



How to Adjust for Spatial Autocorrelation

Jim Grace

U.S. Department of the Interior
U.S. Geological Survey

1

Citation for this specific illustration:

Cahoon, D.R., Olker, J.H., Yeates, A.G., Guntenspergen, G.R., Grace, J.B., et al., 2019, Hurricane Sandy impacts on coastal wetland resilience: U.S. Geological Survey Open-File Report 2018–1142, 117p., <https://doi.org/10.3133/ofr20181142>.

Citation for this general approach:

Bivand, RS, Pebesma, E, and Gomez-Rubio, V (2013) Applied Spatial Data Analyses with R. Springer Verlag

Notes: IP-056512; Support provided by USGS Climate & Land Use R&D and Ecosystems Programs. Formal review of this material provided by Jesse Miller and Phil Hahn. The use of trade names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Last revised 19.06.28.

Why do we care about “Spatial Autocorrelation”?

Issues:

- The raw spatial autocorrelation in data is often explained by simply including the exogenous predictors in the model. Our statistical concern is whether there remains autocorrelation among residuals.
- The reason for our concern is that standard errors, and associated probabilities, depend on an assumption of independence among residuals.
- This module illustrates how to test and correct for spatial autocorrelation when using `lavaan`.
- The principles illustrated can be easily adapted to models developed outside of `lavaan`.



2

There are a number of good references on the topic of spatial autocorrelation, e.g.:

Dale, MRT and Fortin, MJ (2014) *Spatial Analysis: A Guide for Ecologists*. Cambridge University Press

Bivand, RS, Pebesma, E, and Gomez-Rubio, V (2013) *Applied Spatial Data Analyses with R*. Springer Verlag

Bordard, D, Gillet, F, Legendre, P (2011) *Numerical Ecology with R*. Springer Verlag

What is the consequence of spatially-correlated residuals?

- According to Naroll (1961), one of the earliest pioneers in statistics, Sir Francis Galton raised the question of spatial autocorrelation in 1889. Paraphrasing,

“Positive spatial dependence can reduce the amount of information in the observations because proximate observations can partly predict each other.”

- The primary consequence of such reduced information can be expressed as a reduced “effective” sample size.
- One solution is to estimate the magnitude of residual autocorrelation, compute the effective sample size, then replace the value of n (sample size) used for calculating standard errors and p -values with the “effective n ” (described more fully on next slide).



3

A historic reference on this topic is.

Raoul Naroll (1961). "Two solutions to Galton's Problem". *Philosophy of Science* 28: 15–29. doi:10.1086/287778.

What is the consequence of spatially-correlated residuals?

- The details of the issue are, of course, potentially more complex than will be considered here. There are a large number of texts on the subject. The procedures used here were taken from:

Bivand, RS, Pebesma, E, and Gomez-Rubio, V (2013) Applied Spatial Data Analyses with R. Springer Verlag



4

A nice discussion of the general issue can be found in:

Dale, MRT and Fortin, MJ (2014) Spatial Analysis: A Guide for Ecologists. Cambridge University Press

I. The Basic Procedures:

1. Obtain model residuals for endogenous variables of concern.
2. Test for autocorrelation among residuals using Moran's I test.
3. Determine the effective sample size from Moran's I results.
4. Adjust the standard errors for model parameters affected by the autocorrelation.
5. It is also possible to adjust AICc if you are using information and multimodel comparisons (not required if using AIC as it is sample-size independent).

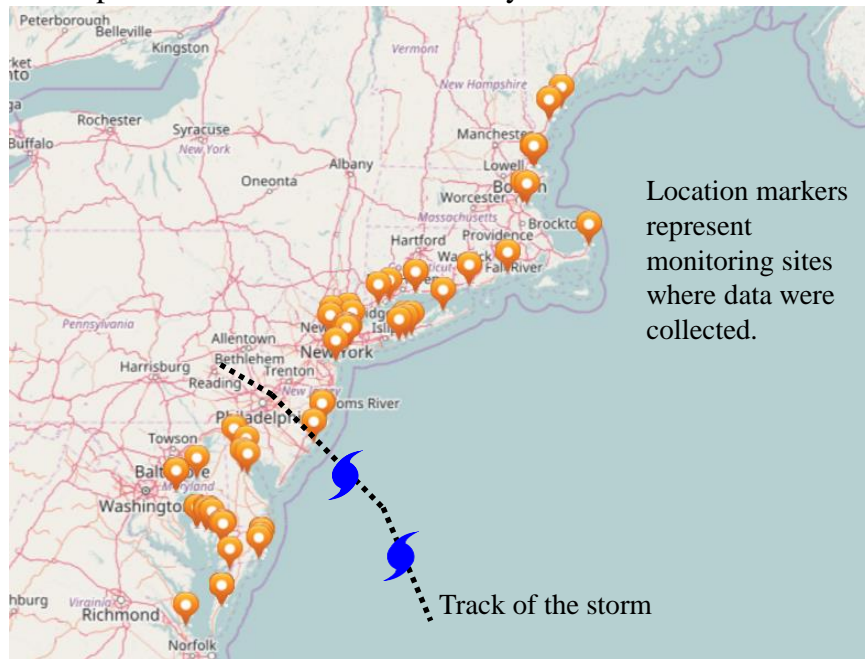


5

It is important to keep in mind that we are discussing properties of the model residuals.

An overview of the process of adjustment is given here.

Example: Effects of Hurricane Sandy on Coastal Wetlands



Slide from Alice Yeates and Glenn Guntenspergen, USGS Patuxent Wildlife Research Center shows the research situation.

Select Model(s) of Interest

There are two ways we can use estimates of effective sample size:

- (1) to adjust standard errors and p-values for a single model
- (2) to adjust AICc comparisons among multiple models.

For this abbreviated illustration, assume we have a model or models for comparison.

Example Model of Interest

```
### Model for potential adjustment
# model 1
mod1 <-lm(change ~ dist.code + I(dist.code^2) + pos + rate.c + surge,
           data = sandy.dat)

### Capture residuals
mod1.res <- resid(mod1)
```



8

We will need some model residuals to evaluate spatial autocorrelation.

Identifying Spatial Autocorrelation

Specifying Neighbor Relations

```
### Determine nearest neighbors from x, y coordinates
## Create x,y coordinate matrix
x      <- latitude
y      <- longitude
xydat <- cbind(x, y)

### Load needed library
library(spdep)

# identify k= nearest neighbors
xy.knn <- knearneigh(xydat,k=2)

# returns neighbors list
xy.nb <- knn2nb(xy.knn)
```

There are a number of ways of defining spheres of neighborhoods. Here we are examining correlations with a pre-identified number of nearest neighbors. Since the monitoring sites are spread along the coast, the neighbors will be located to either side (N or S) of one another, rather than in all directions. Most times I would choose k=4 to select the nearest neighbors, but here use k=2.

10

We need measures of spatial location. This can be used to identify the matrix of neighbor relations. The spdep library has functions that determine a list of neighbors.

Compute Moran's I

```
### Moran's test of model residuals (for mod1)
### {requires residuals and neighbors list}

MoransResults <- moran.test(change.res1, nb2listw(xy.nb,
style="W"))

print(MoransResults)
```

```
> print(MoransResults)

      Moran I test under randomisation

data:  change.res1
weights: nb2listw(xy.nb, style = "W")

Moran I statistic standard dev = 2.0619, p-value = 0.01961

sample estimates:
Moran I statistic      Expectation      Variance
0.175939865          -0.009433962        0.008082731
```



Significant residual autocorrelation is detected.

11

Moran's I index quantifies the degree of spatially-structured correlation in a dataset.

Moran, P.A.P. 1950. Notes on continuous stochastic phenomena. *Biometrika* 37:17-23.

Moran's I is defined as

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (X_i - \bar{X})(X_j - \bar{X})}{\sum_i (X_i - \bar{X})^2}$$

where N is the number of spatial units indexed by i and j ; X is the variable of interest; \bar{X} is the mean of X ; and w_{ij} is an element of a matrix of spatial weights.

The expected value of Moran's I under the null hypothesis of no spatial autocorrelation is

$$E(I) = \frac{-1}{N-1}$$

http://en.wikipedia.org/wiki/Moran's_I.

Formula for corrected sample size (Neff) is

$N_{eff} = N * ((1-I)/(1+I))$

Estimating Effective Sample Size

Estimate Effective Sample Size

```
### Computation of effective sample size
n = 107
eff_ss = round(n/(1+abs(MoransResults$estimate[1])))

print(eff_ss)
```

```
> ### Computation of effective sample size
> 91
```

Full sample size for this example is 107.
Effective sample size is 91.



13

There is a direct connection between Moran's I and effective sample size, as shown in this equation.

As given back in the notes to slide 11,

Formula for corrected sample size (Neff) is

$$N_{eff} = N * ((1 - I) / (1 + I))$$

where I = Moran's I estimated value.

Adjusting Standard Errors and p-values

The Adjustment Formula

The formula below can be used to adjust the standard errors associated with model parameters for the reduced effective sample size.

(Note we intuitively expect standard errors to broaden with reduced sample size.)

$$SE_{adj} = SE \times \sqrt{\frac{N}{N_{adj}}}$$

where:

SE_{adj} is adjusted standard error,

SE is original standard error,

N_{adj} is the effective sample size, and

N is the observed sample size.



15

A basic take home from this formula is that it takes a big drop in effective sample size, resulting from a large amount of residual spatial autocorrelation, to make a big change in standard errors.

The effect = square root of the proportional overstatement of N .

Adjusting Standard Errors (part 1)



```
### Adjust standard errors of model (mod1)
n = 107; n.adj = 91

# Formula for adjustment
se.adj = sqrt(n/n.adj)

# Model
mod1 <- lm(change ~ dist.code +I(dist.code^2) +pos +rate.c
           +surge, data=sem.vars)

# Extract standard errors
std.errors <- coef(summary(mod1))[, 2]

print(std.errors)
```

16

First step in the adjustment process, shown here, involves getting some of the raw materials needed for the calculation:

- adjustment index
- unadjusted standard errors

Adjusting Standard Errors (part 2)



```
# Adjust standard errors
se1.adj = std.errors[1]*se.adj;
se2.adj = std.errors[2]*se.adj;
se3.adj = std.errors[3]*se.adj;
se4.adj = std.errors[4]*se.adj;
se5.adj = std.errors[5]*se.adj;
se6.adj = std.errors[6]*se.adj;

# Print list adjusted standard errors
std.errors.adjusted <- data.frame(se1.adj, se2.adj, se3.adj,
se4.adj, se5.adj, se6.adj)
std.errors.adj.t <- t(std.errors.adjusted)
print(std.errors.adj.t)
```

17

This is just basic computational code.

Adjusted Standard Errors

```
> print(std.errors.adj.t)
```

```
se1.adj    9.8787912  
se2.adj   61.2413906  
se3.adj   93.7365557  
se4.adj    3.3908228  
se5.adj    0.5107856  
se6.adj    3.3763334
```



18

And here are adjusted standard errors, which will be used in next slide.

Illustration of Adjusting p-values

```
### Calculate adjusted p-values for mod1
# get parameter estimates
mod1.ests <- coef(summary(mod1))[, 1] ; mod1.ests
mod1.est1 <- mod1.ests[1]

# compute adjusted p-value for intercept
sel.adj = 9.8787912 #from previous step
n.adj = 91
t1.adj = mod1.est1/sel.adj;
p1.adj <- 2*pt(-abs(t1.adj),df=n.adj-1); round(p1.adj, 4)

# compare to unadjusted (just for information sake)
round(coef(summary(mod1))[, 4], 4)
# proportional increase in p-value
0.9869/0.9858 #very very slight affect

> p1.adj <- 2*pt(-abs(t1.adj),df=n.adj-1); round(p1.adj, 4)
0.9869

> # proportional increase in p-value
> 0.9869/0.9858 #very very slight affect
[1] 1.001116
```



Note that the difference in p-value in this case is very small. ¹⁹

The key here is the use of the base function “pt”, which returns the p-value associated with a particular t-index and sample size.

Adjusting AICc Comparisons

What if we are relying on AICc model comparisons to draw our inferences about what model to select for interpretation?

Competing Models

```
### Competing Models
# Model 1
mod1 <-lm(change ~ dist.code + I(dist.code^2) + pos + rate.c
          + surge, data = sandy.dat)

# Model 2 - drop position variable
mod2 <-lm(change ~ dist.code + I(dist.code^2)          + rate.c
          + surge, data = sandy.dat)
```

21

Here I just show a model with an important term dropped, which will be compared to our base model.

Note that to be extra careful, we could independent estimate the effective sample size for each of the models. Here I am not going to the trouble because it will not affect our model choice.

AICc Comparison Table

```
# AICc table
library(AICcmodavg)

# compare models using AIC
aictab(list(mod1, mod2), c("Model 1", "Model 2"))
```

Model selection based on AICc:

	K	AICc	Delta_AICc	AICcWt	Cum.Wt	LL
Model 1	7	844.72	0	0.98	0.98	-414.79
Model 2	6	852.72	8	0.02	1.00	-419.94

22

*Marc J. Mazerolle (2016) AICcmodavg: Model selection and multimodel inference based on (Q)AIC(c). R package version 2.0-4.
<http://CRAN.R-project.org/package=AICcmodavg>.

The classic AIC is asymptotically unbiased. Thus, it makes sense to adjust for sample size in small samples (=AICc).

When the ratio of parameters to samples is large (i.e., information is low), we use

$$AICc = AIC + \left(\frac{2q(q+1)}{n-q-1} \right)$$

where q = number of estimated parameters in the model and
 n = the number of samples



23

For moderate to small samples (250 samples or less), the AICc seems like a good choice for model selection. This of course depends on the complexity of the model being examined.

With the AICc, we adjust the AIC for the ratio of information/samples to parameters in the model. This is a reasonable suggestion because it takes information to estimate parameters and the ratio of information to parameters is a handy way to discuss sample size recommendations, though such things are truly not simple.

Anyway, the AICc has a more complex parsimony correction term than does the AIC (which is just $2q$), as shown in the slide.

(Note: AICc not theoretically defined for models with multivariate responses, included latent-variable models. It may be useful nonetheless.)

Adjustment of AICc

$$AICc = AIC + \left(\frac{2q(q+1)}{n-q-1} \right)$$

```
### Adjustment
# Extract AIC values (these are not influenced by sample size)
mod1.AIC <- AIC(mod1)
mod2.AIC <- AIC(mod2)

# Bring in adjusted sample size estimate
n.adj = 91

# Extract number of parameters (K) in model from AICc table
mod1.num.params <- aictab(list(mod1), "Model 1")$K #K1 = 7
mod2.num.params <- aictab(list(mod2), "Model 2")$K #k2 = 6

# Compute adjusted AICc
mod1.AICc.adj <- mod1.AIC + ((2*mod1.num.params*(mod1.num.params+1))
                             / (n.adj - mod1.num.params-1))

mod2.AICc.adj <- mod2.AIC + ((2*mod2.num.params*(mod2.num.params+1))
                             / (n.adj - mod2.num.params-1))

# Compute AICc difference (model 1 minus model 2)
mod1.AICc.adj - mod2.AICc.adj

> # Compute AICc difference (model 1 minus model 2)
> mod1.AICc.adj - mod2.AICc.adj
[1] -7.939585
```

This code just implements the formula for a new AICc based on a reduced value of n.