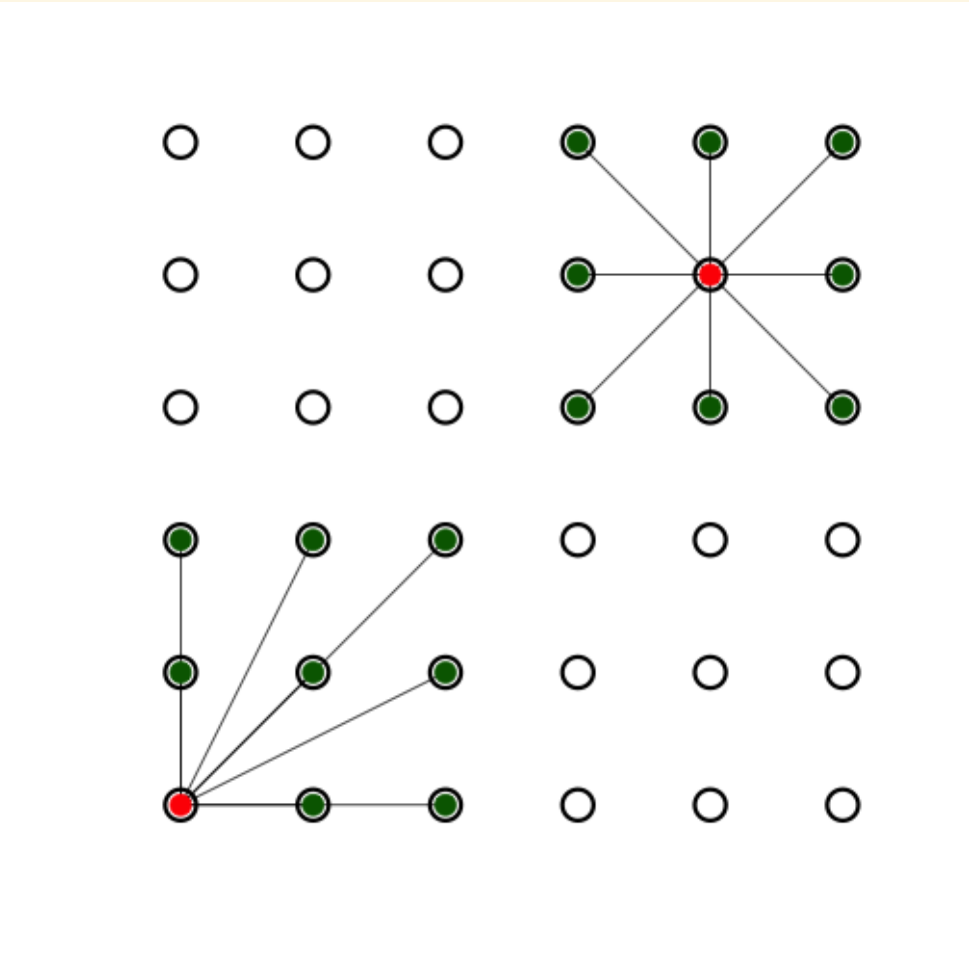


Distance based

All cells within:

- 2.4 units

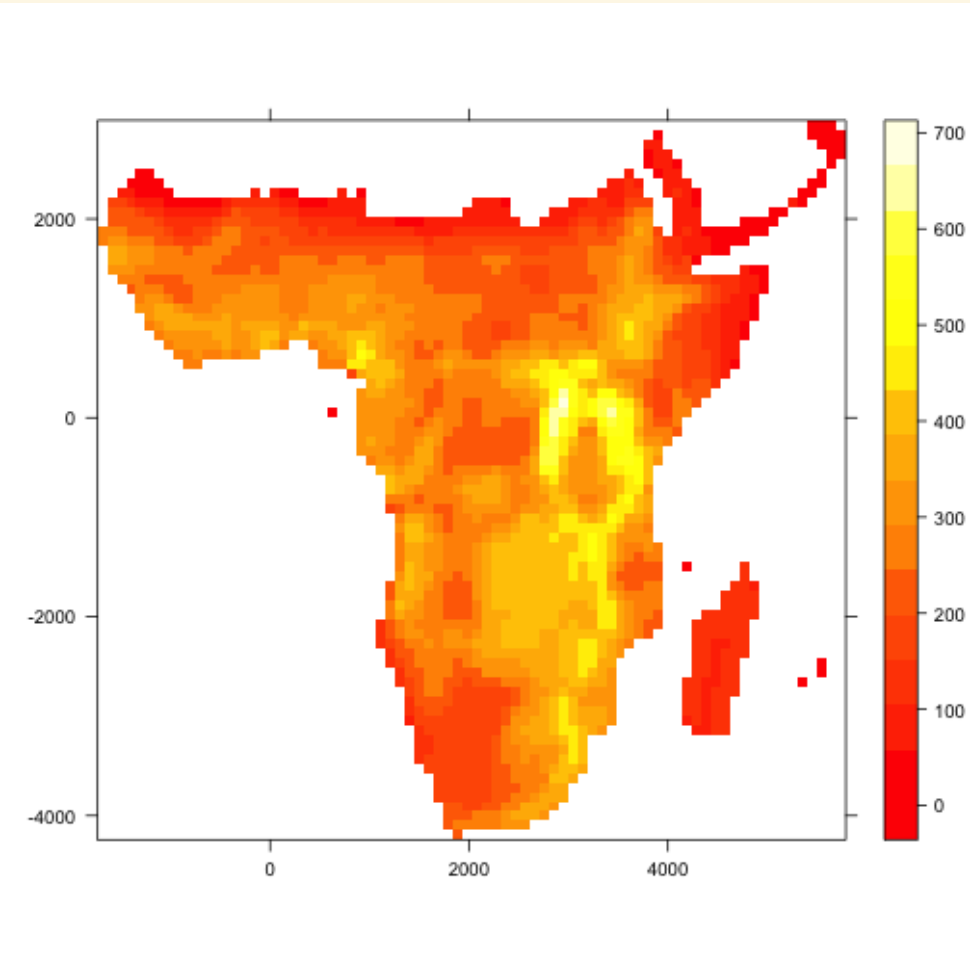
NEIGHBOURHOODS



***k* nearest**

The closest *k* cells

SPATIAL AUTOCORRELATION



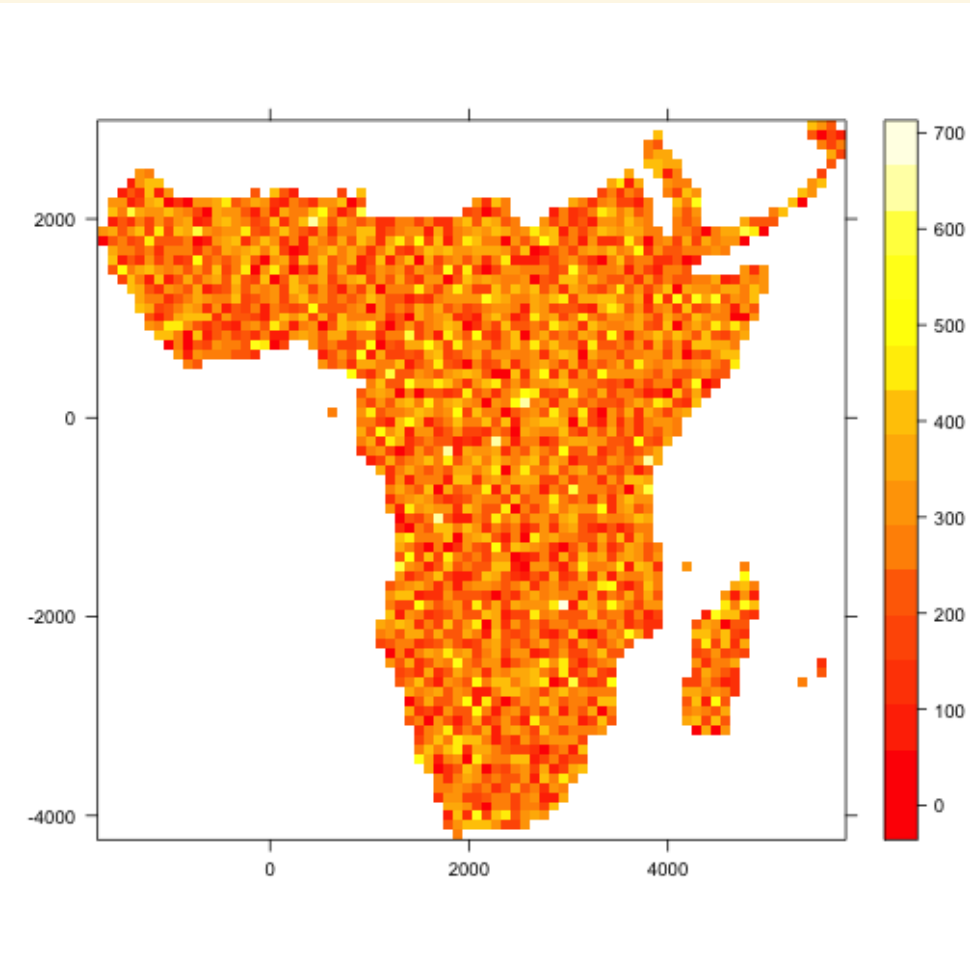
Global Moran's I

- I = 0.922
- p << 0.001

Global Geary's C

- C = 0.070
- p << 0.001

SPATIAL AUTOCORRELATION



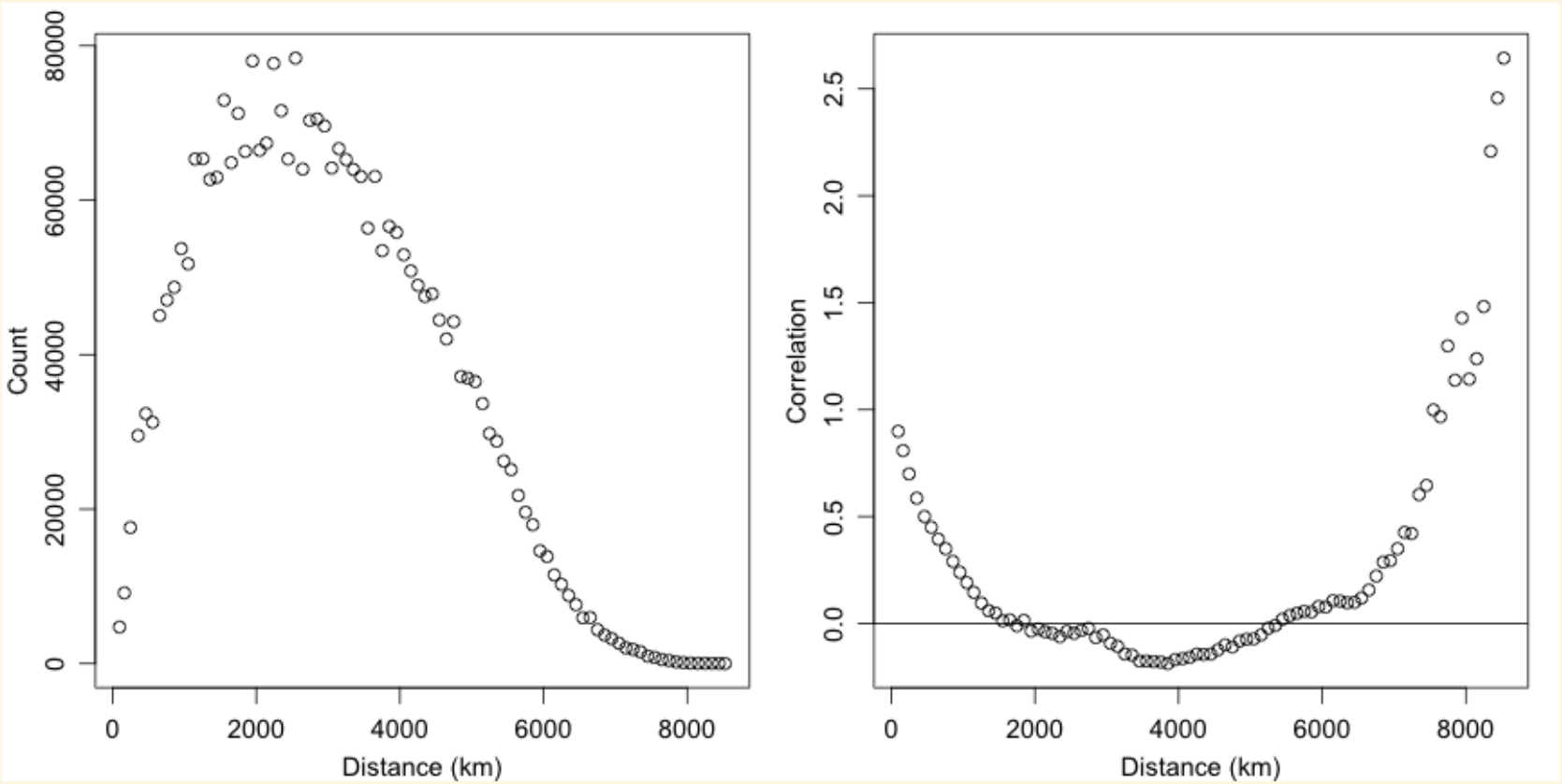
Global Moran's I

- I = -0.002
- p = 0.567

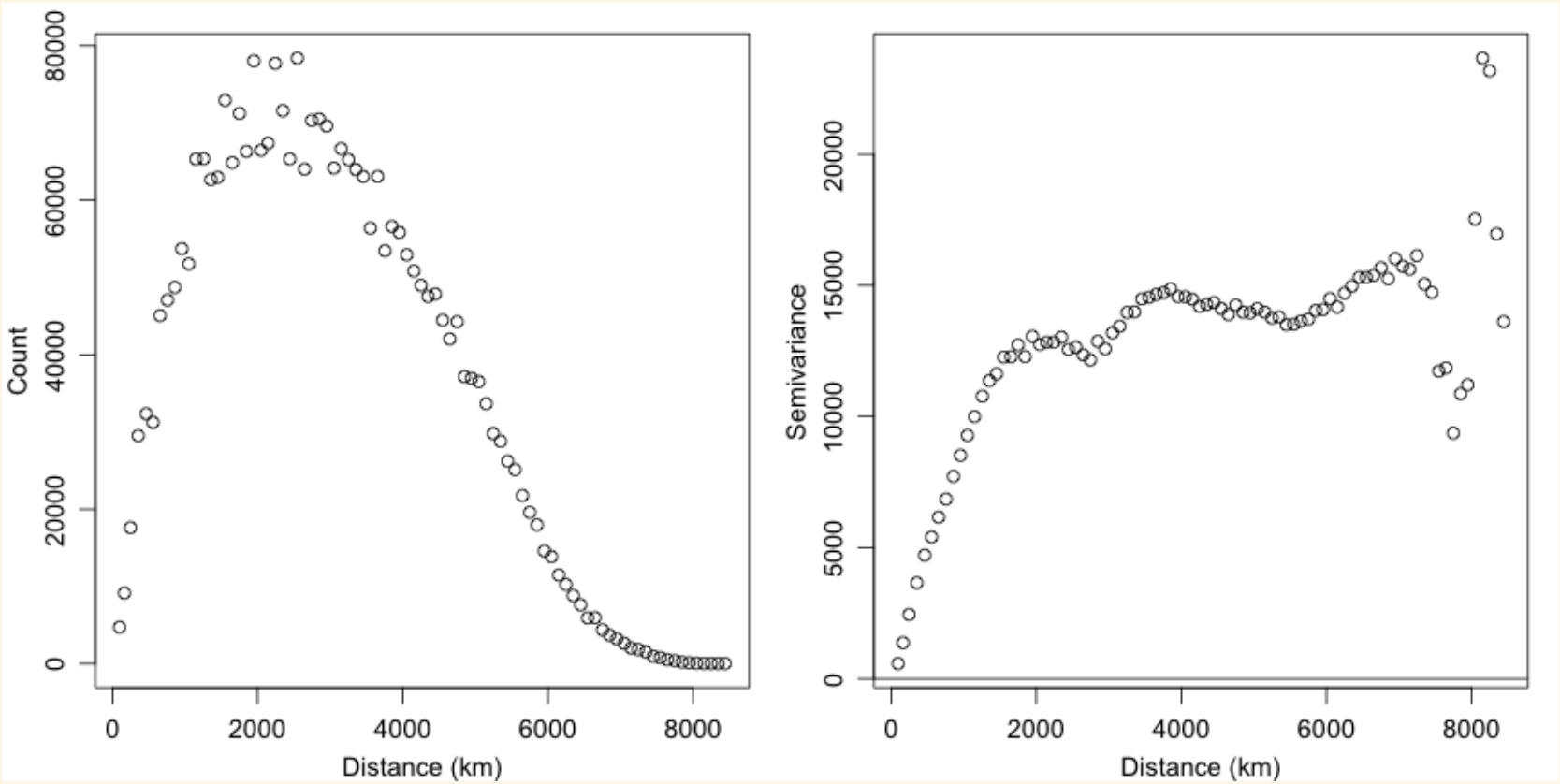
Global Geary's C

- C = 1.004
- p = 0.658

CORRELOGRAM

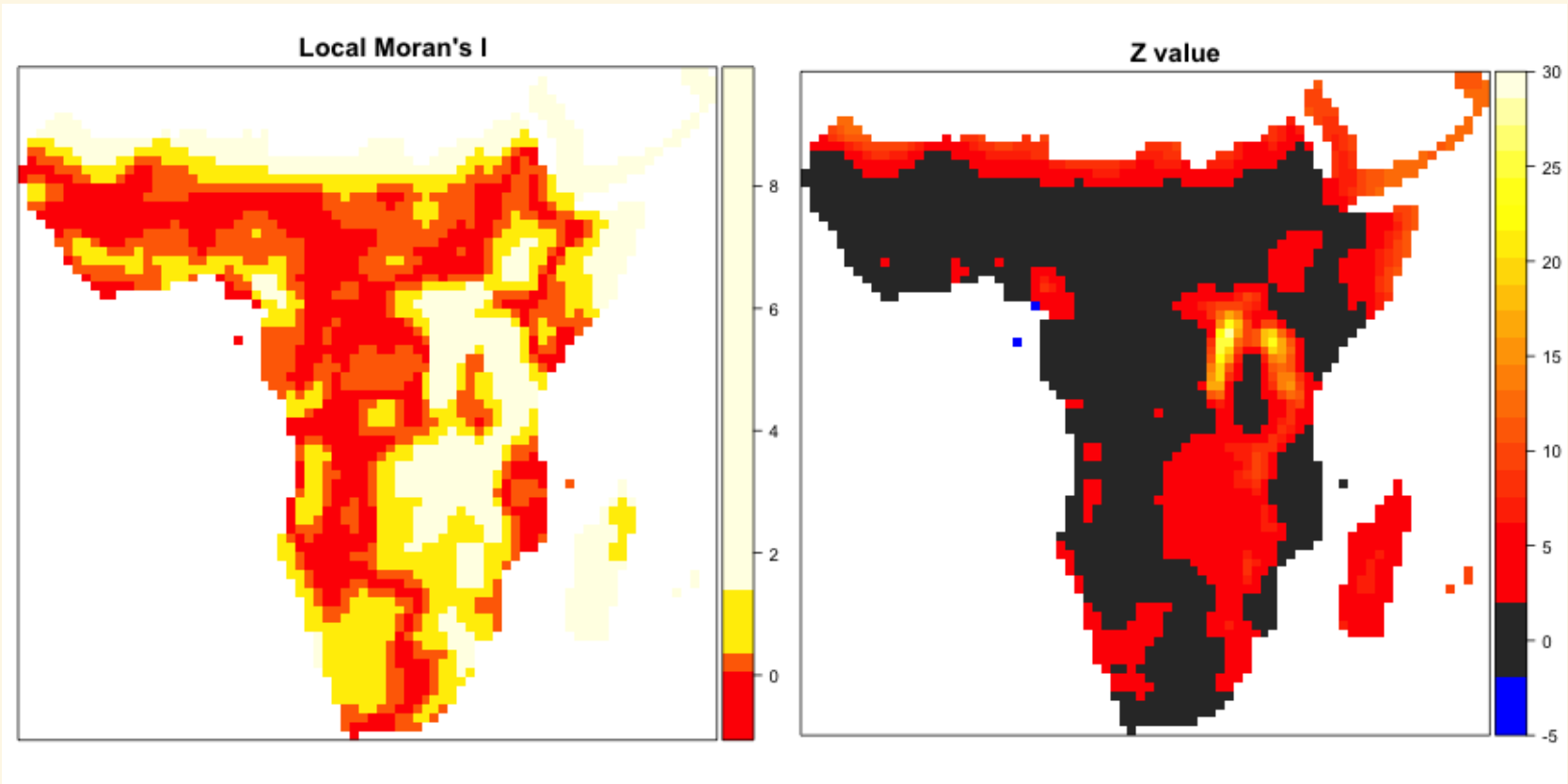


VARIOGRAM


Count
0
20000
40000
60000
80000
0
2000
4000
6000
8000
Distance (km)
Semivariance
0
5000
10000
15000
20000
0
2000
4000
6000
8000
Distance (km)

LOCAL AUTOCORRELATION

Local indicators of spatial autocorrelation (LISA)



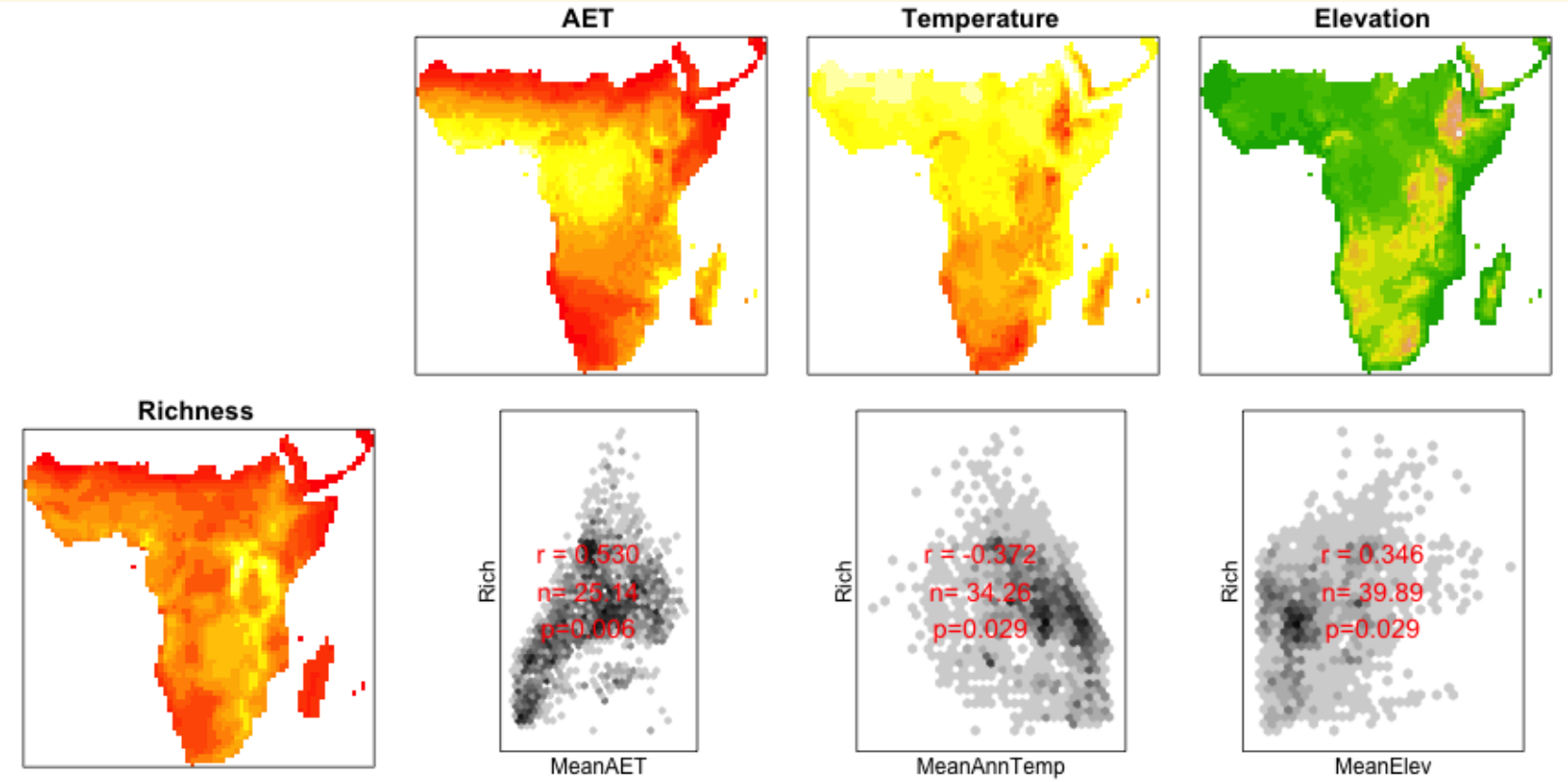
EFFECTS OF SPATIAL AUTOCORRELATION

- Data points **not independent**
- Degrees of freedom reduced:
  - **standard errors and significance testing affected**
- Not equally weighted :
  - **parameter estimation affected**

DEALING WITH SPATIAL AUTOCORRELATION

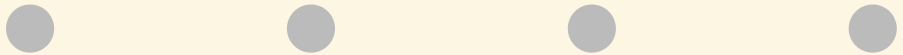
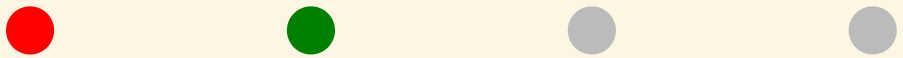
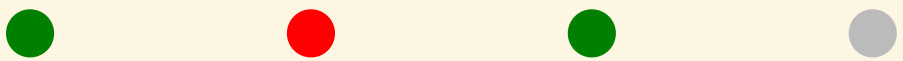
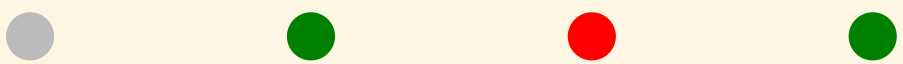
- Modify the degrees of freedom in significance testing
- Account for autocorrelation in models:
  - Simultaneous autoregressive models
  - Generalised least squares
  - Eigenvector filtering
  - Geographically weighted regression

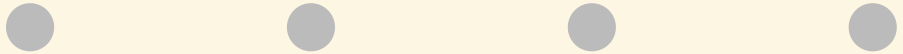
DEGREES OF FREEDOM CORRECTION



SPATIAL AUTOREGRESSION

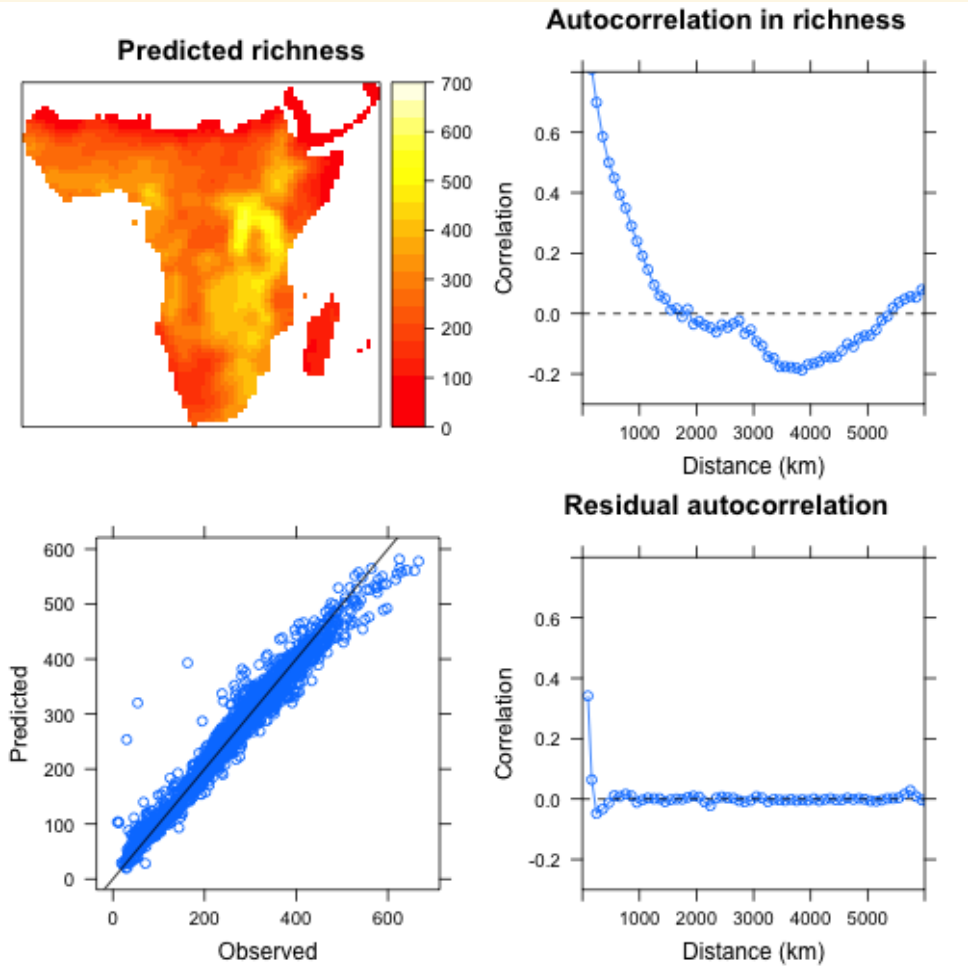
Solve for b:

	
	$bx_1 + \frac{1}{2}bx_2$
	$\frac{1}{2}bx_1 + bx_2 + \frac{1}{2}bx_3$
	$\frac{1}{2}bx_2 + bx_3 + \frac{1}{2}bx_4$



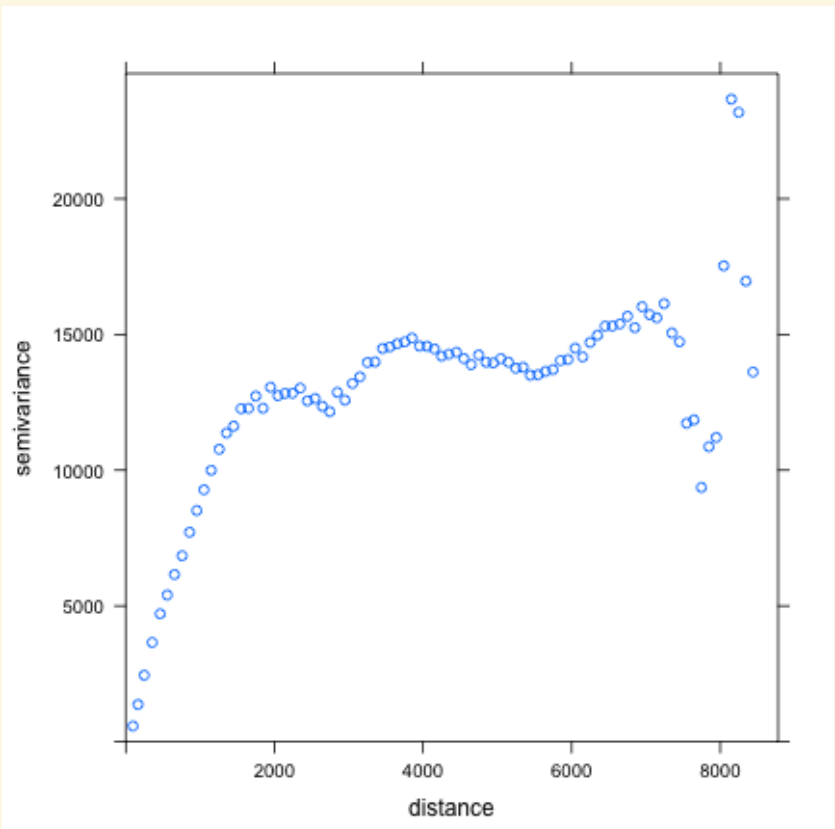
				$\frac{1}{2}bx_3 + \frac{1}{2}bx_4$
x_1	x_1	x_3	x_4	

SPATIAL AUTOREGRESSSION



GENERALISED LEAST SQUARES

```
par(mar=c(3,3,1,1), mgp=c(2,0.8,0))
plot(richVariog)
```

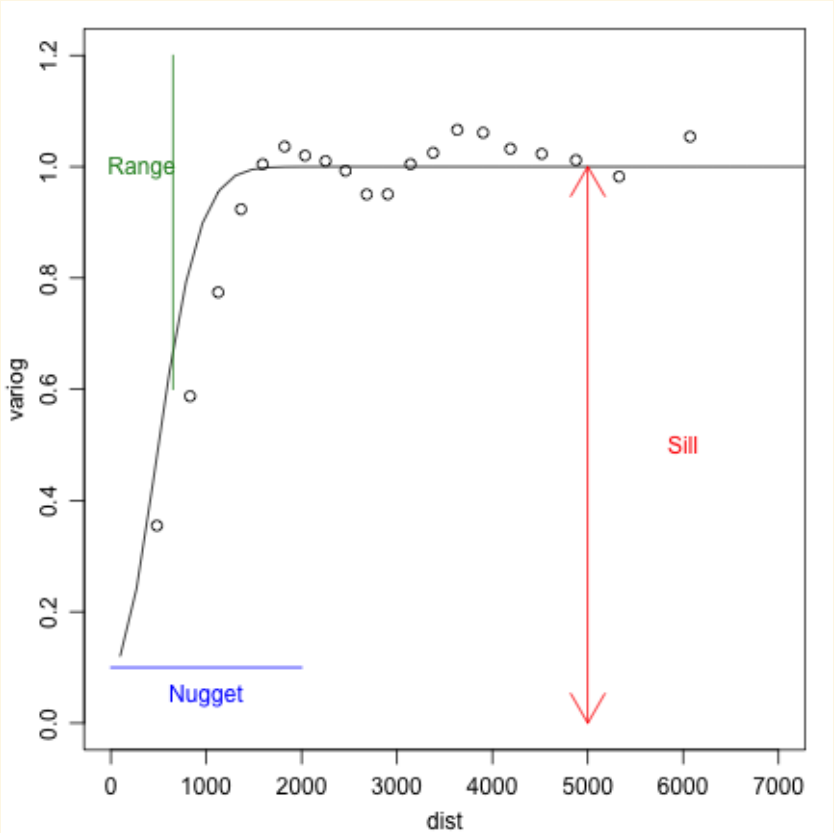


- Correlation structure
- Describe correlation as a function of distance
- Different shapes:
  - Exponential
  - Spherical
  - Linear

GENERALISED LEAST SQUARES

```
par(mar=c(3,3,1,1), mgp=c(2,0.8,0))
plot(variog ~ dist, data=glsGaussVar, xlim=c(0,7000), ylim=c(
lines(variog ~ dist, data=attr(glsGaussVar, 'modelVariog'))


arrows(0,0.1,2000,0.1, col='blue', code=0)
text(1000,0.05, 'Nugget', col='blue')
arrows(650, 0.6, 650, 1.2, col='forestgreen', code=0)
text(325, 1, 'Range', col='forestgreen')
arrows(5000, 0, 5000, 1, col='red', code=3)
text(6000, 0.5, 'Sill', col='red')
```


Range
Sill
Nugget
variog
dist
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0
1000
2000
3000
4000
5000
6000
7000

STATIONARITY AND ISOTROPY

Is the same process happening in:

- different locations (stationarity)?
- different directions (isotropy)?

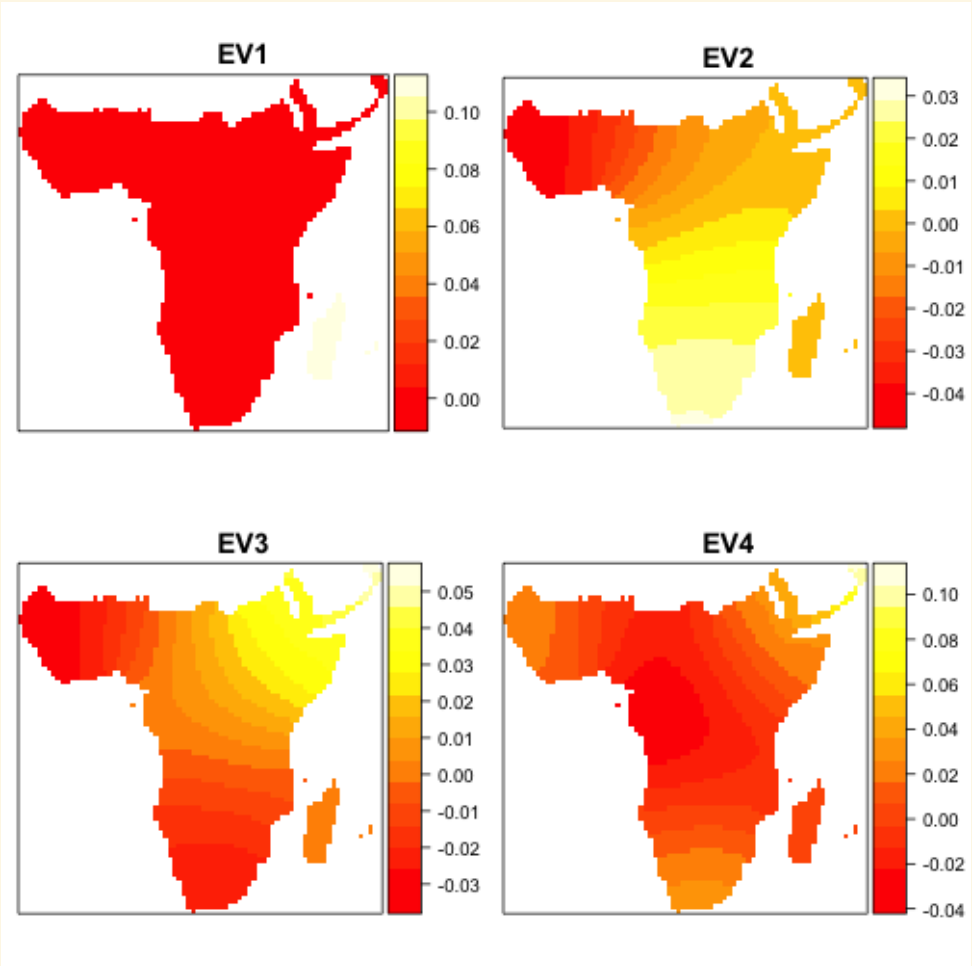
Is the problem in:

- the spatial structure of autocorrelation?
- differences in the actual relationship?

EIGENVECTOR FILTERING

- Transform a spatial weights model into a series of eigenvectors
- Use eigenvectors as variables in the model
- Use a selection process to identify and include only important eigenvectors

EIGENVECTOR FILTERING



EIGENVECTOR FILTERING

```
lm(Rich ~ MeanAET + MeanAnnTemp +
MeanElev
```

	Est	SE	t	p
(Intercept)	189.453	21.329	8.882	0
MeanAET	0.176	0.005	37.342	0
MeanAnnTemp	-4.178	0.722	-5.787	0
MeanElev	0.076	0.005	13.849	0

EIGENVECTOR FILTERING

```
lm(Rich ~ MeanAET + MeanAnnTemp +
MeanElev + Re(spEV1) + Re(spEV2) +
Re(spEV3) + Re(spEV4)
```

	Est	SE	t	p
(Intercept)	80.231	33.003	2.431	0.015
MeanAET	0.182	0.006	31.432	0.000
MeanAnnTemp	0.099	1.141	0.087	0.931
MeanElev	0.078	0.006	12.703	0.000
Re(spEV1)	-1617.625	77.641	-20.835	0.000
Re(spEV2)	963.975	129.208	7.461	0.000
Re(spEV3)	-813.557	95.868	-8.486	0.000
Re(spEV4)	-150.378	100.280	-1.500	0.134

EIGENVECTOR FILTERING

```
lm(Rich ~ MeanAET + MeanAnnTemp +
MeanElev + Re(spEV1) + Re(spEV2) +
Re(spEV3)
```

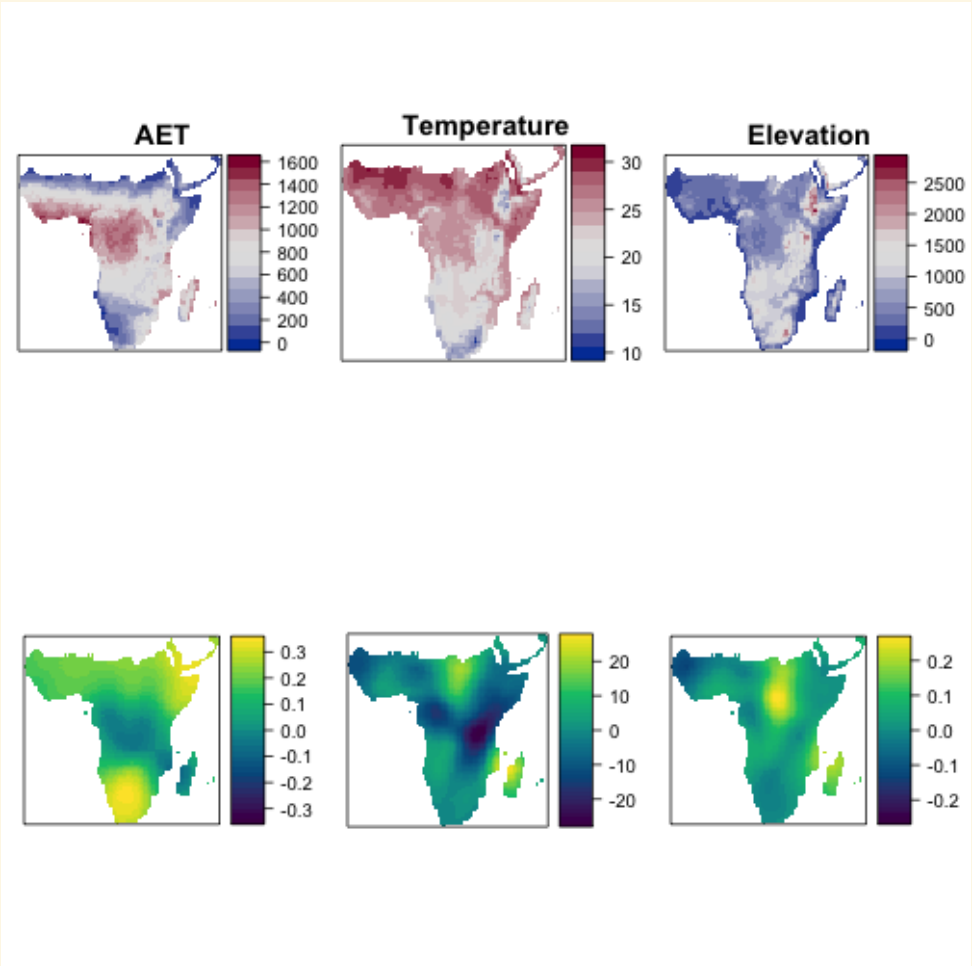
	Est	SE	t	p
(Intercept)	58.387	29.622	1.971	0.049
MeanAET	0.188	0.004	43.675	0.000
MeanAnnTemp	0.748	1.056	0.708	0.479
MeanElev	0.080	0.006	13.783	0.000
Re(spEV1)	-1610.754	77.525	-20.777	0.000
Re(spEV2)	1031.596	121.114	8.518	0.000
Re(spEV3)	-848.190	93.068	-9.114	0.000

GEOGRAPHICALLY WEIGHTED REGRESSION

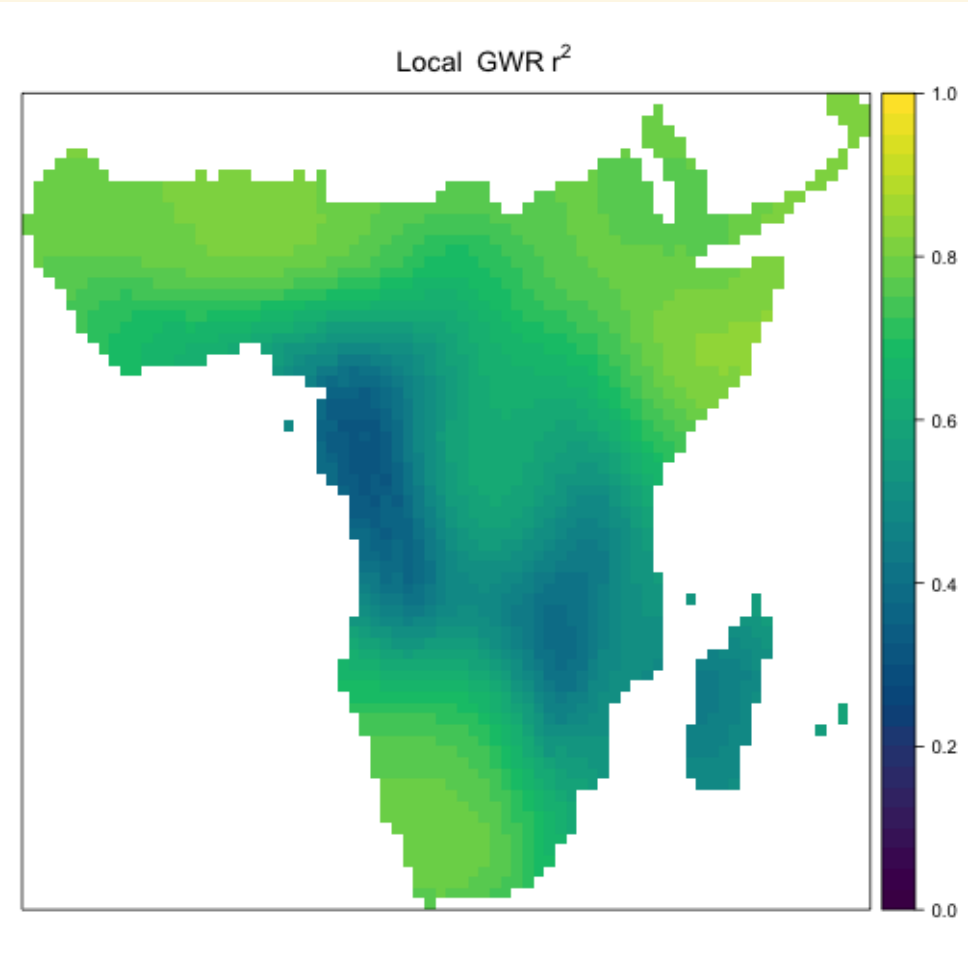
```
## Warning in gwr(Rich ~ MeanAET + MeanAnnTemp + MeanElev, data =
## 0.05, : standard errors set to NA, normalised RSS not availabl
```

```
##    user  system elapsed
## 121.872  10.945 139.256
```

GEOGRAPHICALLY WEIGHTED REGRESSION



GEOGRAPHICALLY WEIGHTED REGRESSION



PROBLEMS

- Profusion of packages: sf, sp, spdep, mgcv, ncf, gstat, nlme, spgwr
- Different data structures
- Sometimes poor documentation
- Speed of calculation (= size of dataset)
- Memory hungry
- Too many options